Quantum computing with Rydberg atoms : control and modelling for quantum simulation and practical algorithms

Lucas Leclerc

université PARIS-SACLAY

Quantum computing with Rydberg atoms: control and modelling for quantum simulation and practical algorithms

Calcul quantique avec des atomes de Rydberg: contrôle et modélisation pour simulation et algorithmes quantiques

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 572, Ondes et Matière (EDOM)
Spécialité de doctorat : Physique
Graduate School : Physique
Référent : Institut d'Optique

Thèse préparée dans l'unité de recherche **Laboratoire Charles Fabry** (IOGS, Université Paris-Saclay, CNRS), sous la direction de **Thierry LAHAYE**, Directeur de recherche, et le co-encadrement de **Loïc HENRIET**, *co-CEO*, Pasqal

Thèse soutenue à Paris-Saclay, le 19 Septembre 2024, par

Lucas LECLERC

Composition du jury

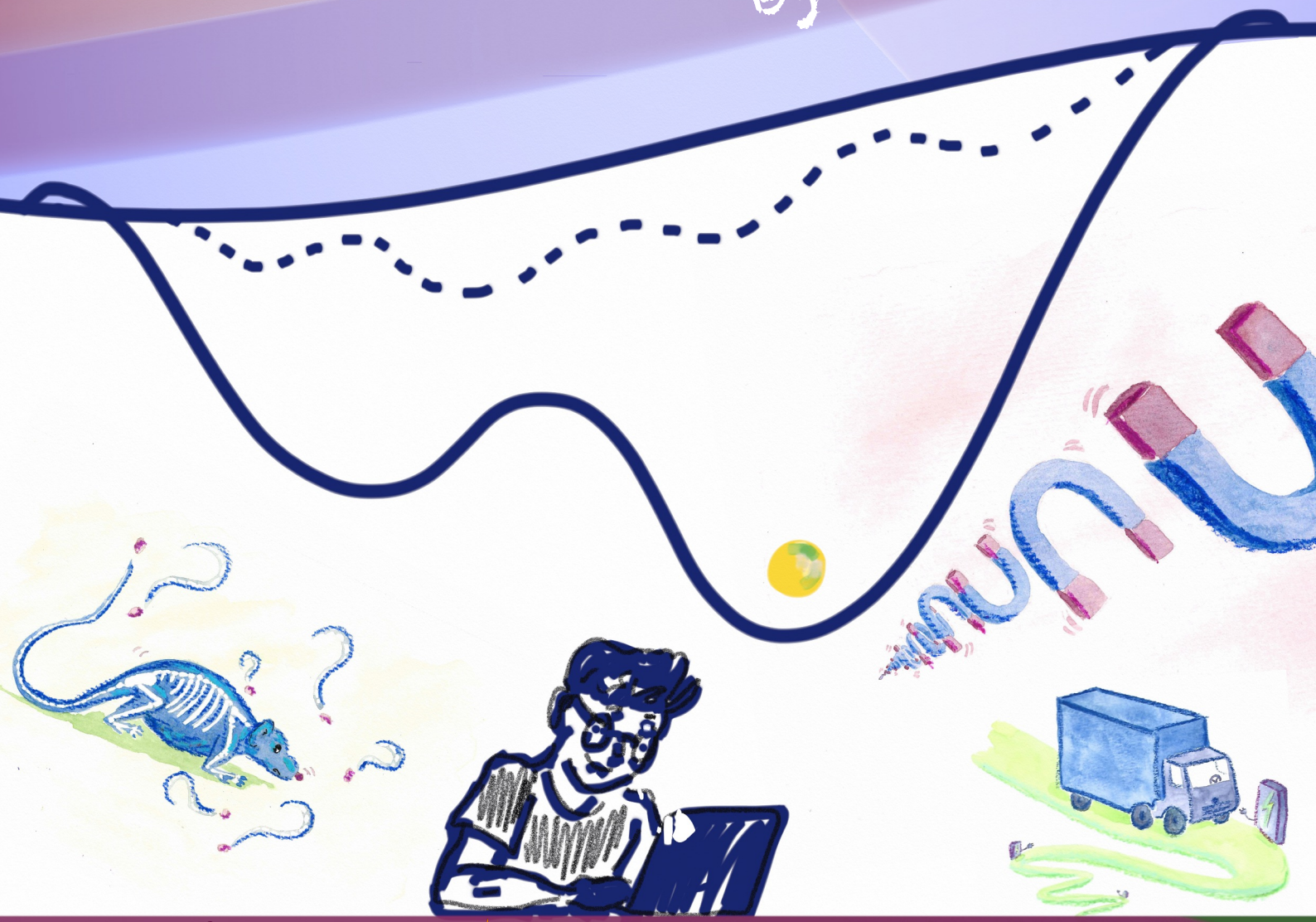
Membres du jury avec voix délibérative

Valentina PARIGI Professeur, Sorbonne Université	Présidente
Benoît VERMERSCH Maître de conférences, Université Grenoble-Alpes	Rapporteur & Examinateur
Guido PUPILLO Professeur des universités, Université de Strasbourg	Rapporteur & Examinateur
Patrick CHEINET Chargé de recherche, Université Paris-Saclay	Examinateur

THESE DE DOCTORAT
NNT : 2024UPASP046

Quantum computing with Rydberg atoms: control and modelling for quantum simulation and practical algorithms

By Lucas Leclerc



université PARIS-SACLAY · cnrs · INSTITUT d'OPTIQUE GRADUATE SCHOOL Paristech · LABORATOIRE CHARLES FABRY · Pasqal

Avant-propos et remerciements

Ce manuscrit présente mes travaux de thèse réalisés entre novembre 2021 et avril 2024 au sein du Laboratoire Charles Fabry de l'Institut d'Optique et de la start-up Pasqal. En écrivant ces lignes qui représentent l'aboutissement de plusieurs années de réflexions et de doutes, d'essais et de ratés, de réussites (parfois) et de ces moments de grâce où tout s'aligne et qui font tout le sel de la recherche, je me dis que tout passe toujours trop vite quand on s'amuse. Une thèse ne se résume pas qu'aux projets entrepris et aux articles publiés. Cette plongée au cœur du monde de la recherche m'a enseigné qu'une conférence sur le calcul adiabatique peut se transformer en road trip déjanté au Nouveau Mexique, qu'une nuit de rush passée au labo peut faire naître une passion pour les figures bien faites, et qu'une implémentation réussie requiert 20% de science, 20% de chance et 120% de bonne communication entre collègues théoriciens et expérimentateurs. Même un pavé de 250 pages, agrémenté d'une centaine de figures, semble insuffisant pour rendre compte de toute cette aventure scientifique, remplie de défis techniques à vous garder éveillé jusqu'à l'aube ou de discussions envoutantes de fin de soirées avec des chercheurs inconnus rencontrés à Madrid, à Detroit ou aux fin fonds des Alpes mais instantanément reliés par cette passion pour la science et cette dépendance au café.

Toute cette aventure n'aurait jamais été possible sans la confiance de Georges-Olivier Reymond et d'Antoine Browaeys, deux exemples inspirants de scientifiques à la fois entrepreneur et chercheur. Merci à vous deux pour cette belle opportunité ! Si j'ai pu naviguer avec assurance pendant ces trois années, c'est surtout grâce aux conseils ininterrompus de mes deux encadrants que je tiens à remercier particulièrement. Thierry, tu as été un directeur hors-pair, toujours présent pour m'introduire aux ficelles du métier de chercheur, pour me pousser à toujours plus de rigueur scientifique et pour répondre à toutes mes questions naïves de physique atomique. Merci pour ton attention, pour ton franc-parler et tes inestimables conseils de vieux d'la vieille. Loïc, tu as été un superviseur extraordinaire, du début de mon stage où tu me promettais que je pourrais tester du contrôle optimal sur une machine de Pasqal (promesse tenue de peu !) à cette fin de thèse où j'éprouve une immense fierté à te voir co-CEO de la boîte désormais (on l'a toujours dit avec Constantin). J'espère qu'un jour tu me révéleras le secret pour, malgré un emploi du temps constamment surbooké, réussir à encadrer simultanément trois thésards et à toujours te soucier de résoudre nos problèmes. Merci d'être un modèle de physicien passionné et exigeant qui nous pousse à toujours faire mieux, merci pour la précieuse indépendance et les responsabilités que j'ai pu avoir à Pasqal grâce à ta confiance et enfin merci de m'avoir appris à profiter de la thèse comme il se doit.

Je tiens aussi à remercier chaleureusement l'ensemble de mes collaborateurs à Pasqal. En quatre ans, cette petite boîte d'une dizaine de personnes réparties dans deux salles de l'IOGS est devenue un formidable fer de lance de l'informatique quantique français, fort de plus de 260 passionnés. La liste de remerciements risque d'être longue alors autant commencer par mon partner-in-crime, Constantin, désormais Dr. Dalyac. De nos discussions peu sérieuses dans la cave d'un hôtel particulier à notre tentative d'effraction dans le laboratoire militaire de Los Alamos (à ça de réussir), une thèse sans toi aurait été beaucoup trop calme. Un énorme merci ! Merci à Henrique pour ses innombrables

conseils de code, à Sebastian pour son soutien dans le projet CACIB et à Louis pour son enthousiasme à tout épreuve. Merci à l'équipe Iroise, Boris et Luiz, pour avoir sélectionné un à un les atomes du projet QEK, et merci à l'équipe Fresnel, notamment Guillaume, Clémence et Lucas, pour votre travail incroyable qui me permet désormais de sélectionner les atomes un par un moi-même. Merci aussi à l'équipe Cloud, notamment, Matthieu, Anne-Claire et Mourad, pour avoir créé une plateforme qui fonctionne (souvent) incroyablement bien ! Je tenais à remercier particulièrement les membres de l'équipe Protocols qui incarne quotidiennement cette volonté de relier théorie et hardware sous la houlette de Lucas Lassablière. Merci d'être toujours à notre écoute et de t'être assuré de ma survie ces derniers temps ! Un énorme merci à Julius pour sa relecture rapide et détaillée (impossible mais vrai) de ma thèse : je ne sais pas ce qu'on ferait sans toi. Merci aussi à toutes les personnes qui ont été de précieux conseils, notamment Adrien, Bruno, Pascal, Davide, Alexandre, Louis Paul, Mauro, Kemal. Enfin un grand merci à mes co-thésards Slimane, Elie et Grégoire et courage pour la fin de la thèse !

J'ai eu l'honneur durant ces années de collaborer avec une belle bande d'expérimentateurs, la célèbre équipe CHADOQ de l'IOGS. Guillaume, Gabriel and Cheng, votre dévouement à la physique est un modèle pour moi, merci pour toutes ces magnifiques données à expliquer. Je veux aussi remercier Britton, Damien, Sarah, Bastien et Igor pour les bons moments passés dans et en dehors du labo !

Je remercie chaleureusement les chercheurs membres du jury, Patrick Cheinet, Valentina Parigi, Guido Pupillo et Benoit Vermersch, qui ont pris le temps de lire ce manuscrit et d'échanger avec moi lors de la soutenance.

Enfin, un immense merci à toutes les personnes qui m'entourent au quotidien ! Merci à tous mes amis qui supportent depuis et pour longtemps encore mes élucubrations de particules quantiques résolvant nos problèmes du quotidien. Philibert, Romain, Lucie, la Liane, le Laitage et tous les autres, merci pour votre présence et votre écoute ! Et bien entendu, merci à Amance et Lova pour cette magnifique couverture ! Pour conclure, je voudrais dédier cette thèse à deux personnes sans lesquelles rien de tout cela n'aurait été possible. A ma mère, tout simplement, qui m'a toujours soutenu dans tous mes projets ! Et à mon grand-père, qui a su partager avec moi son amour de la science sur cette plage de Menton, une équation après l'autre.

Table of Contents

Introduction

Refining our understanding of an unknown system through modelling lays the groundwork for being able to optimally control it and opens the door to a myriad of potential applications, exploiting the once enigmatic and unpredictable effects of this now known system. For instance, quantum systems evolve according to principles seemingly magical or at least counter intuitive when studied in a classical physics framework. Adopting properties from both waves and matter or being able to exist in several states simultaneously, quantum particles can even communicate non-locally between entangled pairs, as experimentally demonstrated for the first time by [1]. By formalising the behaviour of ensemble of particles, such as atoms or photons, quantum theory led to a technological revolution, first enabling in the mid 20^{th} century the development of transistors and lasers, indispensable components of today computers, smartphones, satellite navigation systems and advanced medical diagnostic tools. With ever more advanced experimental control enabling their individual manipulation, quantum systems can improve the accuracy of sensors by order of magnitude, encrypt data in unbreakable ways for secure communication and hold the promise of performing computing tasks that are beyond the scope of conventional computers. While this last hope stems from the exponential complexity required to depict the state of a many-body quantum system, intractable for classical machines, the potential of these highly complex quantum computers may remain untapped if not well operated or if they fail to address pertinent challenges. Among the various quantum systems which can serve as building components of this new class of computers, trapped ions and superconducting circuits are prone contenders with their high fidelity of operations and their fast computing rate. In the recent years, multiple technologies with various advantages have also emerged, making use of photons, polar molecules, nitrogen-vacancy centres, topological particles or atoms trapped in optical lattices, or in tweezers.

This thesis focuses on the latter technology, namely the use of neutral atoms trapped in tweezers and excited to Rydberg states. The wide tunability of the implemented Hamiltonians combined with the geometrical versatility of the platform allows us to study a large variety of many-body phenomena in quantum simulation. We also address the challenges of operating and developing algorithms on current Rydberg-based platforms in the noisy intermediate scale quantum (NISQ) era [2]. Successfully implementing protocols requires a solid grasp of the technical possibilities and constraints of the technology which determines the range of experiments we can conduct. Throughout this thesis, we build several models of Rydberg-based quantum processors, enabling us to emulate the dynamics of large-size systems using tensor networks or to sequentially incorporate various low-level physical effects occurring on the experimental setups to accurately assess their detrimental impact on high-level protocols. By identifying the dominating contributions, we can address the need for experimental corrections or create mitigation techniques using optimal control frameworks, notably for designing faster and more efficient state preparation protocols or entangling quantum gates. These classical tools are of great help in the various applications reachable by analog Rydberg platforms. We illustrate this assertion with enhanced experimental preparation of the antiferromagnetic ground state of the Ising model, a stepping stone for studying many-body localisation phenomena. For benchmark purposes,

we construct digital twins of a noisy processor probing continuous symmetry breaking in the dipolar XY model and developing local rotations useful for quantum state tomography. Then, targeting industry-relevant problems, we improve a variational quantum adiabatic algorithm with a classical Bayesian optimiser to experimentally solve Maximum Independent Set problems on a family of unit disk graphs representing smart charging tasks. We also test the performance of an implemented quantum feature map incorporated in a machine learning classification algorithm used for predicting the toxicity of compounds. Finally, we make the most out of limited quantum resources to design an hybrid quantum-enhanced classifier for fallen angels detection in a financial risk management use case. This thesis demonstrates the versatility of problems that can be tackled by Rydberg-based processors while emphasising the need for accurate modelling and optimised control protocols.

The experimental results gathered during the course of this thesis have been obtained on three distinct quantum processors, two of them, Iroise and Fresnel, belonging to Pasqal and the last one, Chadoq, belonging to the Quantum Optics group at IOGS. Technical possibilities and constraints can thus vary from one chapter to another but for each implementation, the device used will be mentioned. It is also worth noting that Fresnel, on the contrary to both Chadoq and Pasqal prototype, Iroise, was directly accessible through a cloud interface. It was thus possible to perform complex protocols without the need for any manual intervention from experimentalists.

The manuscript is structured as follows.
In chapter 1, we describe the technology enabling to individually trap and control atoms, as well as making them interact by promoting them to Rydberg states. Analysing the various steps needed to perform quantum computation with Rydberg atoms, we also list the potential sources of error impacting each of them.
In chapter 2, we then present three ways of modelling quantum systems, with a focus on Rydberg atoms, in order to be able to classically emulate their dynamics. Balancing the use of these distinct tools allows to predict properties emerging at large sizes or to precisely quantify the detrimental contributions of the experimental noises by building digital twin of the quantum computer under study.
In chapter 3, we use the formalism of optimal control theory to first devise robust protocols for operating optimally noisy quantum computers using the models built in the previous chapter. Second, in order to efficiently tackle more complex tasks, we outline the components of variational quantum algorithms and introduce the use of the Bayesian optimisation technique as a powerful classical optimiser. These emulation and optimisation classical tools are then used throughout the next chapters to either better understand experimental limitations or enhance performance of the various protocols presented.
For instance, in chapter 4, delving in the quantum simulation of the Ising model with Rydberg atoms, we optimise with a variational algorithm an adiabatic schedule to enhance the preparation of the antiferromagnetic ground state of the model directly on the quantum processor of Pasqal, Fresnel. Then, we give some insights on how to use quenched dynamics to probe interesting properties of disordered systems such as many-body localisation.
In chapter 5, we perform extensive noisy emulations of quantum dynamics in the XY model to interpret experimental results obtained on Chadoq, the device of the academic group. Using these digital twins enables to benchmark the utility of new control techniques such as local rotations for state tomography, and validates results obtained at larger sizes in the context of probing a continuous sym-

metry breaking in the dipolar XY model. Reinforced with optimised driving schedules and accurate noise benchmarks, Rydberg platforms become prime synthetic simulators to probe exciting properties in quantum magnetism.

In chapter 6, we then report on the use of those platforms to tackle graph-related problems, notably in the combinatorial optimisation paradigm to which countless industrial applications can be mapped. We explain how graphs can be optimally mapped to atomic configurations and then present, and implement on Fresnel, optimised protocols to solve the Maximum Independent Set problem on unit disk instances.

After commenting on the current limitations and prospects of such quantum optimisation algorithms, we switch in chapter 7 to methods useful in quantum graph machine learning. There, we describe how a quantum feature map can be used, instead of classical kernels, for a binary classification task and report the performance obtained when implementing it on Iroise, the Pasqal prototype, to classify a dataset of graphs representing toxic or harmless molecules.

Finally, in chapter 8, we build a quantum-enhanced machine learning algorithm tackling a financial risk management use case, utilising a random graph sampling method tailored to Rydberg platforms. After showcasing the performance of implementing such algorithm on Iroise and comparing it to classical methods, we comment on the best use of quantum resources as building blocks in larger hybrid classical-quantum algorithms. For current analog Rydberg atom-based quantum computers to become viable solutions for industrial challenges, one imperatively need to take into account the underlying technical constraints in the theoretical possibilities in order to successfully implement relevant algorithms.

1 - From isolated atoms to many qubits: control and interactions

Contents

In this chapter, we detail the inner workings of a neutral-atom quantum processor, from trapping individual atom to engineering quantum dynamics in a many-body interacting system. Inside a Quantum Processing Unit (QPU), many physical processes remain hidden to the user perspective. Understanding the constraints imposed by each of them and precisely modelling their actions on the quantum processing part is paramount for building directly implementable protocols as well as assessing the extent to which hardware imperfections will impact high-level algorithmic results. Although we intend to describe as precisely as possible the physical processes harnessed in such quantum processors, more exhaustive descriptions of the experimental set-ups can be found in the following theses: [3–5].

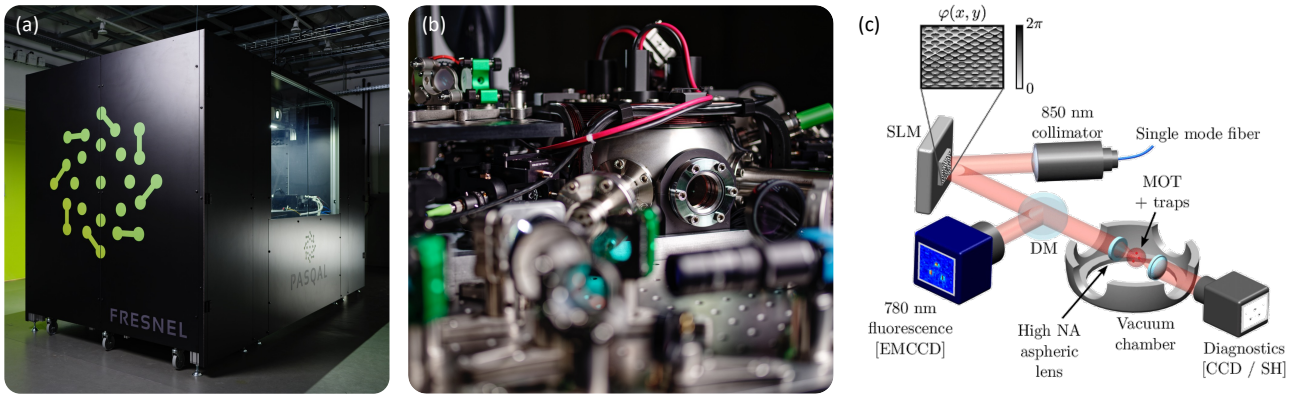


Figure 1.1: **QPU based on neutral atom technology**. (a) Outside of Fresnel, the neutral atom-based quantum computer developped by Pasqal. This shiny massive black box weights around 1.5 ton for dimensions of approximately $1.7 \times 2 \times 3\text{m}^3$. (b) The inside of Fresnel comprises 5 laser sources and more than 5000 optical, electronic and mechanical components. In total, those elements only use 5kW of power, the equivalent of a few kettles. The low consumption of such a device puts it well below current supercomputers ($\sim$ MW) [6] or quantum computers with dilution fridges like Google's Sycamore chip ($\sim$ tens of kW) [7]. All the quantum computation steps occur at room temperature in the vacuum chamber (background). (c) Overview of the main hardware components of neutral-atom based quantum processor.

1.1 - Manipulating neutral atoms with tweezers

Atoms serve as the foundational building blocks upon which qubits are encoded. Therefore, the individual trapping of atoms is a prerequisite for this technology.

1.1.1 - Trapping atoms with light

Isolating an atom from a macroscopic piece of material involves several experimental steps which could each be the subject of an entire chapter. The journey of the atom starts in a block of rubidium (^{87}Rb) metal, heated in either an effusive oven or glass cell, up to 60°C to prepare an atomic vapour. Only atoms collimated along a chosen axis can then exit. Depending on the set up, atoms are then slowed/cooled down using either a Zeeman slower or a 2D Magnetic Optical Trap (MOT) [8], creating an atomic beam with velocity $\sim$ 10 m/s. The atoms end up forming a dilute cloud confined inside a vacuum chamber by a 3D MOT created from a combination of a quadrupole magnetic field and six counter-propagating laser beams. At the center of the vacuum chamber, a 1 mm^3 cloud of 10^6 cold atoms is trapped at temperature of hundreds of μK. To isolate an atom, a far-off resonant laser beam is tightly focused down to a waist $\sim$ 1 μm using a high-numerical aspherical lense located inside the chamber. This effectively creates a trapping potential with Gaussian profile (see Eq. 1.1) with typical depth $\sim k_B \times 1$ mK ($\sim h \times 20$ MHz), well above the MOT temperature. Atoms thus falls into this *optical tweezer* and if undisturbed remain there due to a friction force provided by the MOT light. Moreover, two atoms cannot coexist in one tweezer due to a *collisional blockade* mechanism [9]. Whenever a tweezer becomes doubly occupied, the MOT light facilitates the transfer of the pair to a molecule in an excited state with attractive potential. The two atoms are then accelerated towards each other

with sufficient kinetic energy to expel them both from the tweezer. Overall, a trap has 50% of being occupied by a single atom and 50% of being empty. An even more sophisticated technique, known as *gray molasses* (in contrast to the previously mentioned *red molasses*), allows for an increase in the filling probability up to 90% [10] by using the repulsive branch of the molecular potential, producing enough kinetic energy for only one atom to be expelled from a trap. Another method consists in shielding the atom once in a trap by modifying the intensity profile to prevent another atom to fall in. Once in a tweezer, the single atom is confined in a μm region with temperature $T \sim 10\mu$K and can be further cooled down using techniques such as Raman sideband cooling [11]. Another possibility is to adiabatically lower the trap depth [12] to slow the atom by a factor $\alpha^{1/4}$ at the expense of the atom position dispersion increased by the inverse factor. The lifetime of an atom confined within a tweezer is constrained by two primary factors: random collisions with residual atoms present in the vacuum chamber and the heating induced by the tweezer light scattering off the atom off-resonantly. Mitigating the impact of random collisions necessitates enhancing the vacuum environment within the enclosure, through the implementation of sophisticated apparatus such as a cryostat for instance. Remarkable progress has been made in extending atom lifetimes to unprecedented durations, exemplified by the achievement of a record lifetime of 6000 seconds under optimised cryogenic conditions [13]. Addressing the heating effect induced by off-resonant tweezer light involves periodic cooling of the trapped atoms.

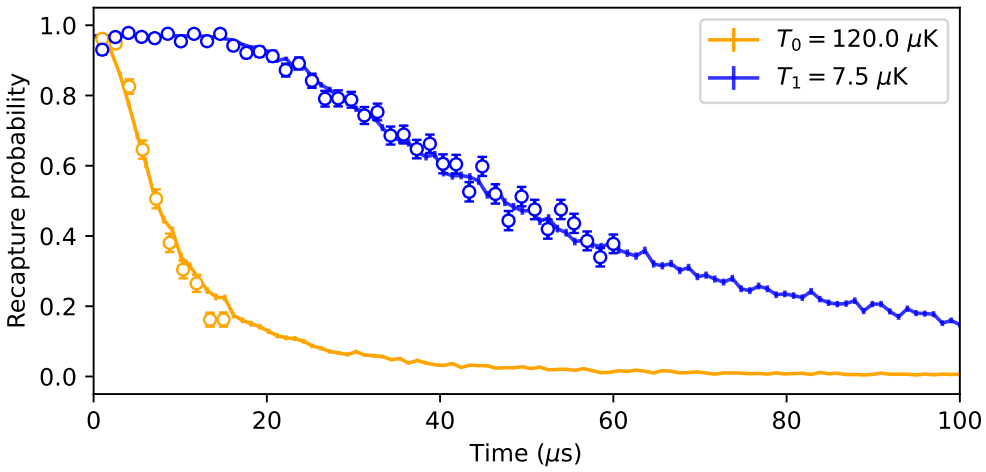


Figure 1.2: **Release-and-recapture experiment**. The probability of recapturing an atom after switching off the traps for a given duration is measured on the QPU (dots) before (yellow) and after an adiabatic lowering (blue). The decrease is fitted (solid line) using a Monte Carlo sampling of positions and velocities of the atoms at temperature T. Before the cooling, the temperature is estimated to be at $T_0 \approx 120\ \mu$K while after the cooling, it has been reduced to $T_1 \approx 7.5\ \mu$K as the atoms have been slowed down by a factor $\alpha^{1/4} \approx 0.25$.

There are several ways to generate a pattern of traps for capturing tens or hundreds of atoms. The options include an array of microlenses with fixed geometry, each of them generating one trap, [15], an Acousto Optic Deflector (AOD) to control multiple beams [16] or a Spatial Light Modulator (SLM) which imprints a adjustable phase pattern to the tweezer beam [17]. With this holographic

Box 1: Temperature in a tweezer

When the traps are turned off, the atoms are in free fall. At that moment, the distribution of velocities $v_0(T)$ can be approximated by a centered Gaussian distribution with standard deviation $\sigma_v(T) = \sqrt{k_B T/m}$ with $k_B = 1.38 \times 10^{-23}$ J/K and $m = 1.45 \times 10^{-25}$ kg for rubidium. In addition, random motion in the traps prior to the dynamics give rise to a thermal spatial distribution $\vec{r}_0(T)$. The latter can be modelled with good approximation by a Gaussian distribution spread around the trap center with $\sigma_{x,y,z}(T) = \sigma_v(T)/\omega_{x,y,z}$ with $\omega_{x,y,z}$ the angular trap frequency along each confinement direction. We consider a trap with depth $U_0/h \approx 20$ MHz, produced by a laser of wavelength $\lambda = 850$nm, with waist $w(z) = w_0\sqrt{1 + z^2/z_r^2}$, $w_0 = 1$ μm being the minimum waist and $z_r = \pi w_0^2/\lambda$, the Rayleigh range. The axial confinement is usually much weaker than the radial confinement with $\omega_{xy}/\omega_z = \sqrt{2}\pi w_0/\lambda$ [14], leading to a larger motional spread along z, i.e. $\sigma_z(T)/\sigma_{xy}(T) \approx 5.2$. In order to measure the temperature of the atoms in the traps, the standard procedure, called release-and-recapture, consists in releasing them by turning the traps off and recapturing them after a time t by turning the traps back on. Varying t enables to reconstruct the loss of atoms as some of them fly too far from the trap. The potential energy of the trap is given by

$$E_p(\vec{r}) = -U_0 \left(\frac{w_0}{w(z)}\right)^2 \exp\left(-2\frac{x^2+y^2}{w(z)^2}\right) \tag{1.1}$$

while the kinetic energy of the atom is $E_c(t) = m\dot{r}(t)^2/2$ with

$$\vec{r}(t) = \vec{r}_0(T) + \vec{v}_0(T)t \tag{1.2}$$

and we consider an atom recaptured after a time t if $E_c(t) + E_p(\vec{r}(t)) < 0$. This model enables to approximate the behaviour of atoms leaving the traps, as shown in Fig. 1.2. It also enables to make the assumption that the atom losses do not decrease significantly for dynamics below 10 μs. The thermal motion related effects happening during the dynamics are described in Box. 10

method, creating multiple traps in a specific geometric configuration requires computing on-the-fly the inverse Fourier transform of the 2D spatial pattern desired in the image plane. The beam acquires the phase pattern by reflecting off of the SLM and consequently diffracting after the lens into many spots, creating layouts of arbitrary traps in 2D and even 3D [18]. The number of traps achievable is constrained by the available laser power (10^4 with current lasers [19]) while the diffraction efficiency limits the total size of the array. The minimal distance between tweezers is constrained by limitations in optical system design, such as diffraction effect and aberrations. The resolution of the microscope objective focusing the light puts a constraint on the theoretical minimum distance between traps. At NA=0.5 and trapping laser wavelength of 850 nm, traps can not be closer than 2 μm and optical aberrations often increases this distance to 4 μm. Overall, achievable inter atomic distances range from few μm to close to 100 μm, still limited by the field of view of the optics. In order to reduce the power needed to produce each trap, one can either start with an even colder cloud of atoms, focus the tweezer light to an even narrower spot using a higher Numerical Aperture (NA) or go closer to

resonance, at the expense of more heating. After this loading stage, each of the N_t traps is randomly filled with success probability p_{fill}. The number of filled traps in an array thus follows a binomial law $\mathcal{B}(N_t, p_{\text{fill}})$, creating random initial configurations of, on average, $N = p_{\text{fill}} N_t$ atoms. While this may be sufficient for some implementations (see Sec. 8.3), many problems require a specific spatial configuration of atoms and thus a way to rearrange them.

1.1.2 - Imaging the atoms

Once a trap pattern has been loaded, a first fluorescence image of the atoms is taken. The scattered light is collected by the same lens which creates the tweezers while a dichroic mirror separates it from the tweezer beam. An electron-multiplying charge-coupled-device (EMCCD) camera collects the photons with a typical integration time of 20ms and two levels of fluorescence can usually be separated by an adjustable threshold as shown in Fig. 1.3. The photon distributions are fitted by two Gaussians whose intersection sets the threshold. Below the threshold, the photons are assumed to be collected from background scatter or camera noise, characterising an empty tweezer while above it, they indicates the presence of a trapped atom. This image will serve as initial condition for the rearrangement process described below.

1.1.3 - Rearranging atoms in a register

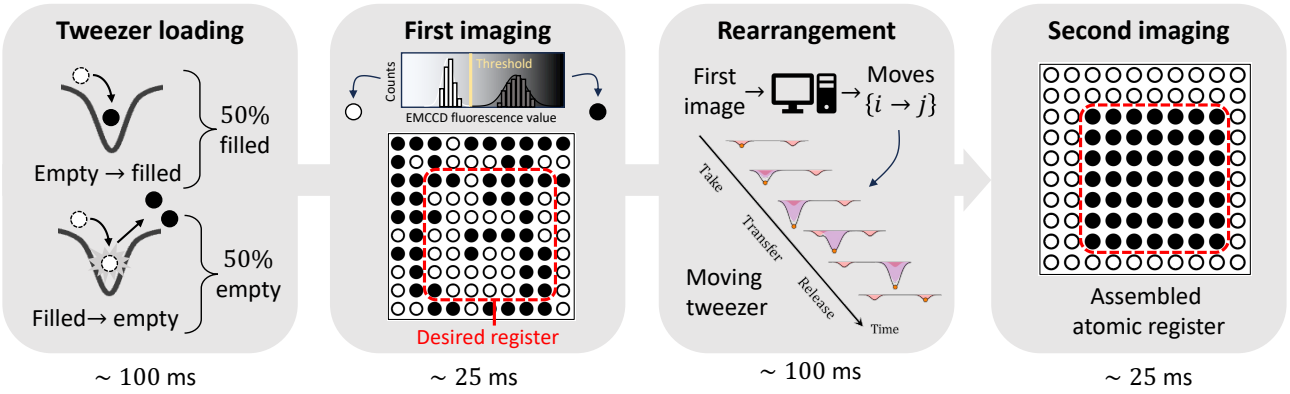


Figure 1.3: **Steps involved in an atomic register arrangement**. During the tweezer loading step, around half of the traps are randomly filled by an atom. A first fluorescence image enables to get the initial configuration of atoms using a predetermined measurement threshold distinguishing filled from empty traps. This initial image is fed to a rearrangement algorithm, computing moves carried out by a moveable tweezer, placing atoms in the desired configuration one at a time. A second image ensures the correct filling of the assembled register.

To initialise an atomic array consisting of N atoms at predetermined positions, two key steps are necessary. Firstly, creating a trap pattern entails arranging N traps at desired locations, along with approximately $N(1/p_{\text{fill}} - 1)$ extra traps positioned arbitrarily. Secondly, a method is required to transport atoms from the arbitrary traps in which they are initially loaded to the traps that need filling. This can be achieved using a programmable *moving tweezer* driven by two orthogonal AODs which can capture, move at 100nm/μs and release one atom at a time with efficiency $p_{\text{move}} = 99\%$.

With the first image giving the random initial filling, an algorithm computes set of moves on the fly to reach the assembled atomic *register*. The ordered set is usually optimised both in terms of total number of moves and travelling distance, with for instance the *Linear Sum Assignment Problem* solver [20] or the Gerchberg-Saxton algorithm [21]. To reduce the duration of the active feedback, the transfer of array pictures and the generation of the moving tweezer signals can be taken care of by dedicated Field-Programmable Gate Arrays (FPGA) while the algorithm runs on an external Graphics Processing Unit (GPU). Thus a rearrangement step with a unique moving tweezer usually takes $\sim$ 100ms for hundreds of atoms [22]. Other algorithms [23, 24] allows only one move per atom to reduce the risk of losing it when it is captured or released or harness parallel control operations to speed the rearrangement. A second fluorescence image is usually taken to ensure a perfectly filled array or at least to post select on experiments with no defects. However, as the arrangement efficiency scales with p_{move}^N, larger arrays might require several iterations of this procedure (imaging followed by rearrangement) to reach perfect filling with the initial number of atoms needed also increasing or a way to continuously load the traps. Those methods have enabled the arrangement of arrays of up to $N = 324$ atoms [25] and even more recently of $N = 1225$ [26]. Other techniques include dynamically changing the SLM pattern [27] with the atoms already trapped, or using a 1D array of moving tweezers controlled by the same two AODs to speed movements when possible [28, 29].

1.2 - Controlling qubits with lasers

Individually trapped atoms constitute the building block of the quantum processor as one can encode a qubit in their internal degree of freedom.

1.2.1 - An atom is a two-level system

Choosing which two levels to label as the logical qubit states $|0\rangle$ and $|1\rangle$ depends both on which interesting properties they may have, such as long lifetime or strong interactions, but also on the capacity of addressing the transition between the two. Coupling two internal states can be done using a laser, characterised by the amplitude $e(t)$, frequency $\omega(t)$ and phase $\phi(t)$ of its electric field, to the atomic transition involved with frequency $E_1 - E_0$. The resulting Hamiltonian of this two-level system in the rotating wave approximation – neglecting fast oscillating terms, valid when $\omega(t) \sim (E_1 - E_0)/\hbar$ – is

$$\begin{aligned}\hat{H}_c(t)/\hbar &= \Omega(t)\left(e^{i\phi(t)}\,|0\rangle\langle 1| + e^{-i\phi(t)}\,|1\rangle\langle 0|\right) - \frac{\delta(t)}{2}\left(|1\rangle\langle 1| - |0\rangle\langle 0|\right) \\ &= \frac{\Omega(t)}{2}\left(\cos\phi(t)\hat{\sigma}^x - \sin\phi(t)\hat{\sigma}^y\right) - \frac{\delta(t)}{2}\hat{\sigma}^z,\end{aligned} \tag{1.3}$$

with $\hat{\sigma}^\alpha$ the Pauli matrices ($\alpha = x, y, z$). The control fields are the Rabi frequency $\Omega(t) \propto e(t)$, the detuning $\delta(t) = \omega(t) - (E_1 - E_0)/\hbar$ and the phase $\phi(t)$ and can thus be varied in time by changing the intensity and frequency of the laser field. These controls are not strictly speaking independent, with for instance the relation $\delta(t) = \frac{d\phi}{dt}$ making it not physically possible to change both δ and ϕ independently at the same time. We will refer in the following to *pulses* which are defined as a modulation of amplitude Ω and detuning δ (or sometimes ϕ) during a given time t_{pulse}.

$\hat{H}_c$ enables to control the state of a qubit, usually represented on the *Bloch sphere*, by sending pulses to it. In the Bloch sphere representation, for each instant t of a pulse, this Hamiltonian

Box 2: Modulating the control fields

The amplitude, frequency, and phase profiles desired for the control field are generated using an Arbitrary Wave Generator (AWG) and are applied to the laser field using two kinds of shaping devices: the Acousto- and Electro- Optic Modulators (AOM/EOM).

- an EOM employs the electro-optic effect occurring when an electric field is applied to a crystal, causing a change in its optical properties. Effectively it can be used as an on/off switch for the amplitude with rise time $\tau_{\text{EOM}} \sim 10$ ns. This enables to switch Ω between $\Omega_{\text{off}} = 0$ and Ω_{on} while fixing a constant value of δ, engineering train of square pulses with the same amplitude.
- an AOM employs the acousto-optic effect, occurring when an acoustic wave interacts with a light wave inside a medium. By varying the frequency and intensity of the acoustic wave, the light wave direction and frequency can be modulated with typically longer rise time $\tau_{\text{AOM}} \sim 100$ ns. This enables to continuously change in time the value of the control fields $(\Omega, \delta \to \Omega(t), \delta(t))$.

Those shaping devices also act as low pass filters due to their finite rise time/bandwidth and the related distortions can be modelled to first approximation by convolving (here denoted by $*$) with the filter response function :

$$\tilde{\Omega}(t) = \Omega(t) * [\sqrt{\pi} f_c^2 e^{-\pi^2 f_c^2 t^2}] \text{ and } f_c = B/\ln\sqrt{2} = 0.48/(\tau_{\text{EOM/AOM}} \ln\sqrt{2}) \tag{1.4}$$

where $\tilde{f}(t)$ stands for the Fourier transform and f_c the modulation bandwidth. The finite bandwidth of the shaping devices gives a lower bound $T_{\min}$ on the duration of a pulse and on the minimum delay between two pulses to keep them well separated as shown in Fig. 1.4.

describes a rotation of the state vector around the axis $\mathbf{\Omega}(t) = (\Omega(t)\cos\phi(t), -\Omega(t)\sin\phi(t), -\delta(t))$ with angular velocity $\Omega_{\text{eff}}(t) = |\mathbf{\Omega}(t)| = \sqrt{\Omega(t)^2 + \delta(t)^2}$. For instance, rotating a qubit state from $|0\rangle$ to $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ amounts to a $\pi/2$-rotation around the y-axis and can be achieved with Ω constant, $\delta = 0$ $\phi = 0$ and $t_{\text{pulse}} = \pi/(2\Omega)$. More generally, at resonance and fixed ϕ, we denote a rotation of angle $\theta = \int_0^{t_{\text{pulse}}} \Omega(t)dt$ around $\mathbf{\Omega}(t) = \Omega(t)\boldsymbol{e}(\phi)$ with $\boldsymbol{e}(\phi) = (\cos\phi, -\sin\phi, 0)$ by $R_{\boldsymbol{e}(\phi)}(\theta)$ which can be decomposed as a rotation around the Bloch sphere's x-axis, conjugated by z-rotations

$$R_{\boldsymbol{e}(\phi)}(\theta) = R_{\boldsymbol{z}}(-\phi) R_{\boldsymbol{x}}(\theta) R_{\boldsymbol{z}}(\phi). \tag{1.5}$$

Adding another z-rotation, virtually applied through a shift in the phase reference frame [30], constructs an arbitrary unitary operator $R(\gamma, \theta, \phi) = R_{\boldsymbol{z}}(\gamma) R_{\boldsymbol{x}}(\theta) R_{\boldsymbol{z}}(\phi)$ which can drive a qubit from an arbitrat initial state to anywhere on the Bloch sphere. Thus with resonant pulses and phase reference frame changes, the state of a single qubit can be fully controlled.

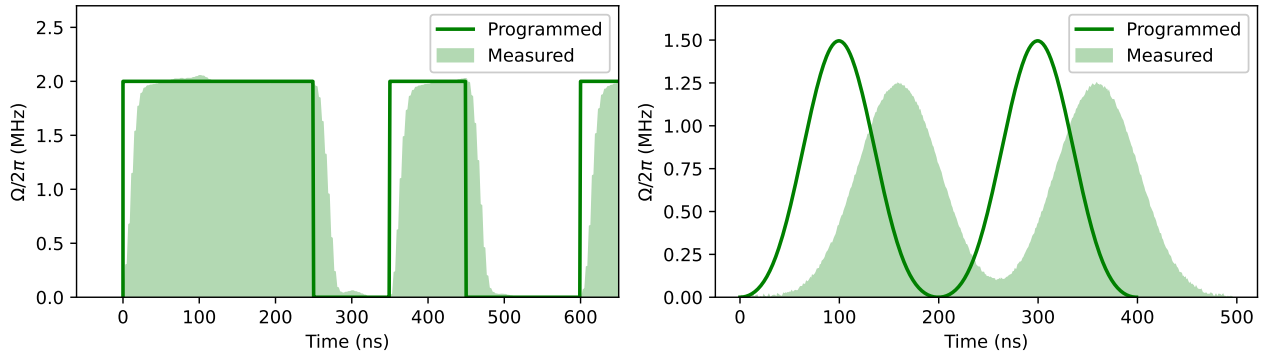


Figure 1.4: **Input vs output controls**. The programmed shape of $\Omega(t)$ (solid lines) is distorted by the finite bandwidth of the shaping device (left-EOM, right-AOM). The output shape (filled) can be measured with a photodiode and the rise time can be approximated : $\tau_{\text{EOM}} \approx 20\text{ns}$ and $\tau_{\text{AOM}} \approx 60$ ns

1.2.2 - Approximating a two-photon transition

While modern lasers are capable of emitting light across a broad spectrum of frequencies, there are still limitations to the range of frequencies they can cover. Some transitions in the energy spectrum of an atom occur at frequencies outside of this range and coupling indirectly two levels can require more complex protocols. For instance, the energy transition used below in Sec. 1.3.3 has a wavelength of 297 nm, making it hard, although possible, to address it with a single laser. However, with a two-photon scheme, i.e. using two lasers labelled "red" and "blue" and an intermediate state $|e\rangle$ as a stepping stool, one can excite an atom from $|0\rangle$ to $|1\rangle$. The Hamiltonian of the three-level system reads

$$\hat{H}_c(t)/\hbar = \frac{1}{2}(\Omega_b(t)e^{i\phi_b(t)}\,|e\rangle\langle 0| + \text{h.c.}) + \frac{1}{2}(\Omega_r(t)e^{i\phi_r(t)}\,|1\rangle\langle e| + \text{h.c.}) - \delta_b(t)\,|e\rangle\langle e| - \delta_r(t)\,|1\rangle\langle 1|\,. \quad (1.6)$$

Dealing with such a system is usually achieved by approximating it by an effective two-level system accounting for the coupling between $|0\rangle$ and $|1\rangle$. The approximation becomes valid if $\delta_b \gg \Omega_{\text{r/b}}$ and the Hamiltonian in Eq. 1.3 is retrieved with the following effective controls

$$\begin{aligned}\Omega(t) &= \frac{\Omega_r(t)\Omega_b(t)}{2\delta_b}\\ \delta(t) &= \delta_r(t) + \frac{\Omega_b(t)^2 - \Omega_r(t)^2}{4\delta_b}\\ \phi(t) &= \phi_b(t) - \phi_r(t).\end{aligned} \quad (1.7)$$

The challenges associated with this technique are twofold. First, the two lasers have to be locked in frequency, in order to be able to coherently drive the transition with a stable phase difference. Laser phase noise (see Box. 3) can thus decrease the driving efficiency. Second, $|e\rangle$ lifetime being usually shorter than the effective pulse duration, part of the population brought in $|e\rangle$ during the dynamics actually leaks back to $|0\rangle$ instead of occupying $|1\rangle$ (see Box. 6). This can be prevented by working far-detuned from the intermediate state (i.e. $\delta_b \gg \Omega_{\text{r/b}}$), effectively populating $|e\rangle$ with negligible amplitude and limiting the decoherence. Thus achieving $R_{\boldsymbol{x}}(\pi)$ only requires to fine-tune the red and

blue laser control parameters to work at resonance and $\phi = 0$. Another method, called STImulated Raman Adiabatic Passage (STIRAP), induces a population transfer from $|0\rangle$ to $|1\rangle$ without populating $|e\rangle$, while working at $\delta_b = 0$ and is described in Box. 14.

Box 3: Laser phase noise

In theory, lasers used in this work are considered as perfectly monochromatic, emitting at one fixed frequency, and with perfectly stable phase. However, in practice, they exhibit a small but finite range of emission, making them not monochromatic, and thus subject to phase noise, resulting in random fluctuations in their phase over time. For a two-photon Rydberg excitation scheme, $\Omega(t) = |\Omega(t)|e^{i\phi(t)}$ and shot-to-shot variations of $\phi(t) = \phi_r(t) - \phi_b(t)$ can lead to loss of contrast when averaging over many shots. The time-dependent phase can be described by random processes fluctuating in time whose behaviours are derived from the laser power spectral density $S_\phi(f)$ [Fig. 1.5(a)]. The frequency stabilisation of a laser can be done by locking it to an ultra-stable cavity at frequency ν_{cav}. Variations from the locked frequency are handled using an active feedback loop operating on the Pound-Drever-Hall (PDH) error signal [31]. This error signal obtained by varying the laser frequency can be characterised by its slope K and one can experimentally measure this signal power spectrum density $S_V(f)$. The phase spectrum is thus derived with the equation

$$S_\phi(f) = S_\nu(f)/f^2 = S_V(f) \times \frac{1 + 4(f/\nu_{\text{cav}})^2}{Kf^2}. \tag{1.8}$$

and realisations of the phase [Fig. 1.5(b)] read

$$\phi_{r/b}(t) = \sum_f df \; 2\sqrt{S_{\phi_{r/b}}(f)}\cos(2\pi f t + r_f) \tag{1.9}$$

with $r_f \sim \mathcal{U}([0, 2\pi])$ [32].

1.2.3 - Constraints on control fields

Several constraints on the control fields directly arises from the hardware limitations. The finite power output of lasers puts an upper bound on achievable $\Omega_{\max}$ when coupling it with a specific transition. As already mentioned in Box. 2, the finite modulation bandwidth of the EOM/AOM devices alters the temporal shape of $\Omega(t)/\delta(t)$ and limits how fast ϕ can be changed in between pulses. In addition, AOMs have a restricted frequency range around which their diffraction efficiency remains reasonable, limiting the absolute value of detuning achievable by $|\delta^{max}|$. Typical values for the lasers used in this thesis are $\Omega_b^{\max}/2\pi \approx 80$ MHz, $\Omega_r^{\max}/2\pi \approx 35$ MHz, $|\delta_b^{\max}|/2\pi \approx 700$ MHz and $|\delta_r^{\max}|/2\pi \approx 10$ MHz. Thus an effective laser coupling with those parameters can not drive an atomic system faster than $\Omega^{\max}/2\pi = 2$ MHz. This value can be increased either by switching to more powerful lasers, increasing $\Omega_{r/b}^{\max}$, or by reducing δ_b at the expense of an increased spontaneous emission effect from the intermediate state.

The precision at which a control operation such as a $\pi/2$ rotation can be achieved is not only limited by the finite resolution of the calibration of control parameters. Fluctuations in the beam-

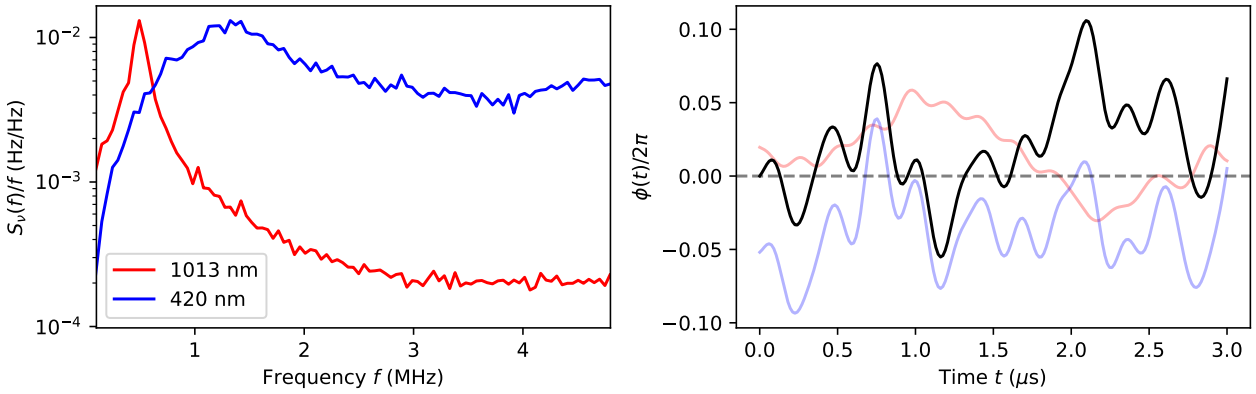


Figure 1.5: **Modelling laser phase noise**. (left) Experimentally measured power spectral density $S_\nu(f)$ of the frequency noise of a 1013 nm laser (red) and a 420 nm laser (blue). (b) Instances of time-dependent random processes $\phi_{r/b}(t)$ (red/blue) consistent with previous spectrums. The black line is the sum of the two coloured random processes, offset to have $\phi(t=0)=0$.

delivery system, such as optical fiber, can alter the value of the Rabi frequency from one shot to another. Effectively, Ω will follow a Gaussian distribution centred around its calibrated value and spread with standard deviation σ_Ω. This deviation is typically of the order of a few percent with a lower bound of 1% achieved with feedback loops. In the case of two-photon transitions, due to light shift terms [see Eq.(1.7)], this can also induce shot-to-shot fluctuations in δ.

In addition to constraints on maximum values, the resolution and stability of those controls is thus limited by several physical phenomena, including the one described in Box. 3. This leads to either slow drifts of parameters over durations longer than a typical experimental timescale, or to fast fluctuations in between repetitions of an experiment. Whereas in the first case, monitoring the controls can enable automatic recalibration using feedback loops, in the second, faster case, parameter values can be recorded and the results of experiments post-selected when a fluctuation threshold is exceeded, to the detriment of requiring more statistics.

Although precisely controlling isolated qubits already constitutes a scientific feat, engineering many-body dynamics requires a way to make those qubits interact with one another.

1.3 - Generating interactions between qubits

In a trapped array, atoms are seemingly isolated from one another. One solution to make them interact at those large scales consists in promoting them to Rydberg states.

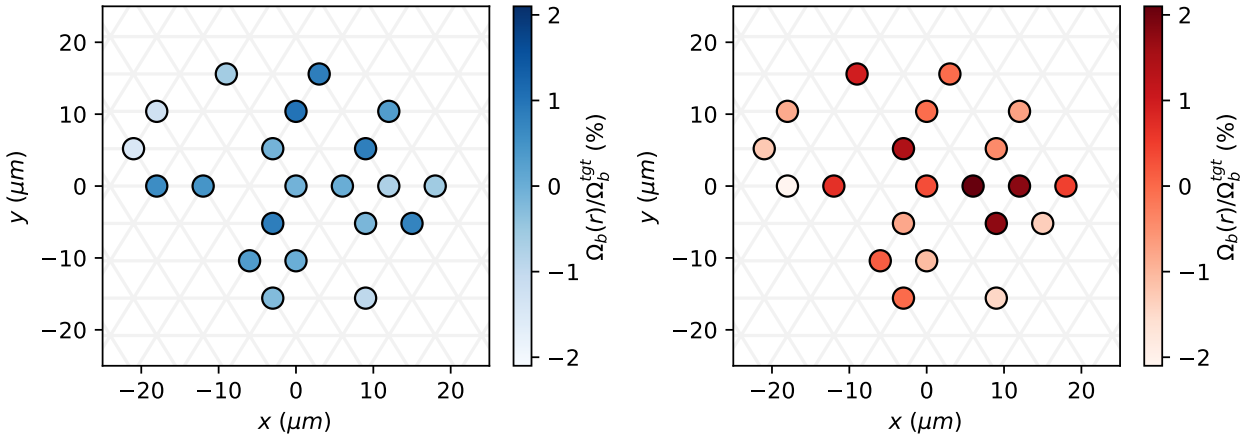


Figure 1.6: **Heatmaps of Rabi frequency** $\Omega_{r/b}(r)$. The Rabi frequency at each site of a trap layout is monitored for both blue (left) and red (right) lasers of a two-photon excitation scheme. Despite waists $w_r = 300 \pm 10$ μm and $w_b = 240 \pm 10$ μm, fluctuations up to 2%, due to optical aberrations, can be observed below 20 μm of distance, larger than those expected from a typical gaussian profile.

1.3.1 - Rydberg states and their properties

Atomic species typically used in neutral atom platforms such as rubidium or cesium share the property of having their valence orbital occupied by only one electron (or two in the case of strontium). Through a multi-photon excitation scheme, this electron can be excited to states characterised by a high principal quantum number n, effectively increasing the radius $\langle r_n \rangle$ of its orbit around the nucleus by several order of magnitude. Such highly excited electronic states are called Rydberg states and most of their properties can be characterised by just a few quantum numbers as explained in Box. 5.

Numerous Rydberg states properties scale with the effective quantum number $n^* = n - \delta(n, L, J)$, with δ a phenomenological quantum defect, introduced to account for the difference with the hydrogen atom. For rubidium, δ mostly depends on L with $\delta(0, 1, 2, 3) \approx 3.13, 2.64, 1.35, 0.016$ and $\delta(L > 3) = 0$. These Rydberg atoms can reach notable sizes as their radius scales with n^{*2}, extending into the micrometer scale at typical values $n \sim 50 - 100$. Moreover, these states remain quite stable, with lifetimes $\tau_n \sim n^{*3}$ extending to hundreds of μs, many orders of magnitude beyond the timescale of an excitation scheme.

The most compelling attribute of Rydberg atoms lies in their high polarisability due to the far orbiting electronic cloud. Nearby atoms or molecules can induce temporary fluctuations in the electron density of this cloud, mixing orbitals and creating charge asymmetries, forming an induced dipole moment in the Rydberg atom. The scalar polarisability $\alpha_n^{(0)}$ scales as n^{*7}, making high-lying Rydberg states more prone to significant energy shifts in the presence of undesired external electric fields. Choosing a Rydberg state in a experiment thus involves balancing Rydberg decay, introduced in Box. 6, and sensitivity to a noisy environment, both limiting the coherence time.

In this thesis we will use the following notation for states:

Box 4: Spatial dependency of laser parameters

Lasers generally exhibit a Gaussian intensity profile: the electric field they are applying to the atoms declines with distance from the focus point. For a circular Gaussian beam of waist w, the electric field can be approximated by:

$$e \propto \frac{w_0}{w(z)} \exp\left(-\frac{x^2+y^2}{w(z)^2}\right), \tag{1.10}$$

with $w(z) = w_0\sqrt{1+\frac{z^2}{z_r^2}}$ and z_r, the Rayleigh range. At the focal point, i.e. $z = 0$, the spatial inhomogeneity of Ω reads:

$$\Omega(r) \propto \exp\left(-\frac{r^2}{w^2}\right). \tag{1.11}$$

Thus, the atoms sit on a spatial gradient of Rabi frequency (and light-shift) and their motion (see Box. 10) can alter the control parameters applied to them. In a two-photon scheme, if both excitation lasers are aligned then the effective waist of Ω_{eff} is obtained from the red/blue waists $w = w_r w_b/\sqrt{w_r^2+w_b^2}$ and the resulting variation in light shifts can be derived from Eq. 1.7. While this Gaussian profile modelling can be a good approximation when working with large systems, i.e. having distance between atoms of the order of w, it will exhibit a flat behaviour for small systems. However, as conveyed by experimental measurements displayed in Fig. 1.6, spatial fluctuations of the order of % can still occur due to ghost reflections of the beam. A possibility to model those effects is to randomly draw a value of $\Omega(r_i)$ at each shot and for each qubit.

- $|g\rangle = |5S_{1/2}, F = 2, m_F = 2\rangle$
- $|e\rangle = |6P_{3/2}, F = 3, m_F = 3\rangle$
- $|nS\rangle = |nS_{1/2}, m_J = 1/2\rangle$
- $|nP\rangle = |nP_{1/2}, m_J = -1/2\rangle$
- $|D\rangle = |59D_{3/2}, m_J = -1/2\rangle$

Considering two atoms separated by $\vec{R} = R\vec{n}$ not coupled initially. The energy levels of the system Hamiltonian $\hat{H}_0$ can be grouped by multiplicity such that

$$\hat{H}_0 |i,\alpha\rangle = E_{i\alpha} |i,\alpha\rangle \text{ and } |E_{i\alpha} - E_{j\alpha}| \ll |E_{i\alpha} - E_{j\beta}|, \tag{1.12}$$

where i, j identifies different levels of the same manifold α. In our case, $|i,\alpha\rangle$ is a Rydberg pair state and a manifold α includes pair states not too distant in energy. When both excited to Rydberg states, the atoms experience a dipole-dipole interaction whose operator reads

$$\hat{V}_{dd}(R) = \frac{1}{4\pi\epsilon_0} \frac{\hat{\vec{d}}_1 \cdot \hat{\vec{d}}_2 - 3\left(\hat{\vec{d}}_1 \cdot \vec{n}\right)\left(\hat{\vec{d}}_2 \cdot \vec{n}\right)}{R^3} \tag{1.13}$$

where $\hat{\vec{d}} = e\hat{\vec{r}}$ is the outer shell electron dipole moment, with $e = -1.60 \times 10^{-19}$C and $\vec{r}$, distance from the reference point, usually the nucleus centre, to the electron. The interaction terms between states

Quantity	Symbol	Unit	Scaling	Values ($n \sim 50$)
Energy	E_n	GHz	n^{*-2}	$\sim 10^3$
Energy splitting	ΔE_n	GHz	n^{*-3}	50
Lifetime	τ_n	μs	n^{*3}	$\sim 10^2$
Orbit radius	$\langle r_n \rangle$	a_0	n^{*2}	$\sim 10^3$
Dipole moment $S \leftrightarrow P$	d	C·nm	n^{*2}	$e \times 100$
Dipole-dipole coeff.	$C_3 \propto V_{dd} \propto d^2$	hMHz·μm^3	n^{*4}	See 1.3.2
Van der Waals coeff.	$C_6 \propto V_{dd}^2/\Delta E_n$	hMHz·μm^6	n^{*11}	See 1.3.3
Ionization field	F_n^{ion}	V/cm	n^{*-4}	50
Scalar polarizability	$\alpha_n^{(0)}$	MHz/(V/cm)^2	n^{*7}	50

Table 1.1: Scaling of various quantities with the effective quantum number n^* of a Rydberg state [33].

$|i\rangle$ and $|j\rangle$ of the same manifold α are obtained by applying perturbation theory to $\hat{H}_0 + \hat{V}_{dd}(R)$ with

$$\begin{aligned}\langle i,\alpha|\, \hat{H}_0 + \hat{V}_{dd}(R)\, |j,\alpha\rangle &= E_{i\alpha}\delta_{ij} \\ &+ \langle i,\alpha|\, \hat{V}_{dd}(R)\, |j,\alpha\rangle \\ &+ \frac{1}{2}\sum_{k,\gamma\neq\alpha} \langle i,\alpha|\, \hat{V}_{dd}(R)\, |k,\gamma\rangle \langle k,\gamma|\, \hat{V}_{dd}(R)\, |j,\alpha\rangle \left[\frac{1}{E_{i\alpha}-E_{k,\gamma}} + \frac{1}{E_{j\alpha}-E_{k,\gamma}}\right] \\ &+ \dots\end{aligned} \tag{1.14}$$

The dipole-dipole coupling induces energy shifts in the Rydberg spectrum of the coupled pair, leading to first- and second-order interaction terms which intensities are derived in Box. 7. The second term represents the direct coupling between levels, from which arises the $\propto 1/R^3$ spin flip coefficient while the third represents the indirect coupling between these two levels via all the k levels of the other multiplicities γ, from which arise the $\propto 1/R^6$ coefficients. Depending on the parity of the pair states involved, some terms can be cancelled leading to two distinct behaviours as explained in the next section.

1.3.2 - XY model from Rydberg-Rydberg encoding

First-order terms in Eq. 1.14 can give rise for specific choices of Rydberg pair states, such as states with consecutive n and opposite parity of L, to strong dipolar interaction terms scaling as n^4 and $1/R^3$. In the following, the qubit states are $|0\rangle = |nS\rangle$ and $|1\rangle = |nP\rangle$ and the transition between the two can be addressed with coupling strength Ω_{MW} using a microwave antenna operating at GHz scale (see Box. 8). The interactions of an array of atoms at position $\boldsymbol{r}$ can thus be mapped to the spin-1/2 XY/R^3 model with Hamiltonian

$$\hat{H}_{\text{XY}}(\boldsymbol{r}) = \frac{1}{2}\sum_{i<j} U_{dd}(\vec{r}_i - \vec{r}_j; n)\left(\hat{\sigma}_i^x\hat{\sigma}_j^x + \hat{\sigma}_i^y\hat{\sigma}_j^y\right) \tag{1.19}$$

where the interaction term reads

$$U_{dd}(\vec{R}; n) = \frac{C_3(n)/\hbar}{R^3} \times (1 - 3\cos^2\theta)/2 \tag{1.20}$$

Box 5: Quantum numbers characterising Rydberg states

Quantum states are usually denoted by a set of quantum numbers describing the properties of the quantum particles involved, such as the electron in the case of Rydberg states. A Rydberg state reads $|nL_J, m_J\rangle$ with

- n, the principal quantum number describing the energy level of the outermost electron. n takes on very large values in Rydberg states, often ranging from tens to a hundred, which indicates that the electron is highly excited and occupies an orbital with an extended radial distribution.
- L, the orbital angular momentum quantum number describing the shape of the electron orbital with $L < n$. While Rydberg states with low L such as S-states ($L = 0$) or P-states ($L = 1$) are commonly used due to the ease of excitation with lasers, circular Rydberg states with highest possible value of L at a given n ($L = n - 1$) exhibit interaction strengths and lifetimes orders of magnitude higher [34].
- J, the electron total angular momentum quantum number, i.e. $J = L \pm 1/2$, which incorporates the spin angular momentum of the electron.

In addition, another interesting number is F, the atom total angular momentum quantum number, i.e. $F = J + I$ with I the nuclear spin quantum number ($I = 3/2$ for Rb). It arises in the context of hyperfine structure of atomic energy levels, particularly in the study of atomic spectra in the presence of magnetic fields. However in Rydberg states, this coupling between J and I vanishes and F is no longer used.
To each angular momentum quantum number $i = L, S, J, F$ is associated a magnetic quantum number m_i describing the orientation of the angular momentum in space, with $|m_i| \leq i$.

with $C_3(n)$ the dipole-dipole interaction constant for $nS - nP$ pair states and $C_3(n) \propto n^{*4}$, $C_3(n = 60)/\hbar = 2\pi \times 3.0$ GHz·μm^3, the pair distance $R = ||\vec{R}|| \sim 10\mu m$ and θ the angle between $\vec{R}$ and the quantisation axis set by the magnetic field $\vec{B}$. As the angular dependency indicates, the interaction strength can thus be engineered to be anisotropic by setting $\vec{B}$ in the atomic plane as in [36]. Each term in the sum corresponds to a coherent exchange between the spin states, transforming the pair state $|0_i 1_j\rangle$ into $|1_i 0_j\rangle$ while conserving the total number of excitations in the system. A typical value of intensity for this interaction (for $\theta = 0$) is $U_{dd}(12.5\ \mu m; 60) \approx 2\pi \times 1$ MHz. This *flip-flop* interaction [Fig. 1.7(a)] enables exploring topological phases of matter [36] and frustration in quantum magnets [37] as studied in chapter 5 or investigating excitation transport [38] within contexts as intricate as the photosynthesis process for instance. It has indeed been proposed that the migration of energy towards light-harvesting complexes could be steered by such XY-like quantum dynamics [39]. More recently, the experimental realisation of a discrete quantum random walk with atoms on optical lattices [40], and promises of exponential speed up with their continuous counterpart [41] have highlighted the potential of using neutral atom platforms to study the continuous XY model.

Box 6: Finite lifetime and decay rate

The spectral linewidth of a state characterises the range of frequencies over which it can absorb or emit radiation. The natural linewidth is directly linked to the spontaneous decay rate, i.e. the probability per unit time that an atom in this state will undergo a transition to a lower energy state by emitting a photon. For instance, the state $|e\rangle$ will naturally decay to $|g\rangle$ with rate $\Gamma_e/3$ ($\Gamma_e/2\pi = 1.41$ MHz). However, other decay channels can lead to populating other states with rate $2\Gamma_e/3$. The finite spectral linewidth of the lasers used in the excitation scheme can also broaden the natural linewidth of a state and affect its decay rate. In the case of Eq. 1.6, the intermediate state decay rate is scaled by $(\Omega_r^2 + \Omega_b^2)/(4\delta_b^2)$.
Although highly stable, Rydberg states also exhibit finite lifetimes which depends both on quantum numbers and surrounding environment temperature T_r due to

- spontaneous emission towards lower energy states such as $|g\rangle$. The finite lifetime of the Rydberg state $\tau_n(0K) \propto n^{*3}$ with corresponding decay rate $\tau_n(0K)^{-1}$ where $\tau_{60S(P)}(0K) \approx 230.3(464.8)\mu s$.
- blackbody radiation happening at finite temperature. Since all objects emit thermal radiation according to their temperature, this radiation can induce transitions between Rydberg states with corresponding rate $\tau_n(T_r)^{-1} - \tau_n(0K)^{-1}$ with $\tau_{60S(P)}(T_r = 300K) = 100.6(134.6)\mu s$.

This finite Rydberg lifetime puts an upper bound on the duration of the quantum processing part and thus on the number and duration of protocols applied to atoms in Rydberg state before they start to significantly decay. It can be linked in an effective model to the relaxation time T_1.

1.3.3 - Ising model from ground-Rydberg encoding

For states with same parity, the first-order terms in Eq. 1.16 vanish, and another kind of interaction arises: the *van der Waals* (vdW) interaction. In the following, the qubit states are $|0\rangle = |g\rangle$, belonging to the ground state manifold of the rubidium atom, and $|1\rangle = |nS\rangle$. Addressing the transition between the two requires a two-photon excitation scheme as shown in Fig. 1.7(b) which can be modelled by an effective laser coupling as explained in section 1.2.2. The interactions of an array of atoms at position $\boldsymbol{r}$ can thus be mapped to the spin-1/2 Ising $1/R^6$ model with Hamiltonian

$$\hat{H}_{\text{Ising}}(\boldsymbol{r}) = \sum_{i<j} U_{vdW}(\vec{r}_i - \vec{r}_j; n)\hat{n}_i\hat{n}_j \tag{1.21}$$

where the interaction term reads

$$U_{vdW}(\vec{R}; n) = \frac{C_6(n)/\hbar}{R^6} u(\theta) \tag{1.22}$$

with $\hat{n}_i = (\hat{\mathbb{I}} + \hat{\sigma}_i^z)/2 = |1\rangle\langle 1|_i$ being the projector of qubit i to $|1\rangle$. $C_6(n)$ is the vdW interaction constant for $nS - nS$ pair states, with $C_6(n) \propto n^{*11}$, $C_6(n = 60)/\hbar = 2\pi \times 137.8$ GHz·μm^6. The pair distance $R = ||\vec{R}|| \sim 10\mu m$ and θ is the angle between $\vec{R}$ and the quantisation axis set by

Box 7: Derivation of $V_{dd}(R)$

A Rydberg state typically writes $|r\rangle = |nL_J, m_J\rangle$ with quantum numbers described in Box. 5. Let $|r\rangle$, $|r'\rangle$, $|r''\rangle$ and $|r'''\rangle$ be four Rydberg states. The coupling between pair states $|r, r''\rangle$ and $|r', r'''\rangle$ is given by:

$$\langle r, r''| \hat{V}_{dd}(R) |r', r'''\rangle = \frac{1}{4\pi\epsilon_0} \frac{\vec{d}_{rr''} \cdot \vec{d}_{r'r'''} - 3\left(\vec{d}_{rr''} \cdot \vec{n}\right)\left(\vec{d}_{r'r'''} \cdot \vec{n}\right)}{R^3} \tag{1.15}$$

This expression can be written as :

$$V_{dd}(R) = \frac{1}{4\pi\epsilon_0 R^3} ||\vec{d}_{rr''}|| ||\vec{d}_{r'r'''}|| \begin{bmatrix} A_0(\theta)\left(\Upsilon_1^{-1}\Upsilon_1^{+1} + \Upsilon_1^{+1}\Upsilon_1^{-1} + 2\Upsilon_1^0\Upsilon_1^0\right) \\ +A_1(\theta)\left(\Upsilon_1^{+1}\Upsilon_1^0 - \Upsilon_1^{-1}\Upsilon_1^0 + \Upsilon_1^0\Upsilon_1^{+1} - \Upsilon_1^0\Upsilon_1^{-1}\right) \\ +A_2(\theta)\left(\Upsilon_1^{+1}\Upsilon_1^{+1} + \Upsilon_1^{-1}\Upsilon_1^{-1}\right) \end{bmatrix} \tag{1.16}$$

$||\vec{d}_{rr''}||$ is the radial part of the dipole matrix element between $|r\rangle$ and $|r''\rangle$ and can be calculated using the *Pairinteraction* Python package [35]. θ is the angle between $\vec{n}$ and the quantisation axis set by the global magnetic field. Υ_1^k are the first order complex spherical harmonics. Their values are computed using the magnetic quantum numbers m_J and the following rule :

$$\Upsilon_1^{k_a}\Upsilon_1^{k_b} \neq 0 \quad iif \quad \begin{matrix} m''_J - m_J = k_a \\ m'''_J - m'_J = k_b \end{matrix} \tag{1.17}$$

The values of Υ_i^k are given in the spherical harmonics table and the angular dependencies ($\phi = 0$) are gathered inside

$$A_0(\theta) = (1 - 3\cos^2\theta)/2,\ A_1(\theta) = 3\cos\theta\sin\theta/\sqrt{2} \text{ and } A_2(\theta) = -3\sin^2\theta/2 \tag{1.18}$$

Note that each term $A_i(\theta)$ couples pair states $|r_a\rangle \leftrightarrow |r_b\rangle$ where the total magnetic quantum number $M = m_J^a + m_J^b$ changes by $\Delta M = \pm i$.

the magnetic field $\vec{B}$. While the vdW interaction also features anisotropic behaviour with $u(\theta)$, the variations remain small for nS states and more pronounced for nP or nD states. A typical value of intensity at $\theta = 0$ ($\pi/2$) for this interaction is $U_{vdW}(5\mu m; 60) = 2\pi \times 8.8$ (8.96) MHz. Using the scaling law $U_{vdW}(R; n) \propto n^{11}R^{-6}$ it is possible to approximately extrapolate to other principal quantum numbers or distances. The vdW interaction strength can thus be finely tuned over a broad range – from a few kHz to hundreds of MHz – by modulating inter atomic distances and selectively utilising Rydberg states.

The pairwise vdW interaction terms in Eq. 1.21 actually engineer the *Rydberg blockade* effect, shifting the energy of doubly excited states $|1\rangle_i |1\rangle_j$. The simultaneous excitation with effective amplitude $\Omega_{\text{eff}} = \sqrt{\Omega^2 + \delta^2}$ of pairs of atoms closer than the *Rydberg blockade radius* $r_b = (C_6(n)/\hbar\Omega_{\text{eff}})^{1/6}$ is thus prevented when $\Omega \ll U_{vdW}(R; n)$ even at resonance (i.e. when $\delta = 0$). In the meantime, interactions terms for atoms separated by far more than r_b can be neglected due to the sharp decay of U_{vdW}.

Box 8: Microwave manipulation

Transitions within the Rydberg manifold commonly occur at frequencies on the order of tens of GHz, as indicated by the energy differences ΔE_n in Table 1.1. Instead of lasers, these transitions are often driven using microwave (MW) radiation. The MW signal is typically produced by mixing a microwave oscillator output by a synthesizer with a radio frequency (RF) signal generated by an AWG. This enables to imprint a temporal shape in the MW signal to create a control pulse $\Omega_{\mathrm{MW}}(t), \delta_{\mathrm{MW}}(t), \phi_{\mathrm{MW}}(t)$. This pulse is radiated by a dipole antenna emitting outside the vacuum chamber and thus globally addresses the atoms with Rabi frequencies up to $\Omega_{\mathrm{MW}}^{max}/2\pi = 20$ MHz.

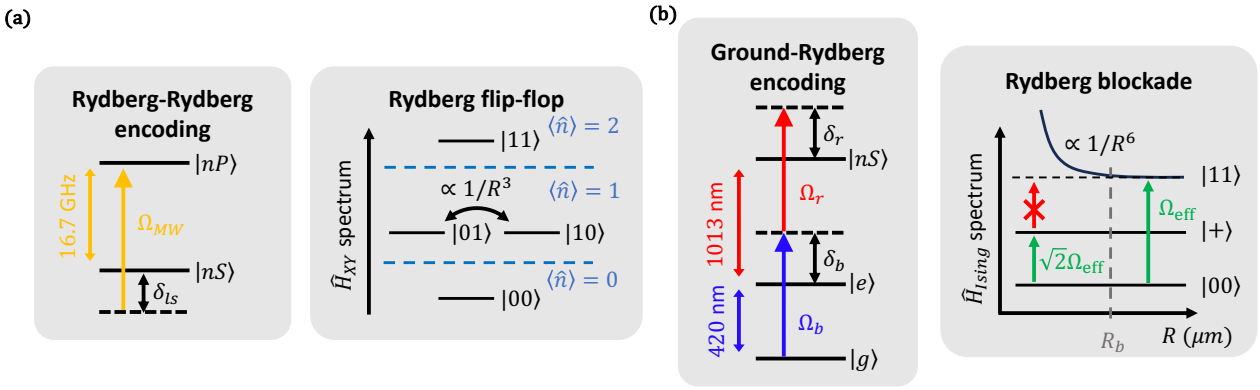


Figure 1.7: **Encoding qubits and effective models**. (a) The Rydberg-Rydberg encoding (using $|0\rangle = |nS\rangle$ and $|1\rangle = |nP\rangle$) gives rise to the Rydberg flip-flop effect (from XY model) which couples states within same magnetisation sector $\langle\hat{n}\rangle$. (b) The ground-Rydberg encoding (using $|0\rangle = |g\rangle$ and $|1\rangle = |nS\rangle$) gives rise to the shift in energy of the doubly excited state which leads to the Rydberg blockade effect. The two-photon process can be approximated by a two-level system when $\delta_b \gg \Omega_{r/b}$.

For a blockaded pair initialised in $|00\rangle$, resonantly addressing the $0-1$ transition (i.e. with $\delta = 0$) drives the system to the so called W-state $|W_2\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ with a coupling enhanced to $\sqrt{2}\Omega$ as conveyed by Fig. 1.7(b). Similarly, for an N-atom blockaded ensemble, states with more than one excitation and especially $|1\rangle^{\otimes N}$, become non-resonant and coupling to $|W_N\rangle = \sum_i |0\cdots 1_i \cdots 0\rangle / \sqrt{N}$ is enhanced by a factor $\sqrt{N}$. This blockade effect can be neutralised for specific values of detuning, when for example a pair of qubits interacting with strength U is subjected to a detuning $\delta \approx U/2$. In this scenario, the doubly excited state energy aligns closely with the ground state energy of the qubit pair, presenting an energy of $-2\delta + U \approx 0$, while the singly excited states experience a shift of $-\delta$. This *facilitation mechanism* may lead to undesired transfers of small population between states $|00\rangle$ and $|11\rangle$, effectively overriding the blockade.

While the Ising model has been extensively studied over the last decades, the understanding of numerous condensed matter phenomena remains challenging, even within this framework. These include the behaviour of geometrically-frustrated arrays [42], quench dynamics of many-body systems

Box 9: vdW terms in XY model

In the Rydberg-Rydberg encoding, the dipole-dipole interaction terms prevail during the dynamics at distances such that

$$d \gg \left(\frac{C_6(n)}{C_3(n)}\right)^{1/3} \underset{n=60}{\approx} 3.6\ \mu\text{m}. \tag{1.23}$$

However, at those distances the van der Waals interaction between similar Rydberg states can still alter the many-body dynamics in the XY model. It can be incorporated by adding to $\hat{H}_{\text{XY}}(\boldsymbol{r})$ the following Hamiltonian

$$\hat{H}_{vdW}(\boldsymbol{r}) = \sum_{i<j} \frac{1}{||\vec{r}_i - \vec{r}_j||^6} \left[C_6(n)u(\theta)(1-\hat{n}_i)(1-\hat{n}_j) + \tilde{C}_6(n)\tilde{u}(\theta)\hat{n}_i\hat{n}_j\right], \tag{1.24}$$

where $\tilde{C}_6(n)$ the van der Waals interaction constant for $nP-nP$ pair states. A typical value of $\tilde{C}_6(n)\tilde{u}(\theta)$ for $n=60$ and $\theta = 0(\pi/2)$ is $-217.1(-84.8)$ MHz$\cdot\mu$m^6 showing that the anisotropic behaviour is more significant for nP states.

[43] and many-body localisation on disordered lattices [44]. The complexity emerging from these problems and their simulations motivate the use of Rydberg platforms as quantum simulators as conveyed by examples of chapter 4. Moreover, as explained in chapter 6, some classes of combinatorial optimisation problems can also be tackled with such devices, by mapping them to the Ising model.

1.4 - Quantum computing with Rydberg atoms

1.4.1 - State initialisation

A crucial step in the quantum computation paradigm is the ability to initialise a set of N qubits to the simple state $|0\rangle^{\otimes N}$[45]. In the MOT, the rubidium atoms occupy their fundamental energy level $5S_1/2$ but are distributed among the Zeeman sublevels m_F of the hyperfine $F=2$ state. Initialising them specifically in $m_F = 2$ is achieved through *optical pumping* where cycles of excitation-deexcitation using a 780nm laser and a repumper end up populating $|g\rangle$. This enables to initialise an array of qubits in the ground-Rydberg encoding with efficiency $(1-\eta)^N$ with $\eta \approx 0.5\%$. For the Rydberg-Rydberg encoding, an extra step is needed to initialise all the atoms in $|0\rangle = |nS\rangle$ for instance. Provided that the atoms are far enough so that vdW interaction is negligible (see Box. 9), a STIRAP protocol (described in Box. 14) enables to globally excite all atoms in $|nS\rangle$ with efficiency $(1-\eta_S)^N$ with $\eta_S \approx 1\%$. While it is possible to initialise an array of qubits in $|\psi(t=0)\rangle = |0\rangle^{\otimes N}$ for both the Ising and the XY models, the preparation efficiencies are thus different.

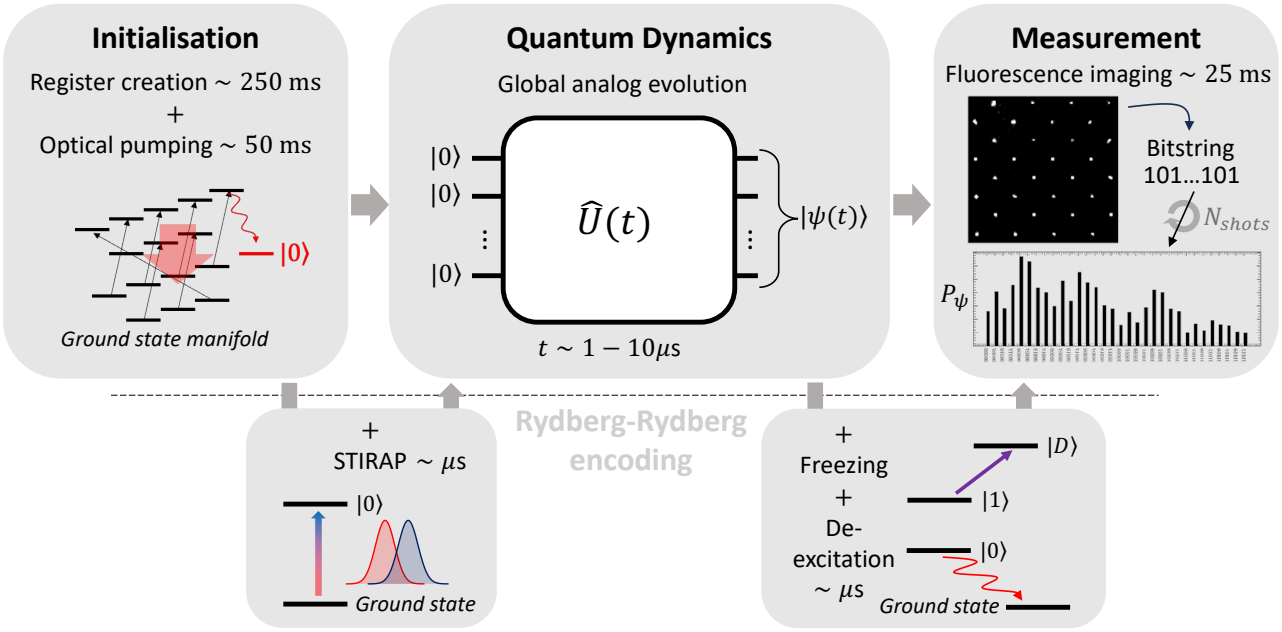


Figure 1.8: **Computation cycle**. In the initialisation phase, atoms are spatially arranged in a register and prepared in a specific initial state with optical pumping (and STIRAP). A global evolution operator $\hat{U}(t)$ produces a quantum state $|\psi(t)\rangle$. During the measurement phase, the qubit basis states are mapped to occupancy of the traps containing the atoms and repeating the global cycle several times outputs a probability distribution P^ψ.

1.4.2 - Performing quantum dynamics

The ability to evolve a quantum system in time starting from the simple initial state given above is at the heart of the quantum computation paradigm. Most quantum algorithms or quantum simulation protocols rely on either creating a specific state $|\psi\rangle$ from $|\psi(t=0)\rangle$ or looking at the characteristics of that state evolving in time. The Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \tag{1.25}$$

governs the dynamics of a quantum system with $\hat{H}$, the Hamiltonian of the system. For an array of N trapped atoms at positions $\vec{r}_i$, the global Hamiltonian reads

$$\hat{H}(t;\mathbf{r}) = \hat{H}_{\text{ctrl}}(t) + \hat{H}_{\text{int}}(\boldsymbol{r}) \tag{1.26}$$

with the control part enabling global coherent driving

$$\hat{H}_{\text{ctrl}}(t)/\hbar = \frac{\Omega(t)}{2} \sum_{i=1}^{N} (\cos\phi\ \hat{\sigma}_i^x - \sin\phi\ \hat{\sigma}_i^y) - \delta(t) \sum_{i=1}^{N} \hat{n}_i \tag{1.27}$$

and the interaction part depending on the qubit encoding, i.e. $\hat{H}_{\text{Ising}}$ for ground-Rydberg and $\hat{H}_{\text{XY}}$ for Rydberg-Rydberg.

The quantum state obtained when globally evolving an atomic system under this Hamiltonian for a duration $t - t_0$ is $|\psi(t)\rangle = \hat{U}(t - t_0)\,|\psi(t = t_0)\rangle$ with

$$\hat{U}(t) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar}\int_{s=t_0}^{t} \hat{H}(s;\mathbf{r})ds\right)\right] \tag{1.28}$$

being the propagator and $\mathcal{T}$ the time-ordering operator. In the analog mode, the global Hamiltonian can be directly fine-tuned by playing with either the control parameters or the atomic positions : it constitutes a resource for computation. This requires the user to delve into the physics of the system and work at a low level, close to the hardware. Most of the work presented in this thesis will be done using this mode.

Box 10: Temperature-induced noise during dynamics

As conveyed in Box. 1, the atoms are subject to thermal motion which affects the quantum evolution in a noisy way. When the traps are turned off, the random motion of the atoms introduces a change in the frequency of the light they emit or absorb. This Doppler effect thus induces random shifts in the detuning of each atom and can be modelled by additional terms δ_i generated from a centred normal distribution with deviation $\sigma_\delta(T) = ||\vec{k}||\sigma_v(T)$, with $\vec{k}$ the wave vector of the laser system and $||\vec{k}||/2\pi \sim \mu\text{m}^{-1}$. Thus for each new shot, the lasers are slightly off resonant and the dynamics of the system differs by the value of the detuning perceived.
In addition, this thermal motion affects shot-to-shot the global Hamiltonian in two separate ways. Firstly, the distance-dependent interaction strength $U(r)$ can be considered time-dependent if the dynamics lasts long enough such that $\sigma_v(T)t$ becomes a small fraction of R, the typical spacing between atoms. Even during short dynamics, taking into account the random initial positions of the atoms may prove to be relevant as $\delta U/U = \alpha \times \delta R/R$ with $\alpha = 6$ for vdW interactions and $\alpha = 3$ for dipole-dipole interactions. Secondly, as explained in Box. 4, the intensity profile of the laser is not constant and in the case of tightly focused addressing beams ($w \sim$ a few μm), moving atoms may deviate far away from the focus point if $\sigma_r(T) \sim w$, resulting in a diminishing control amplitude Ω with time.

1.4.3 - Measurement of observables

During a quantum experiment or algorithm implementation, a target quantum state $|\psi\rangle$ is produced at the end of the pulse sequence and an *observable* $\hat{A} = \sum_i a_i\,|a_i\rangle\langle a_i|$, $a_i \in \mathbb{C}$, needs to be measured to retrieve valuable information on the system. When an observable is measured, the quantum system is randomly projected onto one of the eigenstates $|a_i\rangle$ of $\hat{A}$ with probability $|\langle a_i|\psi\rangle|^2$. The expectation value of $\hat{A}$ over $|\psi\rangle$ reads

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \sum_i a_i|\langle a_i|\psi\rangle|^2. \tag{1.29}$$

Note that after a measurement, the original qubit state is lost, forcing to repeatedly prepare and measure $|\psi\rangle$ to approximate the distribution of $|\langle a_i|\psi\rangle|^2$. Thus probing the result of a quantum algorithm requires repeated applications of said algorithm, or *shots* due to the probabilistic nature of quantum

states. One can then approximate $\langle \hat{A} \rangle$ by averaging the outcomes of the measurements.

On neutral atom devices, for any qubit encoding, one can measure in the computational basis $0-1$ by mapping the qubit basis states $|0\rangle$ and $|1\rangle$ to the occupancy of the trap. The measurement process thus requires the imaging technique mentioned in 1.1.2 which enables distinguishing filled from empty traps with a global fluorescence picture. In the case of ground-Rydberg encoding, this mapping is made naturally when the traps are turned back on at the end of the dynamics. Indeed, the tweezers exert a repulsive ponderomotive potential on the far-orbiting electron of the Rydberg atoms, expelling them from the trapping region. On the contrary, atoms in their ground state are recaptured by the trapping potential given in Eq. 1.1 and appear as bright spots on the detected image. In the case of Rydberg-Rydberg encoding, two extra steps are needed. Let $|0\rangle$ and $|1\rangle$ be the same states as in 1.3.2 and add $|e\rangle$ and $|D\rangle$. First, a *freezing* protocol enables to shelve the population from $|1\rangle$ to $|D\rangle$, which hardly interacts with $|0\rangle$, using MW pulses. The system is quickly "frozen", preventing interaction-induced dynamics between $|0\rangle$ and $|1\rangle$ from happening during the following longer step. Second, a subsequent deexcitation from $|0\rangle$ to $|e\rangle$ is performed by applying a $\tau_{\text{desex}} \sim \mu$s light pulse at resonance: the atoms originally in $|0\rangle$ then quickly decay back to the $5S_{1/2}$ manifold where they can be recaptured and imaged. Atoms in $|D\rangle$ are also expelled by the repulsive ponderomotive potential, effectively mapping the basis states of the XY model to the presence or absence of atoms in the tweezers. This global measurement procedure provides a sufficiently accurate estimate, if repeated over enough shots, of the distribution of $\langle \hat{n}_i \rangle$. Thus for each qubit i, $p_i(0) = 1 - \langle \hat{n}_i \rangle$ and $p_i(1) = \langle \hat{n}_i \rangle$ and we denote P_ψ, the probability distribution describing a perfect measurement of the system in the computational basis. In practice, a quantum system is only sampled within a limited budget of shots, collecting a finite set of bitstring $n_1 \cdots n_N$ and thus subjecting measurements to sampling noise.

The measurement process is inherently flawed by several physical processes which can result in measuring a 1 instead of a 0, leading to *false positive* detection event and conversely to *false negative* (see Box. 11). Background-gas collisions can eject a recaptured atom, emptying a trap and thus flipping a 0 to a 1 in a bitstring with probability $\varepsilon \approx 1\%$. Additionally, the ejection of atoms in Rydberg states typically lasts around a few microseconds, leaving enough time for some of them to decay from $|1\rangle$, effectively flipping a 1 to a 0 in a bitstring with probability $\varepsilon' \approx 5\%$. Reducing the temperature T of the background atoms help lower ε while increasing the Rydberg lifetime by using larger n can lower ε'.

1.4.4 - Cycle rate of neutral atom technology

The major limitation of neutral atom technology naturally arises when summing the durations of all the steps presented above. While the actual quantum dynamics happens at the MHz scale, the necessity to load the tweezers ($\sim$ 100ms), arrange the atoms ($\sim$ 150ms), initialise their state ($\sim$ 50ms) and image them several times ($\sim$ 100ms) for each shot reduces the repetition rate of the useful computation to a few Hz. In addition, the effective repetition rate is even scaled down by the probability of assembling a defect-free quantum register at each cycle. As moving an atom from one trap to another has a finite efficiency η due to losses, reaching a perfect arrangement of N atoms scales as η^N. For instance arranging a 100-node graph may take $\eta^{-N} \approx 3$ times more cycles at $\eta = 99\%$.

Boosting the repetition rate to start closing this frequency gap with others technologies thus carries

Box 11: Effective model for detection errors

The many physical processes inducing bit flips during the measurement phase can be effectively encompassed by two terms $\varepsilon = p(0 \to 1)$ and $\varepsilon' = p(1 \to 0)$. Keeping in mind that this definition depends on the physical states chosen as $|0\rangle$ and $|1\rangle$, the values of the two terms can vary for various experiments but are usually at the percent level. Moreover, when the atoms are addressed locally, those values can become site-specific. Modelling these bit flips can be achieved using the following transfer matrix

$$M_i = \begin{pmatrix} 1-\varepsilon_i & \varepsilon'_i \\ \varepsilon_i & 1-\varepsilon'_i \end{pmatrix} \tag{1.30}$$

As shown in Fig. 1.9, the incorrectly measured distribution is thus $\tilde{P}_\psi = (\bigotimes_i M_i)P_\psi$ where we assume uncorrelated errors. While the detection errors at the single qubit level remain low and easy to correct, they quickly scale with the size of the system. For instance, for $N = 100$ and $\varepsilon = 1\%$, measuring $|\psi\rangle = |0\rangle^{\otimes N}$ is only achieved with an efficiency of $(1-\varepsilon)^N = 36.6\%$. Correcting those errors turns critical for state preparation or algorithmic tasks and requires the inversion of a $2^N \times 2^N$ matrix. While the matrix construction/inversion procedure can be sped up using tensor formalism, the most computational resource demanding aspect lies in building the probability distribution vector of size 2^N. Moreover, due to finite sampling of the state and wrong estimation of $\varepsilon, \varepsilon'$, $(\bigotimes_i M_i^{-1})\tilde{P}_\psi$ may not be a proper probability distribution. While naive methods such as renormalisation or truncation can give sufficient approximation of P_ψ, more advanced methods such as Bayesian reconstruction may prove to yield more accurate results.

substantial significance and requires technical improvements at each step. The loading stage can be fastened by continuously loading the target register using an atom reservoir [26, 46] or improving the filling efficiency from 50% to $\sim 90\%$ using gray molasses [10]. The scaling of the duration required by an atom-by-atom assembler is another bottleneck which can be addressed by either parallel transport of atoms [29] or optimized arrangement algorithms [24]. Fast imaging can be obtained using resonant imaging of free atoms with high numerical aperture optics [47] or replacing EMCCD used to collect fluorescent photons by cryogenic avalanche detectors [48]. One can ultimately achieve fast high-fidelity non-destructive measurement by using cavity-assisted detection such as in [49]. Overall, the combination of all those techniques should raise the effective repetition rate to the 100 Hz or kHz range.

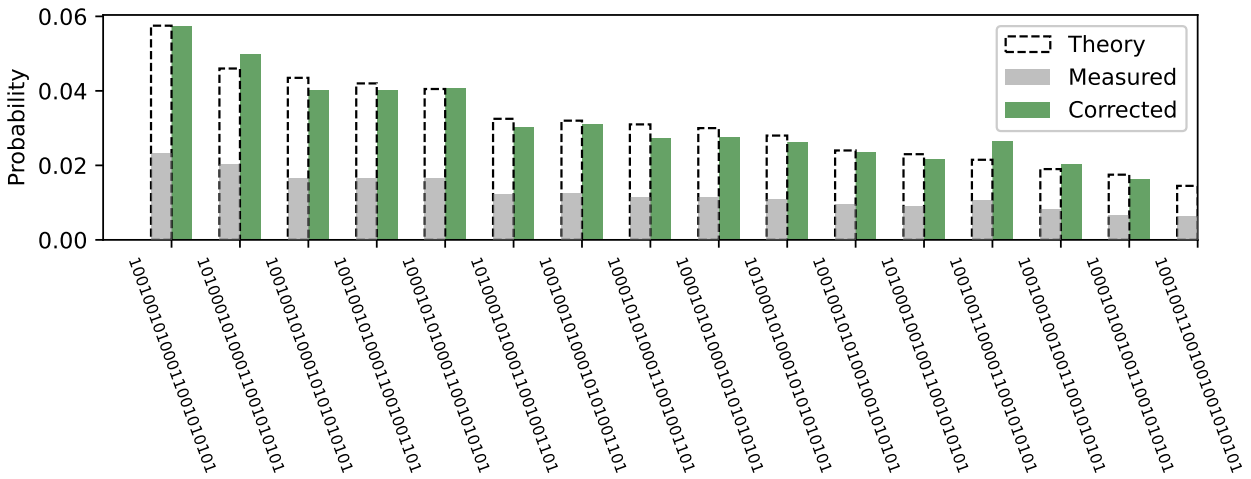


Figure 1.9: A probability distribution P_ψ (dashed) is incorrectly measured. From the measured distribution $\tilde{P}_\psi$ (grey) and known values of detection errors ($\varepsilon = 3\%$, $\varepsilon' = 5\%$), one can approximate the true distribution by a corrected distribution (green) up to 25 qubits in reasonable time. Only the first 16 states with the highest probabilities are shown here.

1.4.5 - Digital mode: gates and circuits

In contrast to analog mode where the evolution is global and continuous in time, the digital mode relies on a discrete set of quantum operations, called *gates*. This mode focuses on a more high level universal vision of quantum computation and allows the user to build hardware-agnostic protocols, called *quantum circuits*, by applying local gates on each qubit independently. Single-qubit gates are essentially rotations (see Eq. 1.5) in the Bloch sphere and two-qubit gates enables to entangle a pair of qubits, usually with the interactions inherent to the considered system. For instance, the X gate flips the states $|0\rangle \leftrightarrow |1\rangle$ of a qubit and is obtained by performing a $\pi-$rotation along the $y-$axis. The controlled-NOT gate (CNOT) acts on a pair of qubits, flipping the state of the second, or target one, i.e. it applies an X gate depending on the state of the first, or control qubit. A quantum circuit is thus characterised by its *depth*, i.e. the number of consecutive layers of gates applied to the qubits, a number to compare with the duration of an analog evolution.

Achieving digital calculations on a neutral atom requires a different basis encoding using long lived hyperfine states $|0\rangle = |5S_{1/2}, F = 1, m_F = 0\rangle$ and $|1\rangle = |5S_{1/2}, F = 2, m_F = 0\rangle$ while still harnessing Rydberg-driven interactions with an additional state $|nS\rangle$ and the scheme described in 1.3.3. Transitions between digital qubits can be driven using a two-photon Raman process $|0\rangle \leftrightarrow |5P_{1/2}, F = 1, m_F = -1\rangle \leftrightarrow |1\rangle$ approximated by a one-photon transition with effective Rabi frequency $\Omega_{01}/2\pi \sim$ MHz [50]. The high-fidelity single qubit rotations already existing for this technology were recently supplemented by two-qubit gates with fidelity reaching 99.5% [51] using optimal control protocols [52], bridging the gap with other digital-focused platforms such as ions and superconducting circuits. This also leads to the prospect of fault-tolerant quantum computing [53], as the effectiveness of error correction codes hinges on maintaining error rates, directly tied to the fidelity of two-qubit gates, below a critical threshold. This threshold is essential to ensure that error correction does not

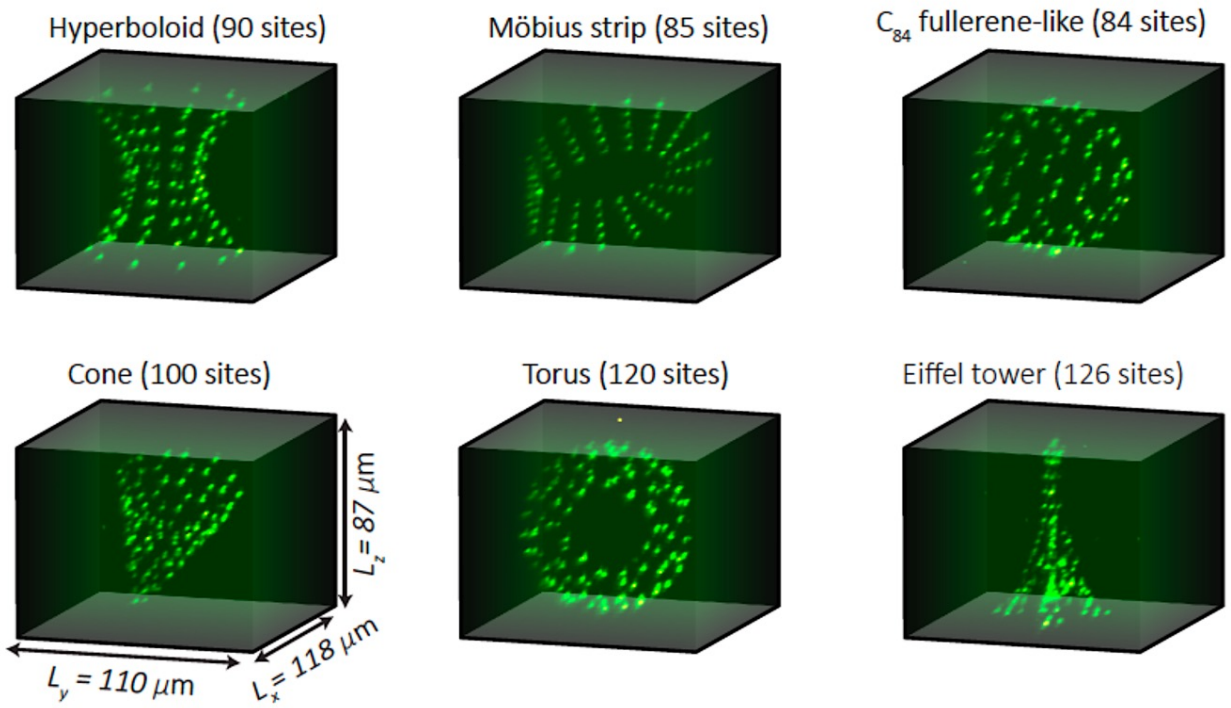


Figure 1.10: **Single-atom fluorescence in 3D arrays** reproduced from [4].

inadvertently introduce additional errors into the system while correcting existing ones. While the applications discussed in later chapters of this thesis will not utilise the digital mode, it's worth noting that the optimal control techniques outlined in chapter 3 offer a means of enhancing the fidelity of certain operations relevant to this field.

1.4.6 - Upcoming hardware improvements

Neutral atom technology is not only closing the gap to other state-of-the-art platforms but also exhibits significant hardware potential for further enhancement. New Rydberg platforms are currently emerging with the notable use of alkaline-earth atoms [54], bi-species arrays of Rb-Cs atoms [55], circular Rydberg states [56, 57] and combination of tweezer technology and cryogenic environments [25]. Some exhibit new features such as local addressing capabilities [58] (see Box. 12), atom shuttling [53], 3D registers [18] [Fig. 1.10], and mid-circuit measurements [59]. We will discuss the use of some of those advances in following chapters.

Box 12: Local addressing in analog mode

To enhance the versatility of analog computing with neutral atoms, the focus can be put on engineering local driving or detuning terms in the Hamiltonian 1.3. Achieving this involves addressing individual atoms with laser light, either resonantly for local driving (Ω-term) or far-off resonance for local detuning (δ-term) through the AC Stark shift effect. Technically, this requires focusing a laser beam onto μm-sized spots overlapped with the atomic register. Controlling the intensity of each spot for maximal flexibility in Hamiltonian parameters poses a significant challenge due to the small scales involved and the need for fast temporal modulation at the ns-to-μs scale. Addressing laser beams using liquid-crystal based spatial light modulators (LCoS-SLM) offers a promising method for spatial shaping. However, these devices face limitations in swiftly switching between configurations. Coupling LCoS-SLM with previously introduced shaping device (EOM/AOM) allows for temporal modulation, inducing in $\hat{H}_{\text{ctrl}}$ the following additional term

$$\hat{H}_{\text{local}}(t)/\hbar = -\Delta(t)\sum_{i=1}^{N}\epsilon_i\hat{n}_i \tag{1.31}$$

with $-20 \leq \Delta(t)/2\pi \leq 0$ MHz and $0 \leq \epsilon_i \leq 1$. The global temporal shape $\Delta(t)$ is controlled by the amplitude modulator, while the LCoS-SLM spatially modulates the light intensity on all sites, providing site-dependent ϵ_i. By applying both a global positive detuning and local negative detunings, we can generate effective local positive and negative detunings. Setting a pattern of ϵ_i values requires the same procedure as changing a pattern of traps, meaning it cannot be changed easily from one shot to another. Notably, it has been used for tasks such as light shifting atoms for initial state preparation [60], adiabatic driving in XY model [37] and measurement of various spin components [58] (see chapter 5). In those works, the detuning term $\delta_{\text{ls}} = \Omega_r^2/(4\delta_r)$ is obtained by lightshifting $|nS\rangle$ from the transition $|e\rangle \leftrightarrow |nS\rangle$ using a 1013 nm laser with parameters Ω_r and δ_r ($\Omega_r \ll \delta_r$).
Alternative approaches for addressability involve acousto-optical deflectors (AODs), either by themselves or combined with LCoS-SLM, allowing neutral atom addressing. Shuttling atoms with AODs to a dedicated zone for addressing with a fixed laser beam is another strategy. Higher refreshing rates for spatial modulation are achieved using micro-electromechanical systems (MEMS) or digital micromirror devices (DMDs) for local Raman transitions or Stark-shift addressing on neutral atoms or ions. Multiplexing approaches, utilising individual modulators for each channel, can be applied in free space with a multi-channel AOM or using photonic integrated circuits (PICs) [61, 62]. The latter case incorporates a steering SLM to match the pattern out of the PIC to the atomic register [63]. Despite the challenges, these techniques provide diverse options for achieving addressability in neutral atom systems.

Part I

Empowering a quantum computer with classical tools

2 - Modelling a QPU : understanding noise contributions for better predictions

Contents

Modelling is a matter of approximation, particularly for complex physical systems such as arrays of trapped atoms, for which keeping an exhaustive record of all the physical effects happening, no matter how minor their actions, would be unthinkable. Real-world phenomena can usually be simplified to abstractions with manageable complexity allowing to first *understand* the underlying mechanisms modifying their apparent behaviours and then *predict* the likelihood of possible outcomes. Current neutral atom quantum computers remain expensive systems to query, both regarding their limited access bandwidth and their relatively low repetition rate. In addition, their noisy behaviours, due to imperfect controls, couplings to the environment or fundamental probabilistic nature, makes the analysis of their outputs a complicated task. Both their high probing cost and the impact of their hardware noises on higher level figures of merit in the NISQ era clearly motivate the need for their modelling by a digital twin. The latter should be faster and cheaper to run than an actual experiment and its representation level should be tunable. Indeed, a crucial aspect of modelling lies in the delicate balance between the accuracy and complexity of the model. By increasingly approximating a system or conversely, by selectively adding effects acting as potential noise sources, one can find the optimal trade-off between computation speed and faithful representation of a noisy QPU, and therefore understanding of its results.

In the following chapter, we will examine the various levels at which we can model a neutral atom quantum computer both in terms of ability to reproduce experimental results and to be run within a reasonable time budget. Those emulation methods will then be used in the following chapters to either explain discrepancies between experimental data and theoretical expectations or to predict scaling of performance of quantum algorithms with the system size.

Some elements from the Pulser description are taken from the following article to which I con-

tributed :

- [64] H. Silvério, et al., "Pulser: An open-source package for the design of pulse sequences in programmable neutral-atom arrays," Quantum, vol. 6, p. 629, 2022.

2.1 - Programming QPUs at the pulse level

A first model useful in the quantum computation paradigm, already discussed in Sec. 1.4.5, consists in abstracting complex quantum operations into a set of fundamental gates. This formalism allows for a consistent framework applicable to various architectures working with qubits, such as superconducting circuits, trapped ions, and neutral atoms. This uniformity streamlines the development of quantum protocols across technologies. However, by specifically focusing on neutral atoms and their programmability, we can start at a lower level of modelling. Our building blocks are not standardised gates but rather atomic positions and laser pulses. The latter give rise to Hamiltonians described in the previous chapter which can easily be programmed and emulated in the following formalism.

2.1.1 - Pulser: a pulse-level library

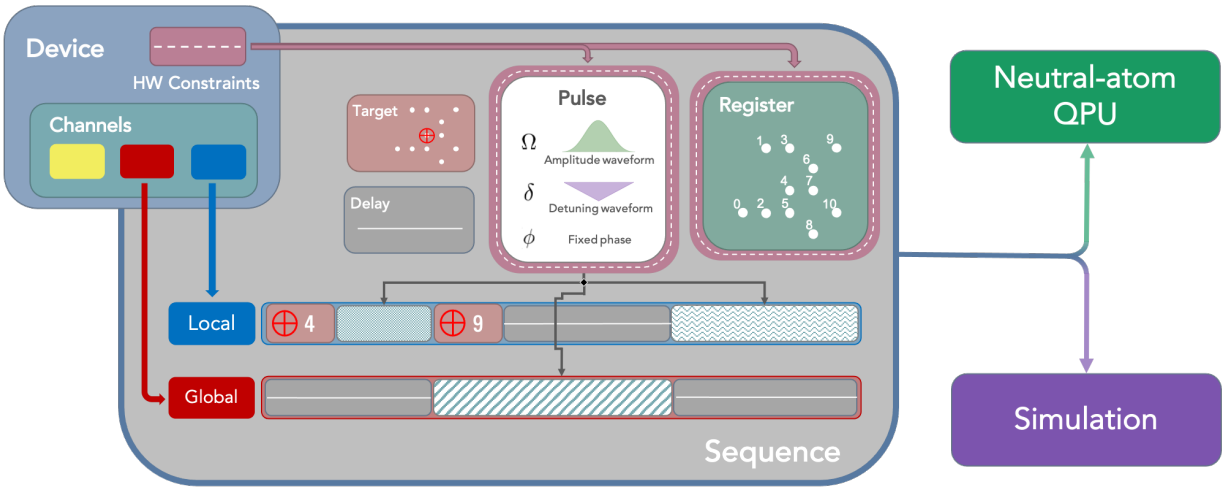


Figure 2.1: **Relationship between the main Pulser classes.** The central object is the `Sequence`, which is linked to a `Device`. The `Device` holds the available `Channels` — which are selected and declared in the `Sequence` — and information of the hardware constraints. These constraints are enforced upon the `Register`, where the neutral-atom array is defined, and upon the `Pulses`. Each `Pulse`, defined by its amplitude and detuning `Waveforms` and a fixed phase, populates the declared channels alongside other commands like `target` — which points local addressing channels to specific qubits — and `delay` — which idles the channel. The resulting `Sequence` can then be sent for execution on the neutral-atom QPU or emulated locally or remotely on a cluster.

Pulser is an open-source Python library, developed by Pasqal, for programming neutral-atom devices at the pulse level [64]. The main goal of this library is to serve as an interface between experienced users and neutral-atom quantum hardware. Using Pulser, users can control all the relevant physical parameters of a pulse-level quantum program. The library also contains several emulation

routines (or `Backends`) for studying and simulating the outcome of pulse sequences for small (using QuTip-based solvers) or larger (using Tensor Network-based methods) systems. We list the main components of Pulser :

(*i*) The `Device` mimicks the QPU itself. It includes a series of specifications characterising the hardware, including the chosen Rydberg level, the available trapping geometries (fixed or arbitrary), the different `Channels` (*iv*) that can be declared as well as constraints on their amplitudes, detunings or `Sequence` (*vii*) duration.

(*ii*) The `RegisterLayout` mimicks the trap pattern created by the SLM. It includes the trap positions and one can be easily created by either providing free-space coordinates or truncating a regular lattice with fixed spacing.

(*iii*) The `Register` mimicks the assembled atomic array. It stores the information about the coordinates of the atoms and their respective ID's, which serve to identify them when applying targeted operations and at measurement. It can be created from selecting a subset of traps of a `RegisterLayout` or directly from scratch by providing free-space coordinates.

(*iv*) The `Channel`s represent the action of the lasers and are organised by addressing (local or global) capabilities and the type of transition (ground-ground, ground-Rydberg, Rydberg-Rydberg) the lasers can be coupled to.

(*v*) `Waveform`s are the basic building blocks of a `Pulse` (*vi*). They can have custom or predetermined shapes, like a ramp or a Blackman window.

(*vi*) `Pulse`s consist of two `Waveforms`, one for the amplitude and one for the detuning. They can be further shifted by a phase. Once a `Pulse` is constructed it has to be added to a `Sequence` indicating which atoms are targeted and which channel will implement it.

(*vii*) The `Sequence` contains the schedule of the pulses in each `Channel`. It is also linked with a `Register` and the `Device` in which it is to be executed. This is the object that can be sent to a real neutral-atom QPU or simulated on a classical computer.

Pulser also incorporates various specialised objects designed for specific tasks, such as limited noisy simulations, local addressing, or qubit masking. Collectively, these features offer maximum flexibility and control over the behaviour of key physical parameters associated with the control and interaction aspects of the Rydberg Hamiltonians, all while adhering to the hardware limitations imposed by the selected device.

2.1.2 - Example of interfacing with a QPU

In the following, we will illustrate how one can access Fresnel, the neutral atom QPU of Pasqal, and run a Ramsey experiment described in Box. 13 to calibrate the detuning and get a first estimate of the dephasing time of the system. We will show how we can parameterise this basic pulse sequence, simulate it using a local emulator and send it to run on a neutral atom QPU through a cloud interface, all using Pulser.

Box 13: Ramsey experiment

A Ramsey experiment involves a straightforward pulse sequence applied to isolated qubits to evaluate the accuracy of detuning calibration. It includes three steps displayed in Fig. 2.2(a). A first resonant $\pi/2$-pulse prepares the state $(|0\rangle + i\,|1\rangle)/\sqrt{2}$ starting from $|0\rangle$, which amounts to putting the qubit on the equator of the Bloch sphere as shown in Fig. 2.2(b). An evolution period with detuning δ and parameterised duration τ rotates the qubit along the equator, preparing the state $(|0\rangle + ie^{i\phi}\,|1\rangle)/\sqrt{2}$ with $\phi = \delta\tau$ being the accumulated phase. Finally a second $\pi/2$-pulse pushes the qubit towards $|0\rangle$ or $|1\rangle$ depending on ϕ. Varying τ enables to measure oscillations at a frequency being the calibrated value of δ. Damping of these oscillations is characteristic of decoherence due to the environment of the free atom. This protocol thus also enables to gain an initial insight of such decoherence whereas a Rabi experiment during which the atom is constantly driven is more sensitive to control fluctuations.

Parameterised Sequence

A significant feature of Pulser used in this example and in the overall thesis lies in the possibility of parameterising `Sequences`. Indeed, many experiments involves scanning a set of parameters possibly over several configurations of atomic registers, while conserving the same overall structure of the applied pulse sequence. For instance, in the Ramsey experiment described above, a waiting time of variable duration is interleaved between two $\pi/2$ resonant pulses. The following code implements such a `Sequence` for a given trap layout, a given `Device` and a fixed value of Ω for the resonant pulses.

```
def parameterised_Ramsey_sequence(trap_layout, device_used, Omega=2*np.pi*1):
    mapp_reg = trap_layout.make_mappable_register(trap_layout.number_of_traps//2)
    para_seq = Sequence(mapp_reg, device_used)
    para_seq.declare_channel("rydberg", "rydberg_global")
    ramsey_duration = para_seq.declare_variable("tau")
    para_seq.enable_eom_mode("rydberg", amp_on=Omega, detuning_on=0)
    para_seq.add_eom_pulse("rydberg", duration= np.pi/2/Omega*1000, phase=0.0)
    para_seq.delay(ramsey_duration, "rydberg")
    para_seq.add_eom_pulse("rydberg", duration= np.pi/2/Omega*1000, phase=0.0)
    para_seq.disable_eom_mode("rydberg")
    para_seq.measure(basis='ground-rydberg')
    return para_seq
```

A first mappable `Register` is created so as to be able to select various qubit registers over which to test this protocol. A `Sequence` takes as an input a `Register` (mappable or not) as well as a `Device` to check during the sequence composition if it remains hardware compatible. A global `Channel` addressing

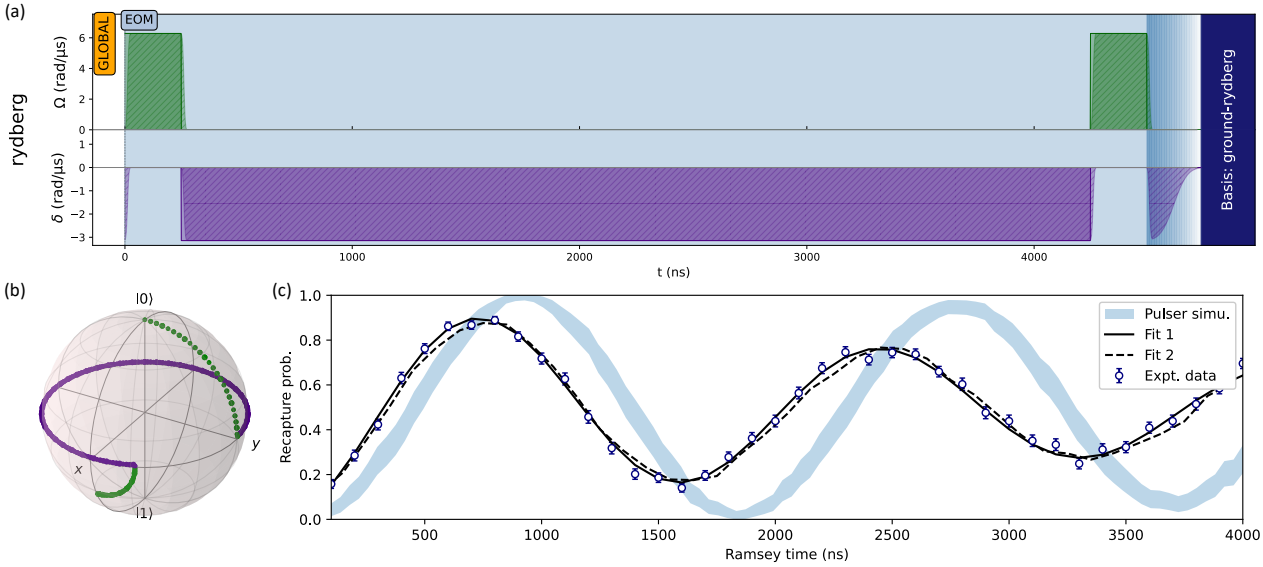


Figure 2.2: **Ramsey protocol** (a) A Ramsey sequence built with Pulser using the EOM mode. (b) Bloch sphere visualisation of the evolution of a qubit under a Ramsey sequence with two resonant $\pi/2$-pulses (green) interspaced by a free evolution time with $\Omega = 0$ (purple). (c) Recapture probability $P_0(\tau)$ obtained when applying the sequence in (a) on an isolated qubit. Experimental data (blue dots) deviates from the simulation with only shot noise (blue area) but can be fitted either (Fit 1) by a damped sine or (Fit 2) by the dynamics obtained using an effective noise model with $\varepsilon = 1\%$, $\varepsilon' = 8\%$, $T_1 = 100$ μs and $T_2 = 4.5$ μs.

the ground-Rydberg transition is labelled and a `Variable` τ is declared. As to create sharp square pulses, we switch to EOM mode, which only allows to work with *on* pulses of fixed amplitude Ω and detuning δ (here 0) or with delays. We then add one $\pi/2$-pulse, delay the sequence for a duration τ, add another $\pi/2$-pulse, switch off the EOM mode and measure the system.

Connecting to the QPU

Once a parameterised `Sequence` has been created, sending it to Pasqal's QPUs is a matter of few lines described below.

```
from pulser_pasqal import PasqalCloud
from pulser import QPUBackend
from pulser_pasqal.backends import RemoteResults

# Provide a connection to the cloud platform using login details
connection = PasqalCloud(project_id, username, password)

# From the connection, one can retrieve available devices and select a trap layout
device_used = connection.fetch_available_devices()["FRESNEL"] # or AnalogDevice
layout = list(device_used.calibrated_register_layouts.values())[0] # Using first Layout

# Create the Parameterised Sequence and select the QPUBackend
```

```
para_seq = parameterised_Ramsey_sequence(device_used, layout)
backend = pulser.QPUBackend(para_seq, connection)

# Several almost isolated atoms are selected to increase the statistics of each shot
isolated_trap_ids = [0,11,13,30,28,32,49,47,60]
qubits = {"q{}".format(i):trap_id for i,trap_id in enumerate(isolated_trap_ids)}
params_list = np.linspace(100,4000,40) # Varying tau between 0.1 and 4 us

# All parameters are gathered inside a list of jobs with 50 shots each
job_params = [{"runs":50,"variables":{"qubits":qubits,"duration":p}} for p in params_list]

# A Batch is created, containing several jobs, and sent to the QPU
results = backend.run(job_params=job_params)
batch_id = results._submission_id

# After some time, the results can be retrieved using the submission id and processed
bitstrings = [res.bitstring_counts for res in RemoteResults(batch_id,connection).results]
```

With credentials and the `Connection` object, one can retrieve informations such as a QPU availability, send batches of jobs to an operative QPU, and retrieve results of a whole experiment, when desired, with a single id.

Fig. 2.2(c) provides the processed results of the above code run on the Fresnel device as well as perfect simulation done with Pulser, easily achieved by switching QPUBackend to QutipBackend.

Fitting noisy results for calibration

The experimental data significantly differs from the perfect simulation due to a miscalibration of δ and decoherence-inducing effects. To calibrate δ we can fit the oscillation with a first effective fit function

$$f_1(t) = Ae^{-t/T}\cos\left(\tilde{\delta}t + \varphi\right) + B \tag{2.1}$$

returning $\tilde{\delta}/2\pi = 0.583(7)$ MHz when $|\delta|/2\pi = 0.5$ MHz and $\varphi/\tilde{\delta} = 0.10(2)$ μs. Measuring $\tilde{\delta}$ enables to recalibrate the detuning and is in fact regularly performed in an automatic way on the Fresnel device, along other calibration jobs. This first fit also gives insight on the noise level as A and B differing from 0.5 reveals the presence of detection errors and a finite value of $T = 4.2(8)$ μs reveals damping due to decoherence. However, it can be complicated to decouple the values of these effective fit parameters in order to extract precise noise levels. The fitted results of several independent and distinct experiments are usually combined. While analytical forms such as Eq. 2.1 are sufficient to analyse the noisy results for some experiments, others necessitate simulating the noisy dynamics of the system by incorporating adjustable levels of noise as conveyed in Sec. 2.3.

Overall, noiseless emulation with Pulser remains relatively fast up to $15-20$ qubits. Simulating a 1 μs long dynamics of $N = 15$ atoms with Pulser takes around a few seconds of classical computation when not considering any additional noise. As N increases, the simulation process gets increasingly time-consuming and approximations are needed to circumvent the exponential scaling in N. The tensor network formalism presented below offers a compact and efficient means of representing quantum states. This enables scaling up the system size to make predictions for larger systems, extending the

reach of emulation to more complex scenarios.

2.2 - Emulation for larger systems with Tensor Networks

In tensor network (TN) methods [65], quantum states are represented as networks of interconnected tensors, capturing the entanglement properties of the system. They are particularly effective in emulating the behaviour of quantum systems using classical hardware, as demonstrated by recent work challenging Google's quantum supremacy claim [66]. TN operations rely heavily on linear algebra, with the size of the matrices involved scaling with the entanglement present in the system being analysed. The relevance of different quantum states in a many-body system's Hilbert space varies, with some being more significant than others in describing physical phenomena. This is particularly relevant in systems where interactions between particles exhibit locality, such as nearest or next-nearest neighbours interactions. Locality constrains the levels of entanglement present in the system and these characteristics make TN well-suited for representing quantum states of the Rydberg systems probed with neutral atom devices.

Since the works in this thesis solely employ widely used TN methods for simulating neutral atom systems, without attempting to develop new ones, we only provide a brief introduction to TN, highlighting the methods which will be used in later chapters.

2.2.1 - Strategy behind TN representations

Let's consider a quantum system composed of N particles with d basis states. The wave function $|\psi\rangle$ describing the system can be written in the following form

$$|\psi\rangle = \sum_{\{s\}} A_{s_1,s_2,..,s_N} |s_1 \otimes \ldots \otimes s_N\rangle \tag{2.2}$$

where defining the value of the d^N complex coefficients of the tensor A is a computationally inefficient definition of the quantum state of a many-body system.

TNs primarily aim at simplifying the representation of such states, while retaining precise description of their predicted entanglement properties. To achieve this, the large tensor A is replaced by a network of N smaller tensors, as illustrated in Fig. 2.3(a), reducing the total number of parameters from an exponential to a polynomial scaling with the system quantities. However, this efficient representation of a quantum many-body state comes with additional complexity. Shaping A to a TN

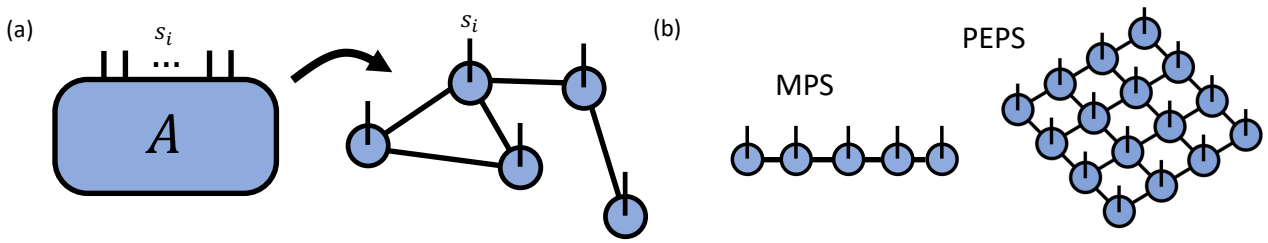


Figure 2.3: **Schematics of tensor networks** (a) A large tensor A is broke down into smaller tensors arranged in a network. (b) Standard MPS/PEPS representation for 1D/2D systems respectively.

introduces extra degrees of freedom, represented by connecting indices among the tensors. These connecting indices encode the structure of many-body entanglement in the quantum state and the number of unique values each index can take quantifies the amount of quantum correlations in the wave function. These indices, often referred to as bonds, have a maximum number χ of possible values known as the *bond dimension*.

Various TN structures have been introduced depending on the system geometry and each exhibits different polynomial scaling with factors such as the system size. The Matrix Product State (MPS) formalism [67] provides efficient representations of 1D systems. The conversion from A to a TN is performed by applying $N-1$ times the Schmidt decomposition. Each of the new tensor $A[i]$ is associated, as displayed in Fig. 2.3(b), with a physical index s_i and two bond indices σ_{i-1} and σ_i. The wavefunction can be written as

$$|\psi\rangle = \sum_{\{s\}} \sum_{\{\sigma\}} A[1]^{s_1}_{\sigma_1} \dots A[i]^{s_i}_{\sigma_{i-1}\sigma_i} \dots A[N]^{s_N}_{\sigma_{N-1}} |s_1 \otimes \dots \otimes s_N\rangle \,. \tag{2.3}$$

The maximum bond dimension χ determines the computation complexity of the simulation, which in this case scales as $Nd\chi^2$. While an exact representation of A can be obtained with $\chi = d^{N/2}$, in most simulation, χ is bounded by a fixed value, serving as a trade-off between approximation and resource consumption. Another structure displayed in Fig. 2.3(b) consists in 2D lattice of TNs and is called Projected Entangled Pair States (PEPS). PEPS TN have primarily been used as ansatz to represent ground states of two-dimensional Hamiltonians [68].

2.2.2 - Useful TN algorithms

Using the TN representation, one can tackle ground state estimation and dynamics solving tasks with algorithms tailored to exploit their simplified structures. This often requires finding optimal contraction methods for combining tensors, incorporating approximations while ensuring stability and convergence to accurate results.

One highly successful algorithm for one-dimensional systems is the Density Matrix Renormalisation Group (DMRG) [69] which can be applied to find low-energy and equilibrium states of 1D Hamiltonians as illustrated in Fig. 2.4(a). It can also be extended to compute excited states or to simulate dynamical, finite-temperature, and non-equilibrium systems. DMRG searches for the lowest energy MPS with a bounded χ. After representing the Hamiltonian as a Matrix Product Operator (MPO), one fixes all but a single MPS tensor and then solves for the minimum eigenvector of the remaining "environment". Individual tensors are then iteratively updated while maintaining stability, often by placing them in canonical form [70].

There also exists efficient algorithms developed for PEPS including routines for optimisation towards low-energy states [71, 72]. Adaptations of the MPS-DMRG techniques in a cylindrical geometry [73] have notably been used to successfully compare experimental results for up to hundreds of atoms [74]. An example of the application of such DMRG algorithm is depicted in Fig. 2.4, where the ground state of an Ising Hamiltonian with nearest-neighbour couplings is obtained for $N = 66$ spins. DMRG is also used in Sec. 5.1.6 to assess order properties of the ground state of the XY Hamiltonian up to $N = 100$ spins.

TN algorithms can also tackle time evolution of large quantum systems. The Time-evolving Block-Decimation (TEBD) method relies on a Trotter-Suzuki decomposition and subsequent approximation

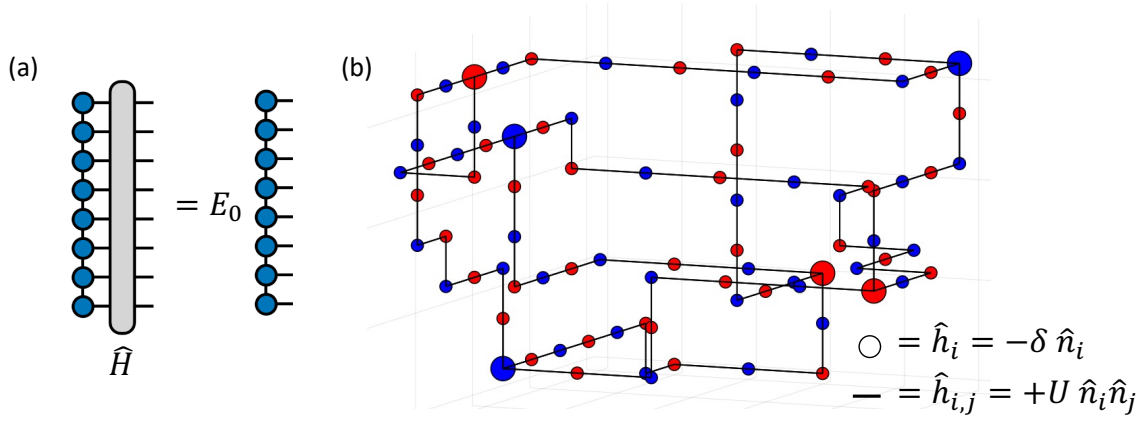


Figure 2.4: **Density Renormalisation Matrix Group.** (a) ground state calculation in TN formalism can be tackled by DMRG algorithms. (b) A 3D system of spins (nodes) interacting with nearest neighbours couplings (edges) can be described by $\hat{H} = \sum_i \hat{h}_i + \sum_{ij} \hat{h}_{ij}$. Its ground state computed using DMRG exhibits alternating upward (red) and downward (blue) pointing spins.

of the time-evolution operator $\exp\{-i\hat{H}\delta t\}$ of a constant Hamiltonian $\hat{H}$. The method exploits the fact that the time evolution operator can be decomposed into a product of local operations, which act non trivially only on a few adjacent sites as showcased in Fig. 2.5(a). TEBD efficiently approximates the action of these operations by applying them to each block separately and then sweeping through the system multiple times, adjusting the boundaries between blocks as needed. Examples of dynamics under constant Ising and XY Hamiltonians computed with TEBD are given in Fig. 2.5(b) and in Sec. 4.5. The local nature of this algorithm however limits its use for more global Hamiltonians. One of the most successful algorithms able to treat long-range interactions, while maintaining a sufficiently small bond dimension, is the Time-Dependent Variational Principle (TDVP) [75, 76]. Its primary aim is to constrain the time evolution to a specific manifold of MPS of a given initial χ by projecting the Hamiltonian onto the tangent space to this manifold and then solves the time dependent SE solely within the manifold.

2.2.3 - Software tools

Many different software packages have been developed in recent years to offer efficient routines for the algorithms mentioned above. A very popular one is ITensor [77], which gathers a strong community of scientists. In addition, it has recently expanded its reach by adding libraries written in the Julia language [78], and by adapting its algorithms to the use of Graphical Processing Units (GPU). Pulser includes a TN-based `Backend` enabling to emulate the dynamics of neutral atom systems up to $N = 100$ within a reasonable computation time, i.e. less than an hour. This emulator [79] utilises the 2-site TDVP method from the `ITensor.jl` package and will be used in Sec. 4.3/6.3.3.

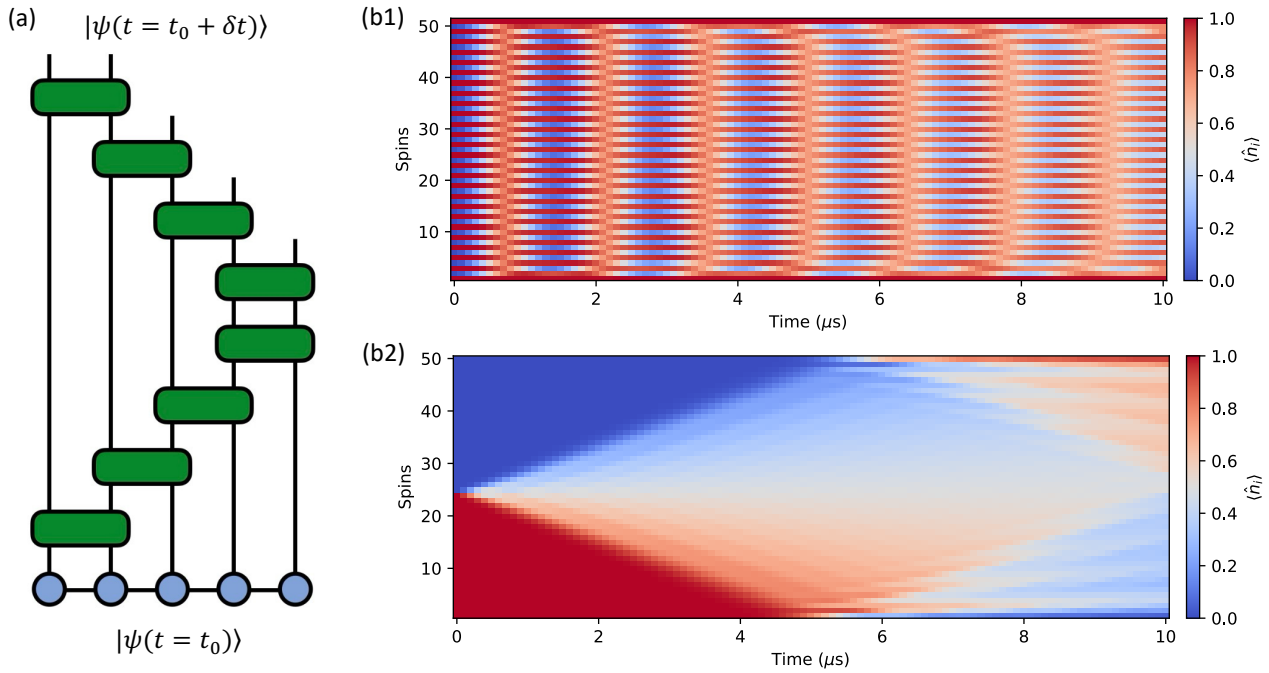


Figure 2.5: **Schematics of TEBD algorithm and examples**. (a) A 1D Hamiltonian with local controls and nearest neighbours coupling can be approximated as constant between t_0 and $t_0 + \delta t$. It is decomposed into small operations (green) sequentially applied to a MPS representation, evolving the state $|\psi(t_0)\rangle$ to $|\psi(t_0 + \delta t)\rangle$. (b) Examples of dynamics of chain of $N = 50$ spins computed with a simple implementation of TEBD using the `ITensor` package for a Hamiltonian constant between 0 and $T = 10$ μs including (b1) controls and Ising interactions or (b2) only XY interactions

2.3 - Emulating noisy dynamics

When benchmarking experimental results with simulated ones, solving the noiseless version of the Schrödinger equation (SE) of Eq. 1.25 might only give qualitative agreement. To bring simulation closer to the experiment, the effects of various physical phenomena acting on the several steps of the quantum dynamics must be taken into account.

2.3.1 - Classification of noises

Most of the noisy processes treated in this thesis have been introduced in boxes of chapter 1 as well as a way to represent them. Here, we will list them according to their related emulation cost. In addition, similarly as in [80], we illustrate their effects on a simple Rabi protocol for which an isolated atom is subjected to a resonant pulse of varying duration, with results displayed in Fig. 2.6.

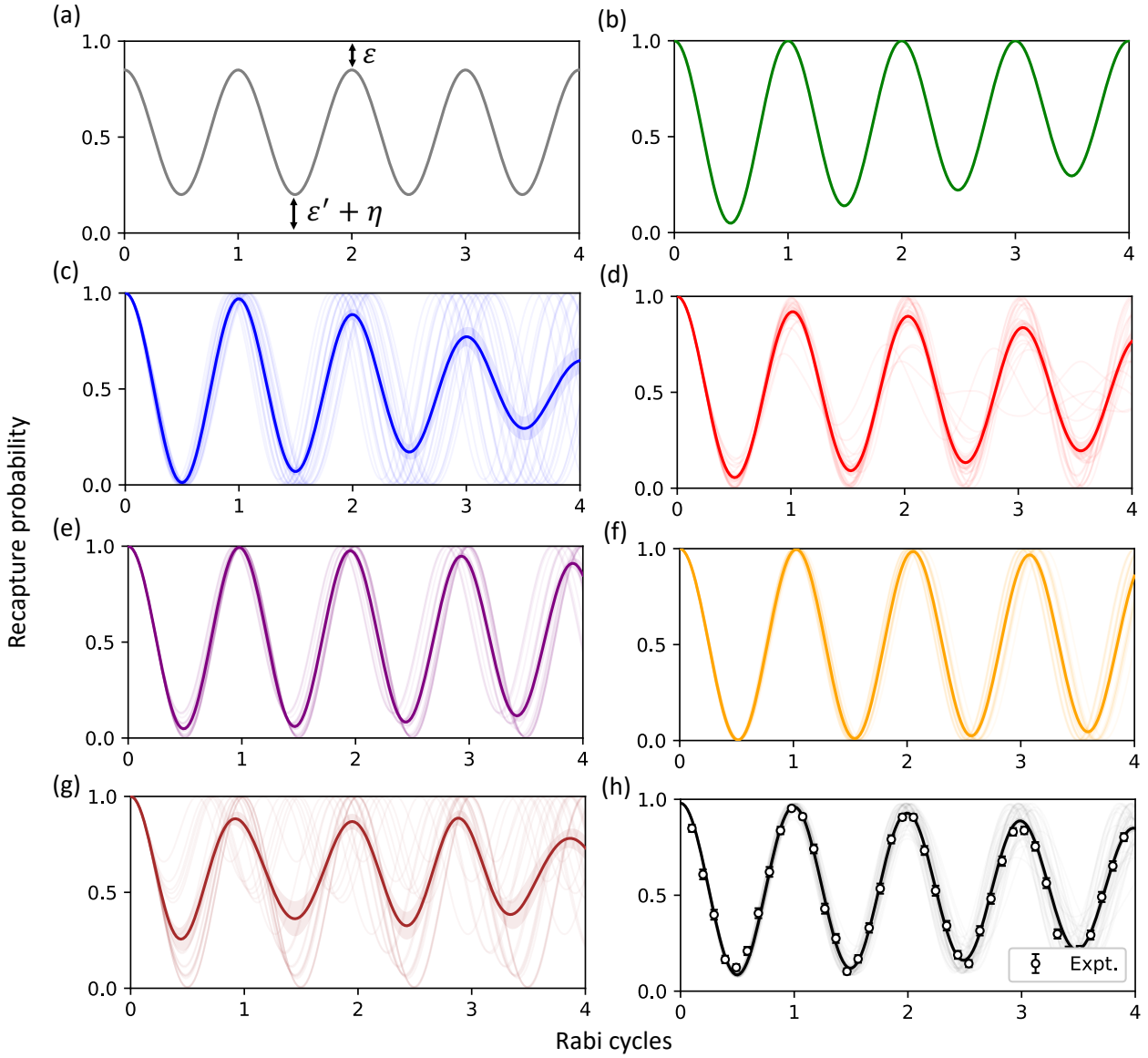


Figure 2.6: **Impact on various noises on Rabi oscillations.** A Rabi protocol is simulated including (a) SPAM errors, (b) decay out of the qubit subspace, (c) Ω random fluctuations, (d) laser phase noise, (e) Doppler shifts, (f) fluctuations of atom positions (g) residual electric field and (h) all of the above with parameter values given in the text so as to match experimental data measured on the Fresnel device.

Cost-free noises

Few hardware imperfections can actually be considered almost "cost-free" to emulate as they only require solving the SE once with a modified Hamiltonian. For instance, incorporating the experimental shape of pulses, as altered by the finite bandwidth of the shaping devices (Box. 2), or as directly measured on the device, can elucidate discrepancies at short times or delayed dynamics, and that with limited additional emulation cost. In the case of large systems, adding the finite waist of the laser beam as position dependant terms in the control Hamiltonian can explain discrepancies between qubit behaviour. Other noises occurring during the measurement phase only impact the measured distributions of bitstrings and are thus handled as post-processing steps. The shot noise on a high-level observable $\hat{O}$ can be obtained by sampling $|\psi\rangle$ M times with n_{shots} and averaging $\langle\hat{O}\rangle$ over the M realisations. It is worth noting that the related cost can rise when repeatedly sampling $|\psi\rangle$ for quite large system sizes. Assessing the impact of detection errors $\varepsilon, \varepsilon'$ (Box. 11) can be achieved by applying the transfer matrix of Eq. 1.30 to the probability distribution P_ψ. However, constructing the transfer matrix can also become resource-intensive as the system size increases. In such cases, it becomes more efficient to probabilistically modify measured bitstrings one by one and utilise a similar sampling-averaging method over M samples of size n_{shots}. Fig. 2.6(a) shows how the detection errors limits the contrast of a Rabi oscillation without inducing any damping (when $\varepsilon, \varepsilon'$ are not time-dependent). All the size-scaling costs mentioned here are usually cheap compared to solving the SE at those sizes, hence their classification as "cost-free" noises.

Master equation formalism

When solving the SE, the quantum system is often considered isolated from its surrounding environment, resulting in its dynamics staying unitary. However, the approximation consisting in neglecting external degrees of freedom becomes invalid when their interactions with the system significantly impact its dynamics, as observed in the damping of oscillations depicted in Fig. 2.2(c). This dissipation of energy into the surroundings can cause uncontrolled decays or incoherent randomisation of the phase and requires to consider non-unitary dynamics of open quantum systems. The latter can be described by the Lindblad master equation (ME)

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}\left[\hat{H}, \rho\right] + \sum_k \gamma_k \left(\hat{L}_k \rho \hat{L}_k^\dagger - \frac{1}{2}\left\{\hat{L}_k^\dagger \hat{L}_k, \rho\right\}\right). \tag{2.4}$$

The density matrix ρ generalises the wavevector formalism describing *pure* states $|\psi\rangle$ to *mixed* states. Mixed states $\rho = \sum_n p_n |\psi_n\rangle \langle\psi_n|$ represent statistical ensembles of possible prepared states $|\psi_n\rangle$ and can also describe states entangled with an environment, where the combined state cannot be impractical to write explicitly. Where $\hat{H}$ describes the unitary part of the dynamics, the set of *jump* operators $\hat{L}_k$ describes the dissipative part of the dynamics, each encompassing an action of the environment on the considered system, happening with rates γ_k. $[\cdot]$ and $\{\cdot\}$ are the commutation and anti-commutation operation respectively.

This formalism enables to effectively take into account, in a two-level system, decay $\hat{L}_1 = |0\rangle\langle 1|$ and dephasing $\hat{L}_2 = \hat{\sigma}^z$ by two coherence times leading to decoherence rates $\gamma_1 = 1/T_1$ and $\gamma_2 = 1/T_2$. This effective model, when combined with the detection errors encompassed by ε and ε', may be sufficient to reproduce experimental data such as the oscillations in Fig. 2.2(c) or the finite contrast in Fig. 4.7.

In some cases, the two-level approximation is not sufficient and the picture must be supplemented by additional levels coupled to the others by coherent processes or spontaneous decay. Taking into account spontaneous emission of the intermediate level in the ground-Rydberg encoding extends the two-level system $|g\rangle, |nS\rangle$ into a four-level system $|g\rangle, |e\rangle, |nS\rangle, |g'\rangle$ where the two-photon transition is addressed by two lasers. $|g'\rangle$ is a *garbage* state not participating in the dynamics and representing states living in the ground state manifold where the atom can decay to. Decays from $|e\rangle$ to $|g\rangle$ and $|g'\rangle$ have branching ratios of 1/3 and 2/3, respectively. Multiplying these by Γ_e we get the individual decay rates. Fig. 2.6(b) showcases the effect of the intermediate spontaneous emission on the Rabi oscillation: as the garbage state becomes more and more populated, less and less population can be driven towards the Rydberg state, damping the oscillation in an asymmetric way.

The finite lifetime of the Rydberg state is usually modelled by $\hat{L}_{nS\to g} = |g\rangle\langle nS|$ with rate $\gamma_{nS\to g} = 1/\tau_n(0K)$. Including the blackbody radiation from $|nS\rangle$ to various Rydberg states requires the addition of another garbage state $|r'\rangle$, effectively figuring the Rydberg manifold. This decay is figured by $\hat{L}_{nS\to r'} = |r'\rangle\langle nS|$ with rate given in Box. 6. The additional emulation cost of adding these states and solving the ME can be staggeringly high as it scales as d^{2N} where d the number of basis states and N the system size. Several physical processes, such as laser phase noise or Doppler effects, can be encompassed inside dephasing terms such as $\hat{L}_2$ in an effective model. Nevertheless, another approach, less scaling-intensive and allowing more detailed simulation of these processes is described below.

Monte Carlo sampling/trajectories

Most of the physical processes happening before or during the dynamics can be treated with a Monte Carlo (MC) sampling approach. Experimental parameters $\boldsymbol{\theta} \in \Theta$ varying between each shot can be sampled M times and for each realisation the related Hamiltonian can be constructed and solved. Averaging the measured observable over large enough M number of repetitions enables to extract the expected behaviour as well as confidence intervals. M should be chosen large enough to prevent a single realisation with parameters sampled far from their mean value from significantly affecting the averaged dynamics. Although the emulation cost scales with M, it can be parallelised since the results of the repetitions do not influence one another. One caveat to consider nonetheless is the presence of sampling noise in the result, which can especially affect derivative calculations using finite-difference methods.

Some of the noises that can be managed with MC sampling include :

- imperfect state preparation, with finite efficiency of optical pumping for instance (see Sec. 1.4.1). For each atom, the outcome of a Bernoulli variable $\mathcal{B}(\eta)$ indicates whether it is considered or not during the dynamics. Fig. 2.6(a) highlights that for an isolated qubit, a finite η causes a contrast reduction in a similar way to ε', the two being then indistinguishable with this protocol.

- shot-to-shot fluctuations of control parameters. Values of Ω and δ are obtained by repeatedely drawing scaling factors from Gaussian variables $\mathcal{N}(1, \sigma^2_{\Omega/\delta})$. Fig. 2.6(c) shows the damping induced by averaging oscillations of various frequencies due to fluctuations of Ω.

- laser phase noise (Box. 3). At each shot, a time-dependent random phase process can be constructed with initial value drawn from $\mathcal{U}(0, 2\pi)$. Fig. 2.6(d) shows the damping induced by

averaging oscillations where the two excitation lasers have random phases processes given by the spectrums of Fig. 1.5(a).

- thermal motion of atoms. Initial positions and velocities (Box. 1) are shifted using Gaussian variables $\mathcal{N}(\boldsymbol{r}_0/\boldsymbol{v}_0, \boldsymbol{\sigma}^2_{r/v})$. Those values are considered when emulating Doppler shifts (Fig. 2.6(e)), time-dependent interactions or the influence of spatial profile in the case of local controls with small waist (Fig. 2.6(f)).

- influence of residual electric field. A value of electric field e drawn from a centered Gaussian $\mathcal{N}(0, \sigma_e^2)$ shifts the Rydberg level $|nS\rangle$ energy by an amount $\alpha_n^{(0)} e^2/2$ (Fig. 2.6(g)) similarly to Doppler effect. In addition, inhomogeneities of the electric field can also induce a spatial variation of this detuning shifts. Effects related to residual electric fields will not be considered in the following.

In addition, decay processes can also be included in this picture by promoting jump rates γ_k from the ME formalism to jump probabilities. At each small time step δt of the dynamics, the evolved state $|\psi(t)\rangle$ can either be randomly projected using one of the specified jump operators $\hat{L}_k$ with probability $\delta p_k(t) \ll 1$ or continues the unitary dynamics: this constitutes the quantum-trajectories MC approach [81]. Averaged over enough realisations the result should converge towards the one obtained with the ME. The MC method is advantageous over the ME approach regarding memory consumption as it only needs to store the state vector $|\psi\rangle$, rather than the entire density matrix ρ. As the number of considered noises increases and more and more configurations of experimental parameters are possible, increasing M enables to better sample Θ but can become cumbersome regarding the emulation cost. When choosing between the ME and the MC methods, it is essential to estimate which quantity will scale the worst between the state size and the number of samples/trajectories required for convergence.

Detailed noise model for Rabi oscillations

Fig. 2.6(h) showcases experimental data of Rabi oscillation performed at $\Omega/2\pi = 1$ MHz and measured on Fresnel with a noisy emulation combining the ME and the MC approaches. This combination is faster than a full MC approach as the system size is $N = 1$. We include the various decays described in 2.3.1 as well as the processes described by the other panels of Fig. 2.6 : $\varepsilon = 1\%$, $\varepsilon' = 8\%$, $\eta = 0.5\%$, $\sigma_\Omega = 3\%$, $T = 20$ μK and laser phase noise with already mentioned spectrum. Such simulation being quite fast to run, it can be incorporated into a fit method so as to extract the various noise parameters directly from this data. However, several of them can not be separated by only Rabi oscillations and the results of other experiments should also be benchmarked at the same time. In the following, we carefully benchmark a STIRAP protocol so as to build a model useful in an optimisation procedure performed in chapter 3.

2.3.2 - Benchmarking a STIRAP protocol

The STImulated Raman Adiabatic Passage (STIRAP) protocol enables to prepare a qubit in $|nS\rangle$ with efficiency η_{STIRAP}. This preparation constitutes an essential step for initialisation in the Rydberg-Rydberg encoding. In the following, we benchmark a STIRAP procedure utilised on the Chadoq device in order to explain with an accurate noise model the finite efficiency experimentally obtained.

STIRAP for robust Rydberg excitation

The STIRAP protocol is detailed in Box. 14 and the resulting dynamics is displayed in Fig. 2.7(b). For this specific implementation with $\Omega_r^{\max} = \Omega_b^{\max} = 2\pi \times 50$ MHz, $T = 0.2$ μs and $\alpha = 1/6$, the accumulated phase $\mathcal{A}_{\Omega_{\text{rms}}} \approx 10\pi$ which yields a finite preparation fidelity of 98.8%. Indeed, the protocol is not perfectly adiabatic and the instantaneous state (solid lines) deviates from the dark state (dashed lines) inducing diabatic errors. In this noiseless model, the fidelity can be increased arbitrarily by extending the protocol duration T to lower diabatic errors. However, accounting for relaxation T_1 and dephasing T_2, a trade-off must be made as increasing T also makes the system more prone to decoherence.

The main advantage of the STIRAP lies in its insensitivity to many parameters, including fluctu-

Box 14: STIRAP protocol

The STIRAP [82, 83] protocol is a two-photon scheme which enables to efficiently transfer atoms from their ground state $|g\rangle$ to a Rydberg state $|nS\rangle$ without populating the intermediate state $|e\rangle$. Diagonalising the three-level Hamiltonian of Eq. 1.6 on the double resonance condition, i.e. $\delta_b = \delta_r = 0$, reveals that the ground state of the system can be written as

$$|\psi_0(t)\rangle = \cos\theta(t)\,|g\rangle - \sin\theta(t)\,|nS\rangle \;\text{ with }\; \tan\theta(t) = \frac{\Omega_b(t)}{\Omega_r(t)}, \tag{2.5}$$

$\theta(t)$ being the *mixing* angle of this *dark* state. Finding a θ evolution respecting $|\psi_0(0)\rangle = |g\rangle$ and $|\psi_0(T)\rangle = |r\rangle$ implies starting with $\Omega_r \gg \Omega_b$ and ending with $\Omega_b \gg \Omega_r$ which results in the counter-intuitive pulse sequence displayed in Fig. 2.7(a). It is worth noting that the transfer only happens between αT and $(1-\alpha)T$ but the finite ramping time of pulse shaping experimental devices extends the protocol duration. The pulses can be spaced closer together by reducing α to limit the lost time $\approx 2\alpha T$ at the expense of the adiabaticity and therefore the efficiency of the protocol. Indeed, for the system to remain at all time in $|\psi_0(t)\rangle$, the following local adiabatic criterion must hold at any time :

$$\Omega_{\text{rms}}(t) \gg |\dot{\theta}(t)| = \frac{|\dot{\Omega}_b(t)\Omega_r(t) - \dot{\Omega}_r(t)\Omega_b(t)|}{\Omega_{\text{rms}}^2(t)} \tag{2.6}$$

with $\Omega_{\text{rms}}^2(t) = \Omega_r^2(t) + \Omega_b^2(t)$. As the protocol is sped up (reducing α or T), the other instantaneous eigenstates of the system can become populated instead of $|\psi_0(t)\rangle$ and the final instantaneous state differs from $|nS\rangle$. Integrating Eq. 2.6 gives $\mathcal{A}_{\Omega_{\text{rms}}} \gg \pi/2$ giving a more global condition and $\mathcal{A}_{\Omega_{\text{rms}}} \propto \Omega_{\text{rms}}^{\max} T$.

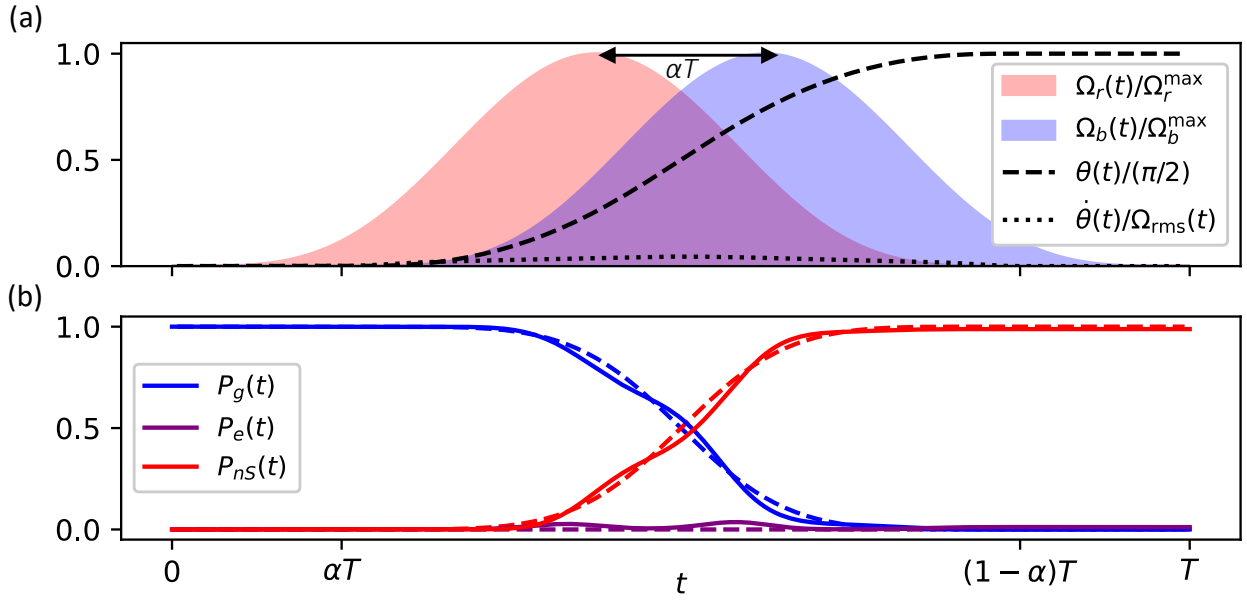


Figure 2.7: **STIRAP protocol and induced dynamics.** (a) Temporal evolution during a STIRAP protocol of the controls amplitude $\Omega_{r/b}$ (red/blue), of the mixing angle θ and of the ratio $\dot{\theta}/\Omega_{\text{rms}}$ assessing the protocol adiabaticity. (b) Temporal evolution of population in basis states $|g\rangle$ (blue), $|e\rangle$ (purple), $|nS\rangle$ (red) when staying in the ground state $|\psi_0(t)\rangle$ (dashed) and obtained when solving the SE (solid).

ations of $\Omega_{r/b}$, δ_r and the synchronisation between the two pulses characterised by αT. To illustrate the latter, the protocol is implemented on Chadoq with $\Omega_r^{\max}/2\pi = 36$ MHz, $\Omega_b^{\max}/2\pi = 85$ MHz, $T = 1.3$ μs and $\alpha = 0.185$ on a square grid of atoms spaced by 12.5 μm. Both the two-photon detuning δ_r and the spacing between pulses τ are varied around their respective central values, 0 MHz and $\tau_0 = \alpha T = 0.24$ μs. Fig. 2.8 depicts the simulated and experimentally measured excitation probabilities. In the noiseless case, the excitation probability plateaus close to 1 for $|\delta_r|/2\pi \leq 10$ MHz and $\tau - \tau_0 \in [-0.15, 0.24]$ μs. For comparison, the excitation fidelity of a $\pi-$pulse (gray line) drops quicker with $|\delta_r|$ and τ when using the same parameters for $\Omega_{r/b}$ and $\delta_b/2\pi = 700$ MHz. The asymmetry around $\delta_r/2\pi = 0$ MHz of the experimental curve on Fig. 2.8(a) is distinctive of a miscalibration of δ_b and helps estimating the non zero value of $\delta_b/2\pi \approx 13$ MHz. Although this protocol should exhibit a preparation fidelity of $\eta_{\text{STIRAP}} = 99.9\%$ at $\delta_r/2\pi = 0$ MHz and $\tau = \tau_0$, the experimental implementation is essentially limited by the altered shape of pulses (-0.6%) and spontaneous emission from the intermediate/Rydberg state (-1.7%). The model also includes shot-to-shot fluctuations of $\Omega_{r/b}$ and δ_r to validate that those almost do not contribute. Accounting for detection errors of the order of $\varepsilon = 2\%$ and $\varepsilon' = 5\%$ reduces the measured simulated efficiency to 93%. However, close to resonance, the experimental efficiency is measured to be around 87%. This discrepancy is believed to be due to unwanted couplings with other states of the intermediate manifold occurring for large values of Ω_b and is reduced when increasing the detuning. This could also be caused by time-dependent depumping effects causing transfers from $|nS\rangle \rightarrow |e\rangle$. Accounting for the latter would require non

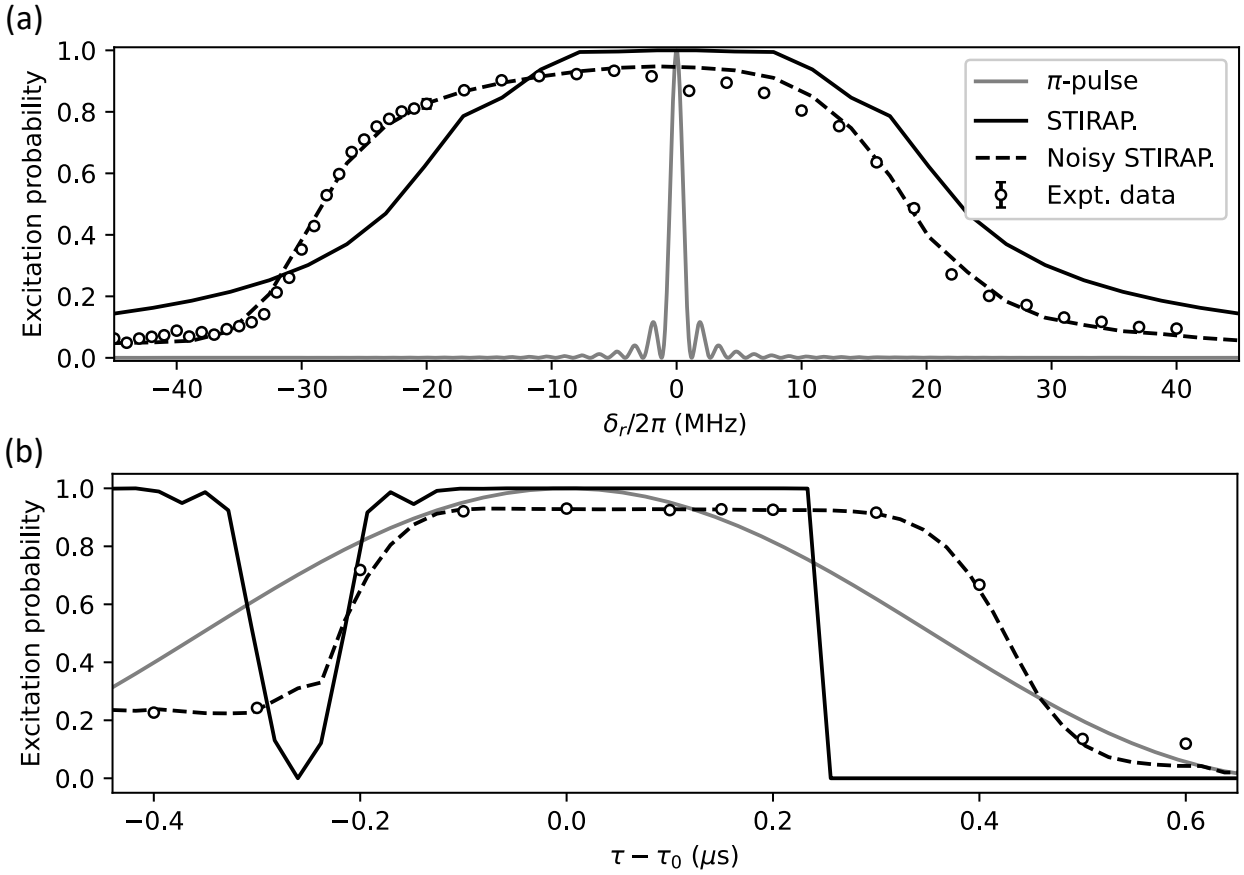


Figure 2.8: **Robustness of STIRAP.** Excitation probability obtained when varying (a) the two-photon detuning δ_r and (b) the duration between the two pulses τ for either (gray) a π-pulse using square shapes for $\Omega_{r/b}(t)$ or (black) a STIRAP protocol. For the STIRAP, experimental data obtained on Chadoq (dots) is benchmarked against noiseless simulation (solid) and noisy simulation (dashed) with parameters given in the text.

trivial time-dependent collapse operators $\hat{L}(t)$ to be included in the simulation. A solution to improve efficiency at resonance (followed by the optimisation procedure performed in next chapter) is to reduce both the amplitude and the duration of the pulses. The model built in this part will be used as a digital twin encompassing the main effects of the device for this protocol.

2.3.3 - Summary of hardware constraints and noise levels

Finally, we gather typical values of hardware constraints and noise levels of a neutral atom QPU in the following Table 2.1.

Specifications		
Symbol	**Unit**	**Value**
Ising Hamiltonian		
n	-	60
$\Omega^{\max}/2\pi$	MHz	2.5
$\lvert\delta\rvert^{\max}/2\pi$	MHz	10
$U^{\max}/2\pi$	MHz	8.9
Traps and atoms		
$N_t^{\max}$	-	61
p_{fill}	%	55
$N_{\max}$	-	30
$d_{\min}$	μm	5
$d_{\max}$	μm	40
p_{move}	%	99
Repetition rate	Hz	$0.8-1$
Laser and pulses		
$\Omega_b^{\max}/2\pi$	MHz	85
$\Omega_r^{\max}/2\pi$	MHz	30
$\delta_b^{\max}/2\pi$	MHz	450
$\lvert\delta_r\rvert^{\max}/2\pi$	MHz	10
w_b	μm	240
w_r	μm	300
w	μm	180
τ_{EOM}	ns	20
τ_{AOM}	ns	100
$T_{\min}$	ns	50
$T_{\max}$	μs	6

Noise		
Symbol	**Unit**	**Value**
SPAM		
ε	%	$1-2$
ε'	%	$5-8$
η	%	<1
Calibration		
Spacing	%	1
Ω	%	2
$\delta/2\pi$	MHz	0.15
Decay		
$\Gamma_e/2\pi$	MHz	1.4
$\Gamma_{r\to g}/2\pi$	kHz	0.7
$\Gamma_{r\to r'}/2\pi$	kHz	0.9
Shot-to-shot fluctuations		
σ_Ω	%	3
T	μK	$10-20$
α	-	$1/100$
$\sigma_r(T)$	nm	$50-100$
σ_z/σ_{xy}	-	5.2
$\sigma_v(T)$	nm/μs	$30-40$
$\lvert\lvert\boldsymbol{k}\rvert\rvert/2\pi$	μm^{-1}	1.2
$\sigma_\delta(T)/2\pi$	kHz	50
Effective model		
T_1	μs	100
T_2	μs	4.5
1D ξ	sites	3

Table 2.1: **Device specifications and noise levels measured on Fresnel.** Values updated in April 2024.

In this section, we have seen how detailed modeling of the noise impacting a quantum system can help understand some unexpected behaviours when implementing a protocol. Several examples of this approach are given in chapters 4, 5 and 7 of this thesis. The numerical twins, once tuned with the quantum device they describe, can provide useful predictions on the dynamics and its limitations but always up to a relatively small system size. TN methods provide, at the other end of the spectrum, simulations of larger systems but including noise in the latter remains a challenging issue.

3 - Optimal control for improved driving protocols

Contents

Improving the desired signal of a quantum experiment conducted on a noisy device can be approached from two angles. Once the influence of low-level noises on higher-level performance has been modelled, specific improvements to the hardware can be undertaken, albeit at sometimes considerable expenses. If no further enhancements are feasible and given that the noise intensities are monitored and kept below a specified threshold, one may explore the development of more *robust* control protocols tailored to operate optimally within these established conditions.

Classical control theory is a unified framework focusing on designing controls to optimise the behaviour of classical dynamical systems within a set of constraints, such as speed, resilience or resource economy. Quantum control theory [84, 85] expands upon these principles into the domain of quantum systems. Originally developed in the context of molecular physics [86] and nuclear magnetic resonance (NMR) [87], it has then been applied to countless quantum computing problems such as finding improved protocols for quantum gates and optimising cost observables in the context of variational quantum algorithms (VQA).

In this chapter, we provide a brief description of useful existing methods in quantum optimal control and apply them to various examples involving Rydberg atoms, including time-optimal and amplitude-robust controlled-Z (CZ) two-qubit gates, detuning robust single-qubit rotations and STIRAP protocols. Then we explain the need for gradient-free efficient optimisation techniques when implementing a VQA, as current neutral atom quantum processors can be viewed as noisy black boxes, expensive to probe.

3.1 - Quantum Optimal Control: methods and examples

3.1.1 - General framework for quantum control

In quantum optimal control (QOC), the target operation to achieve on a system can usually be encompassed inside an operator $\hat{U}_{\text{tgt}}$ which acts on a set of basis states $|\psi_s\rangle$. Such an operation can for instance be a quantum gate such as the ones described in Sec. 1.4.5. To realise such a target, one need to apply a protocol, i.e. a set of time-dependent control functions $u_c(t)$ associated to the control Hamiltonians $\hat{H}_c$. We focus here on bilinear systems whose dynamics are linear with respect to the states and the controls. The total Hamiltonian $\hat{H}(t)$ reads in this formalism

$$\hat{H}(t) = \hat{H}_0 + \sum_c u_c(t)\hat{H}_c, \tag{3.1}$$

with the time-independent part $\hat{H}_0$ usually being referred to as the *drift*. The action of this Hamiltonian on the basis states can be obtained using the related unitary operator (see Eq. 1.28) such that $|\psi_s(t)\rangle = \hat{U}(t)|\psi_s\rangle$. For a protocol lasting a total duration of T, the operation fidelity $\mathcal{F}$ (infidelity $\mathcal{I}$) is obtained by comparing the resulting unitary $\hat{U}(T)$ to $\hat{U}_{\text{tgt}}$, i.e.

$$\mathcal{F} = \left| \frac{\text{Tr}\left\{\hat{U}_{\text{tgt}}^\dagger \hat{U}(T)\right\}}{\text{Tr}\left\{\hat{U}_{\text{tgt}}^\dagger \hat{U}_{\text{tgt}}\right\}} \right|^2, \quad \mathcal{I} = 1 - \mathcal{F} \tag{3.2}$$

Fundamentally, methods in QOC rely on the variational principle to find sets of optimised pulses minimising the infidelity over a set of basis states, while taking into account several constraints. The functional J encompasses this problem with any required constraints implemented via Lagrange multipliers and reads

$$J(\{|\psi(t)\rangle\}, \{u(t)\}) = \mathcal{I} + \sum_c \int_{t=0}^{T} g_{\text{ctrl}}(u_c(t))dt + \sum_s \int_{t=0}^{T} g_{\text{state}}(|\psi_s(t)\rangle)dt. \tag{3.3}$$

g_{ctrl} is a running cost on the control fields and can for instance penalise excessive amplitudes or include knowledge on the control noise to mitigate using the filter function formalism [88]. g_{state} is a state-dependent running cost and can encode time-dependent control targets [89] or penalise population in a forbidden subspace [90].

QOC framework therefore includes three main components : a parameterisation of the control fields, a functional to minimise and an optimisation method. Most algorithms thus iteratively updates the control pulses based on the system dynamics and the desired optimisation goal, relying on gradients of the functional to guide it towards the optimum $\nabla_{u_c,\psi_s} J = 0$. While these ingredients can be selected independently to perform the task, leading to methods such as the ones described in Sec. 3.2.3, they can also be constructed collectively in order to obtain analytical arguments for convergence such as in the Krotov method.

3.1.2 - Krotov method for monotonic convergence

The Krotov algorithm [91] has been applied to various quantum mechanical problems [92, 93] and is based on a rigorous examination of the conditions for calculating the updated controls in order to guarantee a monotonic convergence of J.

Theory

At each iteration k, the Krotov algorithm performs a forward-backward propagation of the states. The co-states $|\chi_s(t)\rangle$ are obtained by evolving the target states $\hat{U}_{\text{tgt}}|\psi_k\rangle$ backward in time under the adjoint Hamiltonian $\hat{H}^\dagger$. The condition for monotonic convergence of J reads

$$\left.\frac{\partial g_{\text{ctrl}}}{\partial u_c(t)}\right|_{(k)} = 2\text{Im}\left[\sum_{s=1}\left\langle \chi_s^{(k-1)}(t)\right| \left(\left.\frac{\partial \hat{H}}{\partial u_c(t)}\right|_{(k)}\right) \left|\psi_s^{(k)}(t)\right\rangle\right]. \tag{3.4}$$

The clever choice made by the Krotov algorithm implies a link between the control constraints and the update scheme such that

$$g_{\text{ctrl}}(u_c^{(k)}(t)) = \frac{S_c(t)}{\lambda_{a,c}}\left(u_c^{(k)}(t) - u_c^{(k-1)}(t)\right)^2. \tag{3.5}$$

The (inverse) Krotov step width $\lambda_{a,c}$ governs the overall magnitude of the update and if too large, $u_c(t)$ changes minimally each iteration, leading to slow convergence. Conversely, overly small values may induce instability in numerical implementation. The update shape function $S_c(t)$ helps enforce boundary conditions on the controls: if both the guessed field at previous iteration and S_c transition smoothly around 0 and T, this characteristic will be retained throughout the optimisation. Combining Eq. 3.4 and Eq. 3.5 gives a first-order update scheme which, from an initial set of controls $u_c^{(0)}(t)$, ensures a monotonic convergence towards optimised ones. In the following, we test an implementation of the Krotov method using the `krotov` Python package [94] to optimise a global two-qubit CZ gate for neutral atoms.

Example : CZ gate with global addressing

An efficient and easily implementable two-qubit controlled gate is a prerequisite for building algorithms in the digital quantum computing paradigm. Several realisations of a CZ gate have been proposed for neutral atom technology, such as the $\pi - 2\pi - \pi$ [95] or the Levine protocols [96]. While the first requires local addressing on each qubit, the second utilises a global scheme with a phase jump to achieve $\mathcal{F} \geq 97\%$. A more recent time-optimal (TO) protocol has enabled to improve $\mathcal{F}$ to 99.5% [51]. Indeed, speeding up a protocol involving Rydberg states enables to reduce the performance losses due to the Rydberg finite lifetime. We reproduce here (and in the following example using GRAPE) the theoretical results obtained in [52].

The objective is to design through optimal control a two-qubit CPHASE gate acting on a $3-$level system consisting of two hyperfine $|0\rangle$, $|1\rangle$ and a Rydberg state $|nS\rangle$ by shaping the following time-dependent controls : the detuning δ or the phase φ. In order to be as fast as possible, $\Omega(t)$ remains constant at a maximum value Ω_{max} and the two qubits interacts when in $|nS\rangle$ through an interaction term with strength $U = 10\Omega_{\text{max}}$. The target operator reads $\hat{U}_{\text{tgt}}(\alpha) = |00\rangle\langle 00| + e^{-i\theta}(|01\rangle\langle 01| + |10\rangle\langle 10|) + e^{-i(2\theta+\alpha)}|11\rangle\langle 11|$, with θ being a free-parameter adjustable with a 1-qubit phase gate. For $\alpha = \pi$, $\hat{U}_{\text{tgt}}$ implements a $CZ-$gate.

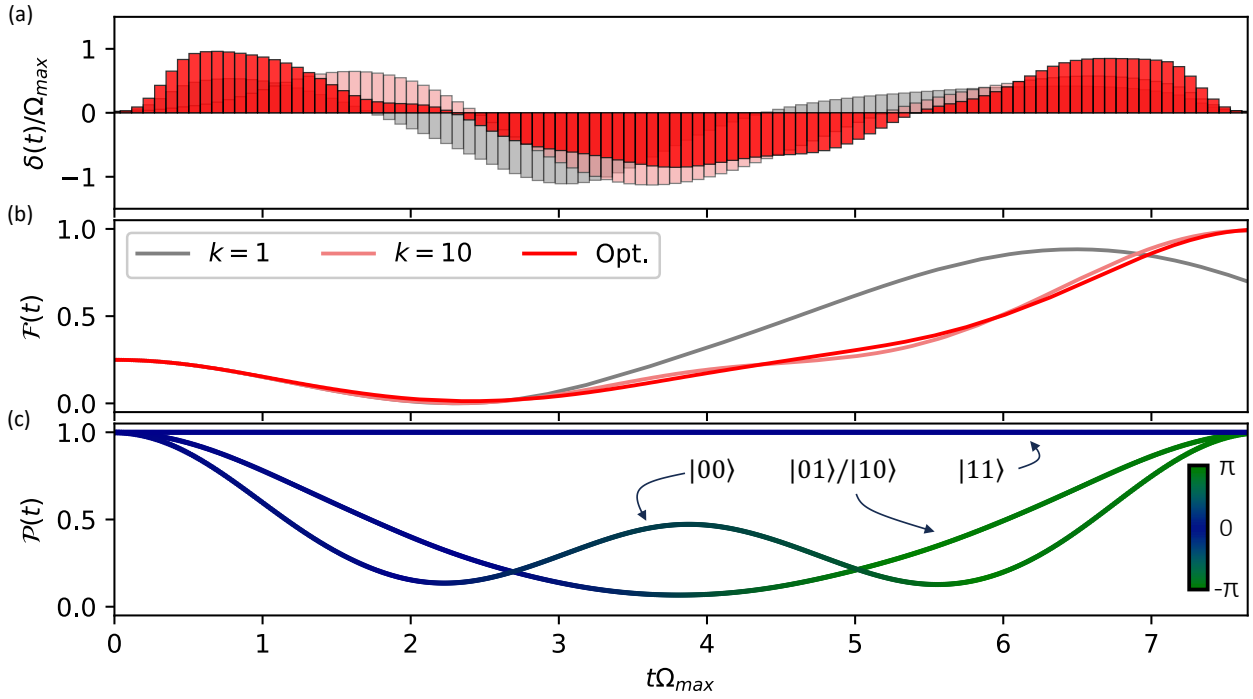


Figure 3.1: **Krotov method applied to global CZ gate.** (a) Evolution of the optimised control δ at various iterations, including (grey) $k = 1$, (pink) $k = 10$ and (red) $k = 60$. The pulses are discretised in $n_t = 200$ time steps. (b) Temporal evolution of the fidelity of the operation for various iterations. (c) Temporal evolution of the probabilities of each basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ under the optimised protocol. The solid lines are coloured according to the instantaneous phase of each state (blue→green).

The analytical monotonic convergence result of Krotov is only valid for continuous controls and one requires to discretise their shapes in numerical optimisations. However a sequential scheme, well described in [94], can still guarantee its validity with controls as piece-wise constant functions over thin enough time intervals. For this specific implementation using the Krotov method, we optimise on δ, set φ to 0 and fixes $T \approx 7.612/\Omega_{\max}$ μs as this is the optimal time found in [52] (compared to $T \approx 8.585/\Omega_{\max}$ μs in [96]). We initialise the optimiser with a constant shape for $\delta(t)$ and select a flattop update shape $S(t)$ with rising time 50 ns. After 60 iterations, the algorithm outputs $\mathcal{I} \approx 10^{-3}$. Fig. 3.1(a) showcases the updates on the initial constant shape of $\delta(t)$ after 1, 10 and 60 iterations, settling on an oscillatory behaviour reminding of the phase profile found in [52]. Each iteration of the algorithm improves the final fidelity as exemplified in Fig. 3.1(b) where $\mathcal{F}^{(k)}(t)$ is plotted for the corresponding pulses. We can check with Fig. 3.1(c) that for the optimised pulse the phase of the basis states implements $U_{\text{tgt}}(\pi)$ up to a global phase of π. While further improving the fidelity can be ensured by Krotov monotonic convergence, the cost of additional iterations can be prohibitive. In addition, a thin enough time discretisation requires a high value of λ_a, meaning that controls are updated by only small values. Another challenge thus arises in the balanced selection of the step

width, which can accelerate convergence but lacks a systematic tuning method beyond trial and error.

3.1.3 - GRadient Ascent Pulse Engineering (GRAPE)

While Krotov method devises a sequential time discretisation through a clever construction, the GRAPE method circumvents this challenge by initially discretising the functional before employing variational calculus. The GRAPE method was first introduced to design NMR pulse sequences [97] but its use extended to quantum computing, included in neutral atom technology [98]. As it utilises a direct analytical expression for the gradient by discretising the continuous control fields, GRAPE can efficiently find a suitable solution in the parameter space with improved convergence speed.

Theory

For a time interval discretised in n_t steps of size Δt, the resulting unitary $\hat{U}(T)$ can be written as product of constant operators $\prod_n \hat{U}_n$ with $u(t) = u_n$ for $t \in [(n-1)\Delta t, n\Delta t]$. The gradient of a functional, such as $\mathcal{I} = 1 - |\langle\psi_{\text{tgt}}|\psi(T)\rangle|$, can thus be calculated exactly at iteration $k-1$ with

$$\begin{aligned} \left.\frac{\partial \mathcal{I}}{\partial u_n}\right|_{(k)} &= -\frac{\partial}{\partial u_n} \langle\psi_{\text{tgt}}| \hat{U}_{n_t}^{(k-1)} \dots \hat{U}_n^{(k-1)} \dots \hat{U}_1^{(k-1)} |\psi(0)\rangle \\ &= -\left\langle \chi^{(k-1)}(n\Delta t) \right| \frac{\partial \hat{U}_n^{(k-1)}}{\partial u_n} \left| \psi^{(k-1)}((n-1)\Delta t) \right\rangle \end{aligned} \tag{3.6}$$

On the contrary to Eq. 3.4, the gradient computation only requires knowledge of the previous iteration and all control parameters u_n can be updated at once, making GRAPE a *concurrent* scheme instead of a *sequential* one like Krotov. Meeting halfway, there also exist *hybrid* schemes that interleave concurrent and sequential calculation of the gradient [99]. Using classical gradient descent optimisations technique such as Broyden-Fletcher-Goldfarb-Shannon (BFGS) method, one can then iteratively locate the optimum of the functional. One downside of GRAPE compared to Krotov is that the optimised controls may not be approximately continuous, complicating in that case their implementation on hardware. In the following, we test an implementation of the GRAPE method using the gradient-based optimiser of the Boulder Opal software [100, 101] which relies on an efficient tensor formulation of the objects at hand to provide powerful optimisation. Building on the problem of finding the TO CZ gate, we then apply this formalism to build a protocol mitigating the influence of amplitude noise in the controls.

Example : improving the robustness of CZ gates

We perform a similar optimisation procedure as for the Krotov case in order to minimise the infidelity of a CZ gate on two qubits. We perform several optimisations for various values of $4 \leq T\Omega_{\max} \leq 10$ in order to locate the time optimum. On the contrary to the `krotov` package, we can easily constrain δ to remain within $|\delta(t)|/\Omega_{\max} \leq 5$ (mimicking the ratio reachable on Fresnel). The pulses are split into $n_t = 100$ steps and smoothed with cut-off frequency $n_t/2T$. We perform the optimisation on either φ or δ separately while keeping the other to zero.

Fig. 3.2(a) shows the infidelity obtained at each optimisation with a clear drop around $T_{\text{opt}}\Omega_{\max} = 7.612$ rad giving the TO gate protocol. Below T_{opt}, the gate time is too short to obtain the proper phase accumulation of a CZ, while above, each optimisation procedure can produce a working pulse where the minimal infidelity obtained only depends on convergence criterion of the procedure. The TO

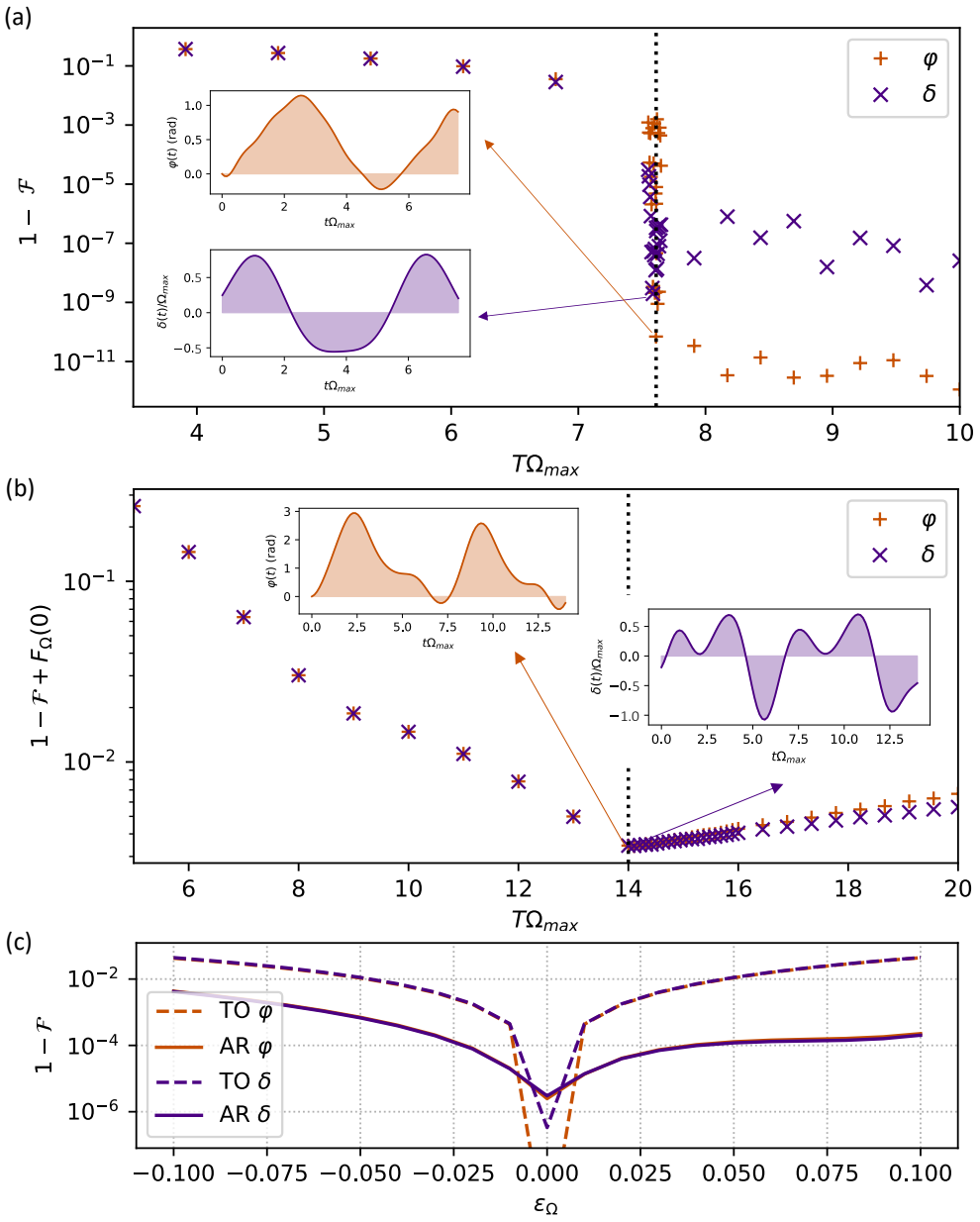


Figure 3.2: **GRAPE method for finding time-optimal and amplitude-robust protocols for the CZ gate.** (a) For each total duration T, the final infidelity is plotted when optimising (purple) $\delta(t)$ or (orange) $\varphi(t)$. The TO pulses obtained at the infidelity drop are given in the inset. (b) Similar plot for the AR pulses. (c) The robustness of the infidelity against amplitude fluctuations ε_Ω is given for the four pulses highlighted in (a) and (b).

shape in detuning closely resembles the one obtained in Fig. 3.1(a) while the phase shape reproduces the result of [52].

Apart from finding optimal pulses in the noiseless case, QOC can also tackle the mitigation of

experimental noise, such as fluctuations of laser intensities. This can be translated into a quasi static amplitude noise $\Omega_{max} \rightarrow (1+\varepsilon_\Omega)\Omega_{max}$ with ε_Ω of the order of a few %. In [102], the authors apply the same framework to the task of finding a TO amplitude-robust (AR) pulse. We incorporate the sensitivity of the controls to this amplitude variation via a corresponding noise operator $\propto \hat{\sigma}_x$ and use the filter function formalism [88, 100], adding $F_\Omega(0)$ to the functional in order to tailor the procedure to produce a robust pulse. Fig. 3.2(b) shows the minimised value of obtained for various total durations with a clear inflection around $T\Omega\text{max} = 14$ rad giving the fastest amplitude-robust gate protocol. The two approaches, optimising on φ and δ are giving similar results and the optimised shapes are comparable to the ones obtained in [102]. Fig. 3.2(c) displays the robustness of the highlighted protocols when applied with $\Omega_{max}(1+\varepsilon_\Omega)$. While the TO protocols performs very well at $\varepsilon_\Omega = 0$, the infidelity quickly increases to 10^{-2} at $\varepsilon_\Omega = 5\%$. Although less efficient in the noiseless case, the AR protocols allow to gain at least an order of magnitude for $|\varepsilon_\Omega| \leq 10\%$. Such approaches on building robust pulses can be extended to various other noises, such as detuning fluctuations as tackled in the following section.

3.1.4 - Chopped RAndom Basis (CRAB) optimisation

Both Krotov and GRAPE rely on computing or estimating the gradients of the functional relative to the controls. However, relaxing on the condition of having access to those gradients, one can also parameterise controls in more various ways including sums of basis functions such as in CRAB.

Theory

The essence of CRAB [103] lies in the specific choice of the parameterisation of the controls in terms of a truncated randomised basis of functions recasting the problem from a functional minimisation to a multi-variable function minimisation, manageable by direct-search methods such as the ones described in Sec. 3.2.3. In this case, each control of Eq. 3.1 can be written as

$$u_c(t;\boldsymbol{\alpha}_c) = \sum_m \alpha_c^m f_m(t) \tag{3.7}$$

where f_m is a family of functions such harmonic ones characterised by amplitude, frequency and phase or a sequence of Gaussians. The optimisation is then performed over the $\boldsymbol{\alpha}_c$. While for some parameterisation, a gradient can still be calculated at relatively low cost, for more complex basis, the use of gradient-free methods requires to set an upper bound on the size of the family to limit the parameter space. Without knowledge on the gradients, the optimisation usually converges slower, especially for large $|\boldsymbol{\alpha}_c|$, but can be sped up by slightly randomising the hyperparameters of the basis functions at each iteration. Constraints on the control pulses can be transformed into constraints over the $\boldsymbol{\alpha}_c$ and robust pulses can be found by averaging the process over a broad array of possible noise values.

Example : $\pi/2-$pulse robust to detuning shifts

Another strong building block of digital quantum computation is the ability to perform resilient single-qubit gates, i.e. rotations such as described in Eq. 1.5. Due to environmental noise or instabilities in the drive, the detuning component can suffer from random shifts in its value. This problem was tackled in the context of NMR with the authors of [104] introducing general rotation pulses, called Band-selective Uniform Response Pulses (BURP) and optimised so as to be robust under shot-to-shot fluctuations of the detuning. The latter utilises the following harmonics decomposition for the driving field

$$\Omega(t;\boldsymbol{\alpha}) = \frac{\theta}{T}\left[\alpha^0 + \sum_{m=1}^{m} \alpha^{2m-1}\cos(\omega_m t) + \alpha^{2m}\sin(\omega_m t)\right] \text{ with } \omega_m = 2\pi m/T \tag{3.8}$$

with T the protocol total duration and θ the rotation phase. While continuity of the controls are straightforward by construction, we can implement constraints such as $\Omega(t=0;\boldsymbol{\alpha}) = \Omega(t=T;\boldsymbol{\alpha}) = 0$ $\rightarrow \alpha^0 + \sum_{m=1} \alpha^{2m-1} = 0$ by incorporating them into the global function to optimise. We are especially interested in applying single qubit x rotations to the system, i.e. $\hat{U}_{\text{tgt}} = \hat{R}_{\boldsymbol{x}}(\theta)$ with for instance $\theta = \pi/2$.

This parameterisation enables to construct several pulses, displayed in Fig. 3.3(a), implementing $\hat{U}_{\text{tgt}}$ such as :

- the square pulse, with $\boldsymbol{\alpha} = [1]$,
- the Blackman pulse, with $\boldsymbol{\alpha} = [1, -1/(1-a), 0, a/(1-a), 0]$ with $a = 0.16$,
- and the Universal BURP pulse, with the $1 + 8 + 8 = 17$ parameter values given in [104].

We implement a CRAB method to find an optimal pulse using as many parameters and as robust as the BURP instance while speeding it up. We use the tensor formalism of Boulder Opal to compute the dynamics occurring with $\Omega(t;\boldsymbol{\alpha}^{(k)})$ at various values of constant detuning δ_0 up to $\Omega_{\max}$, averages the obtained infidelity while including constraints such null values at the boundaries and time-optimality and perform the optimisation using the gradient-free method available in Boulder Opal.

The optimised pulse given in Fig. 3.3(a) is 8 times longer than a simple square pulse but 40% faster than the BURP one. Its benefit is highlighted in Fig. 3.3(b) where the deviation from a perfect $\pi/2$ rotation along x averaged over values of detuning shifts up to δ_0 is plotted from the various pulses. The square pulse excels primarily in nearly noiseless scenarios, whereas the Blackman pulse initially lags behind in this domain but compensates a bit for it when larger shifts are occurring. While both BURP and our optimised pulse are beaten by the square one at $\delta_0 = 0$, if the shifts can be as high as $\Omega_{\max}$, they help gaining two orders of magnitude in robustness. The shorter optimised pulse also outperforms BURP from one order of magnitude in the noiseless case but is slightly less efficient when $\delta_0/\Omega_{\max} \geq 0.4$. The dynamics of the Blackman and of the optimised pulse are given in Fig. 3.3(c1) and (c2) respectively for various values of shifts. As the simpler pulse can not prevent the qubit state from increasingly rotating along the z-axis, the optimised pulse leads to distinct dynamics, ensuring that regardless of the shifts, the qubit ends up near the target state. Another interesting problem is to consider pulses robust to interactions induced-shifts when wanting to perform single-qubit rotations in an array of interacting qubits.

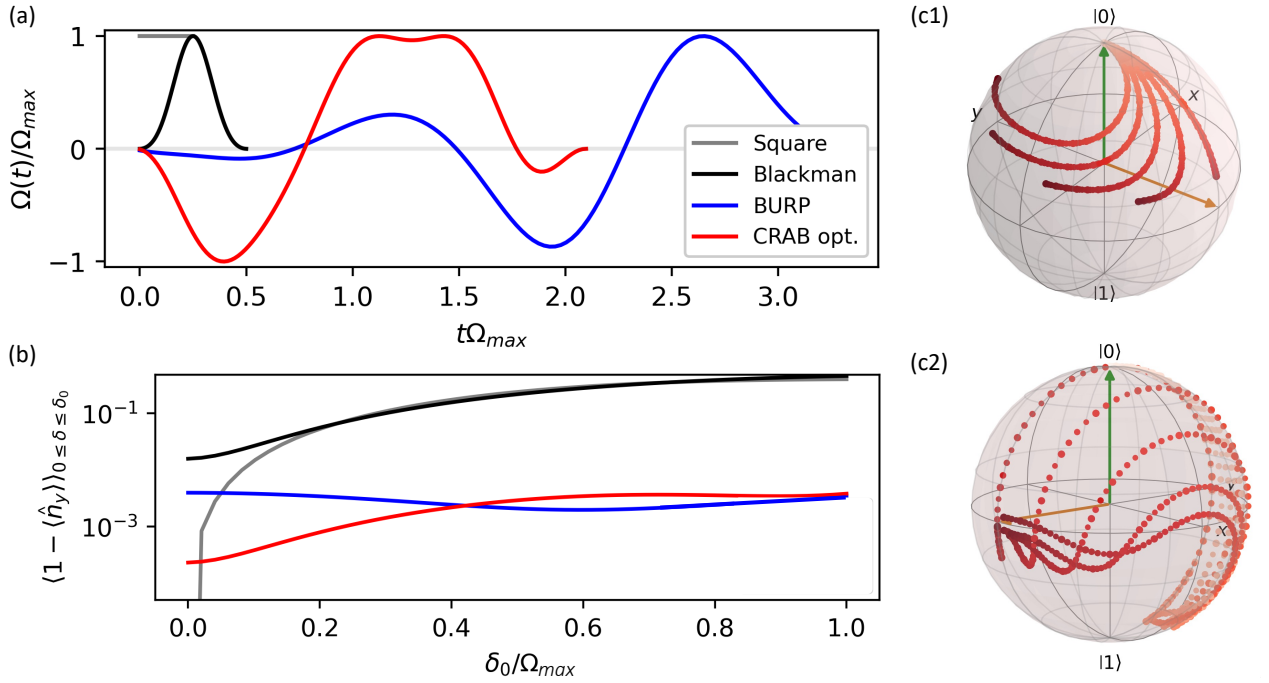


Figure 3.3: **CRAB method for robust single-qubit $\pi/2-$pulse.** (a) The harmonics parameterisation enables to construct (grey) a square pulse, (black) a Blackman pulse, (blue) the BURP pulse and (red) the CRAB optimised pulse. (b) Deviations from the perfect rotation obtained when using the pulses given in (a) and averaged over realisations with detuning shifts $\delta \in [0, \delta_0]$. (c1) Bloch sphere representation of the qubit dynamics (increasingly red points) with the Blackman pulse, starting from the initial state $|0\rangle$ (green) and targeting $(|0\rangle + i\,|1\rangle)/\sqrt{2}$ (yellow). (c2) Same with the CRAB optimised pulse.

3.1.5 - Improving STIRAP efficiency with QOC

The latter examples demonstrates the utility of QOC methods for designing time-optimal and robust gates for digital computing with Rydberg atoms. In addition, QOC can also tackle the optimisation of protocols like the STIRAP (see Box. 14) benchmarked in Sec. 2.3.2.

As a reminder, the system at consideration is a 5-level system with basis states $|g\rangle, |e\rangle, |nS\rangle, |g'\rangle$ and $|r'\rangle$ where the efficiency of the targeted transfer from $|g\rangle$ to $|nS\rangle$ can be encompassed in the fidelity $\mathcal{F}(t) = |\langle nS|\hat{U}(t)|g\rangle| = \mathcal{P}_{nS}(t)$. Anticipating on the experimental testing of an optimised STIRAP protocol, we can define a measurable fidelity:

$$\mathcal{F}_m(t) = 1 - \sum_{k=g,g'} \mathcal{P}_k(t) = \sum_{k=e,nS,r'} \mathcal{P}_k(t) \tag{3.9}$$

since differentiating between $|g\rangle$ and $|g'\rangle$ (or between $|e\rangle$, $|nS\rangle$ and $|r'\rangle$) is not easily achievable with the used measurement scheme. This fidelity is useful for benchmarking simulations and the device. However when the STIRAP is only the first step of a more complex protocol, we are more interested in the theoretical $\mathcal{F}(T)$. With the model described in Sec. 2.3.2, the experimental STIRAP protocol, given in the right inset of Fig. 3.4, achieves $\mathcal{F}_m(1.3\mu s) = 0.9756$ when correcting for detection errors.

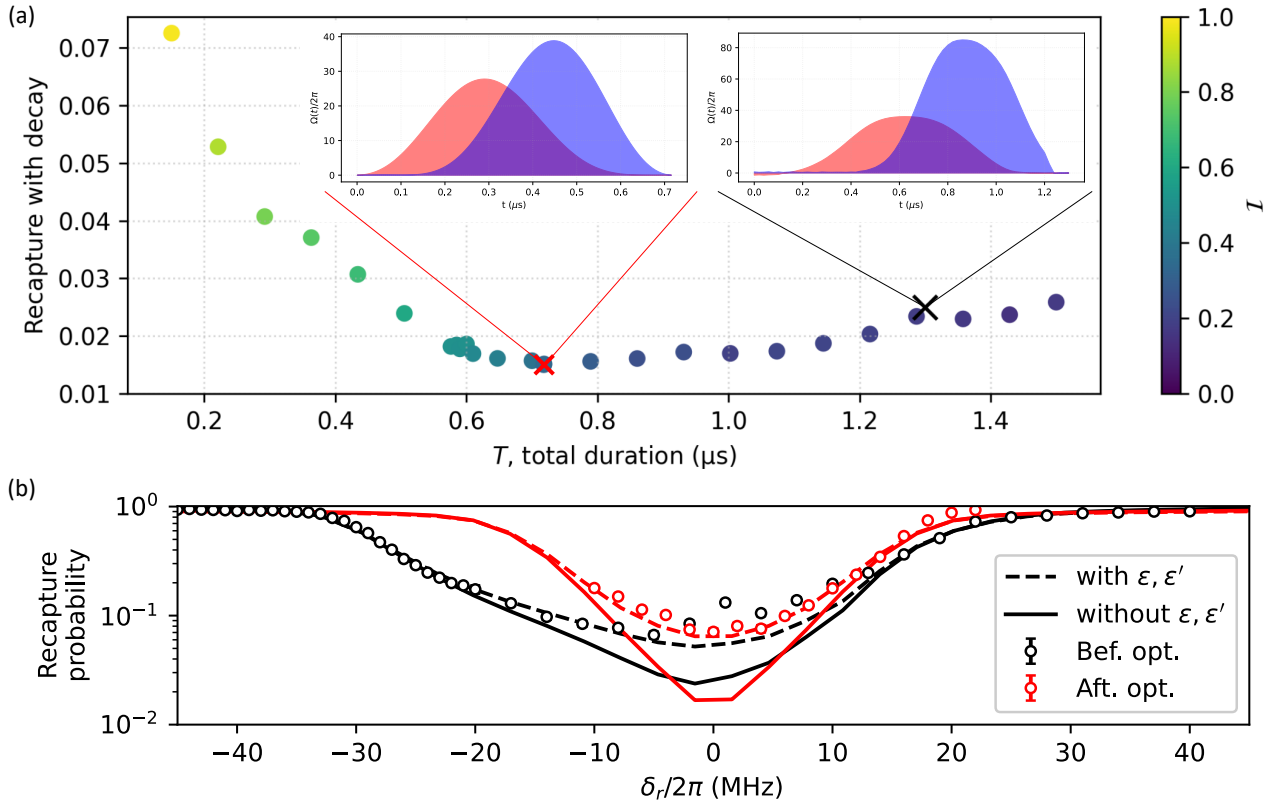


Figure 3.4: **Optimisation of a STIRAP protocol with quantum optimal control.** (a) For each total duration T, an optimised protocol is obtained, yielding a noiseless infidelity $\mathcal{I}$ (coloured points). The measurable fidelity $\mathcal{F}_m$ is obtained when emulating the protocols with a realistic noise model. (insets) Experimentally used and time-optimal STIRAP protocols with the time dependent amplitudes of $\Omega_{r/b}(t)$. (b) Recapture probability $(1 - \mathcal{F}_m(T))$ obtained when varying the two-photon detuning δ_r for both protocols highlighted in (a). Experimental measurements (dots) are benchmarked against noisy simulations with detection errors (dashed) and the corrected recapture probabilities are also given (solid).

This sets a baseline for our optimisation algorithm both in terms of pulse duration and fidelity to reach. We focus on shaping the two driving fields in the two-photon scheme, $\Omega_{r/b}(t)$, with a GRAPE-like method using Boulder Opal while also letting δ_b as a free parameter to optimise within reasonable bounds $|\delta_b|/2\pi < 15$ MHz. The cost to minimise is the following :

$$\mathcal{J} = 1 - \mathcal{F}(T) + \frac{1}{T}\int_{t=0}^{T} |\langle e|\psi(t)\rangle|^2. \tag{3.10}$$

While the first term accounts for the closeness to the target, the second implements a cost penalty for populating too much the intermediate state, as it could lead to unwanted decays. $\Omega_{r/b}$ are built using piece-wise constant pulses with $n_t = 100$ segments, filtered by a low-pass of cutoff frequency $\propto 1/\tau_{AOM}$ and multiplied by a cosine envelope to ensure $\Omega_{r/b}(0) = \Omega_{r/b}(T) = 0$. In order to reduce

the influence of the finite lifetime, we want to implement a TO protocol. To avoid solving a master equation at each step of the optimisation (and also because gradient estimation is costly within this framework), this time optimisation is either done by adding another penalty term $\propto T$ (at the cost of having to fix another hyper parameter) in $\mathcal{J}$ or by performing several optimisations each at a fixed time T and emulating the corresponding dynamics with the master equation formalism to choose T_{opt}, the time enabling to achieve the target while not letting too much time for the decoherence effects to strongly affect the system.

As the total duration available increases, the optimiser reaches improving values of $\mathcal{I}$ in the noiseless case. However, when using the noise model derived from the experimental benchmark, including various decay channels for instance, an optimal time can be found as shown in Fig. 3.4(a). The pulse obtained by the optimiser at T_{opt} is shown in Fig. 3.4(left inset) and achieves (with correction of the detection errors) $\mathcal{F}_m(0.718\mu s) = 0.9847$, improving the measurable STIRAP efficiency by $\Delta\mathcal{F}_m = 0.9\%$ in simulation. The free parameter δ_b is set to 0 by the optimiser. We then implement this protocol on the Chadoq device. The latter takes as inputs AWG voltage signals and perform several steps to generate the controls. Some of those steps can be characterised by an experimentally measured transfer function while others like the modulation by the AOM are not easily reversible. Thus, the experimental controls sent to the atoms might slightly differ in shape from the theoretical ones. $\Omega_{r/b}^{max}$ are set and measured to be at the right level while δ_b is set to 0 with ± 1 MHz inaccuracy. The experimental recapture rates obtained while sweeping δ_r are shown in Fig. 3.4(b) and are in agreement with the noisy simulations. Fluctuations in calibrating δ_b, ε and ε' can explain the few discrepancies. Interestingly, we notice that around $\delta_r \approx \delta_0$ (at the resonance) the simu./expt. discrepancies are less pronounced in the optimised case. This may be explained by the smaller power needed for this optimised protocol, not being enough to couple unwanted transitions between $|g\rangle$ and others hyperfine levels. The optimised protocol also seems less robust than the previous one but exhibits better results in the constrained range of $|\delta_r/2\pi| \leq 3$ MHz. However, while the optimised pulse is shorter and requires less power, it remains hard to precisely measure its expected benefit in terms of measurable fidelity (+0.9%). Looking at the theoretical fidelity $\mathcal{F}$, the gap between the two pulses is in fact larger by 1.7% meaning that the benefit of the optimised protocol is not fully measurable by $\mathcal{F}_m$ on the device. A possible procedure for amplifying this signal would be to compute an optimised inverse STIRAP protocol and apply sequentially the two several times. By measuring the expected decay of the fidelity after enough steps, we could trace back from experimental measurements the theoretical improvement. Another possibility would be to switch to a CRAB like approach with a simpler parameterisation, only optimising on the maximum amplitudes, pulse widths and time separation of the pulses. This could enable to directly optimise the protocol on the Chadoq device.

3.2 - Running variational quantum algorithms on a QPU

Model-based optimal control techniques described above are especially useful when the controlled system can be modelled accurately and its dynamics solved analytically or emulated at low cost. However, for systems of larger sizes, i.e. with more qubits, where the emulation cost is prohibitive or for systems with noisy output hard to characterise, one can resort to closed-loop experimental optimisations. This approach operates independently of a Hamiltonian or other mathematical description of the quantum system, making it highly effective in scenarios with unknown noises, such as caused by unexpected couplings, interfering energy levels and control distortions.

3.2.1 - Building a hybrid closed-loop

The term closed-loop refers to the iterative process over parameters $\theta = (\theta^1, \ldots, \theta^i, \ldots) \in \Theta$ between a classical optimiser, tasked to compute and minimise a cost function $C(\theta)$ through direct queries to the experimental quantum system, producing the states $|\psi(\theta)\rangle$. If Θ only contains a finite set of elements, the optimisation is said to be *discrete*, otherwise it is said to be *continuous*. The optimiser iteratively adjusts candidate controls $u(\theta)$, incorporating experimental measurement results into its decision-making process. The iterative loop ends after meeting fixed convergence criteria, such as exceeding the total number of iterations or cost evaluations allowed (which can be different if the decision-making process requires to probe several points to adjust its model) or having variations $||\theta^{(k+1)} - \theta^{(k)}||$ or $||C(\theta^{(k+1)}) - C(\theta^{(k)})||$ below a predefined threshold. Once θ_{opt} has been found, such that

$$\theta_{\text{opt}} \approx \underset{\theta \in \Theta}{\text{argmin}}\ C(\theta), \tag{3.11}$$

the prepared state, the optimised controls or the minimum cost can be used in subsequent problems. This concept of closed-loop is especially used in the formalism of VQA, first introduced in the digital paradigm. On the contrary to standard optimal control problems, focusing on building the optimised controls $u(\theta_{\text{opt}})$, VQAs aim at finding approximate solutions, embedded in $|\psi(\theta_{\text{opt}})\rangle$, to quantum computational problems. The parameterisation typically adopts the structure of a quantum circuit, often consisting of a series of arbitrary unitary rotations denoted as $R_{\boldsymbol{e}(\phi)}(\theta^i)$ (refer to Eq.,1.5), which are subject to optimisation. However, with the emergence of the neutral atom technology, a number of analog-based VQA proposals have emerged, with some of them detailed in Sec. 6.2.1. Gradient-based methods have been especially used for the VQA classical optimisation part but can suffer from the appearance of barren plateaus depending on the expressiveness of the parameterisation, as described in Box. 15. The control parameterisations $u(\theta)$ introduced above in the QOC paradigm can be extended to the analog VQA and additional ones will be used in Sec. 4.2.4 and Sec. 6.2.1.

Box 15: The problem with gradient descent methods for VQA

The choice of the parameterisation of the quantum evolution, or *ansatz*, whether in the form of circuits or continuous controls, directly influence the landscape of the cost function. A significant challenge arises when this landscape appears relatively *flat* as gradient-based classical methods rely on following the gradients $\nabla_\theta C$ to navigate the parameter space. Unsurprisingly, this phenomenon can occur when the evolution is over-parameterised and some parameters have little influence on the overall dynamics. However, a larger problem arises when the gradients relative to the parameters θ^i all exponentially vanish with N the number of qubits involved, causing the optimisation process to stall: this problem is referred to as barren plateaus [105]. Specifically, it has been observed that highly expressive ansätze tend to exhibit flatter cost landscapes, making them more challenging to optimise [106]. Recent findings have also established a connection between the dynamical Lie algebra of an ansatz and the occurrence of barren plateaus [107]. The effect of barren plateaus can be formalised by

$$\mathrm{Var}_\theta \left(\frac{\partial C(\theta)}{\partial \theta^i} \right) \in O(1/b^N) \tag{3.12}$$

for $b > 1$. Combined with the Chebyshev's inequality, the latter implies that the gradient of the cost function will be, on average, exponentially small since

$$p\left(\left| \frac{\partial C(\theta)}{\partial \theta^i} \right| \geq c \right) \leq \frac{1}{c^2} \mathrm{Var}_\theta \left(\frac{\partial C(\theta)}{\partial \theta^i} \right) \tag{3.13}$$

with $c > 0$. Consequently, gradient-based optimisation classical algorithms struggle to identify the direction of steepest descent, as the gradient provides weak or no guidance. The presence of barren plateaus poses a significant challenge for VQAs with parameterised quantum circuits and gradient-based optimisation especially when applied to larger and larger systems. It can lead to slow convergence, increased sensitivity to noise and errors, and a higher likelihood of getting stuck in sub-optimal solutions.

3.2.2 - A QPU is just an expensive noisy black box

Finding the optimum value of a cost function $C(\theta)$ requires a well-crafted strategy to probe the parameter space Θ. Without prior knowledge, finding the optimal solution could be done by evaluating C for every possible combination of parameters. However sufficient for small discrete spaces, this simplest scheme is quickly rendered unpractical on larger space by the curse of dimensionality. Indeed, let imagine a classical computer evaluating C at a given set of parameters θ in less than 1 μs. If θ is comprised of 20 parameters, such as the control slices in GRAPE, and each of them can only take 10 values (due to finite resolution of the signal generator for instance), then an exhaustive search would still take a few million years. Furthermore, the functions considered here call on quantum computers for their evaluation and thus bears a significantly higher monetary/time cost. Optimising such *expensive* functions imposes fast converging methods, sparing on the number of calls. If C is a *black box* function, lacking known particular characteristics such as linearity, differentiability or concavity, traditional methods, including gradient-based approaches, often offer limited assistance.

Nevertheless, certain techniques, like the finite differences and the subsequent parameter shift rule [108, 109], provide the means to estimate gradients for such functions in certain scenarios, albeit requiring additional costly evaluations of C. Finally, two successive evaluations of C at the exact same set of parameters may not yield identical results. Indeed, the high-level cost C might be impacted by low-level noises such as finite sampling of the quantum state produced, fluctuating controls or inherent random physical processes happening during the dynamics. Examples of scenarios where the expensive noisy black box optimisation problem arises abound, including tuning hyperparameters in large machine learning models, performing lab experiments in biomedical research, optimising investments portfolio in the finance industry or running a VQA on a NISQ quantum computer. In the rest of the section, we will focus into choosing an efficient classical decision-maker for the future hybrid loops used for antiferromagnetic state preparation in chapter 4 and combinatorial problems in chapter 6.

3.2.3 - Gradient-free optimisation of black box functions

Although, the gradient-descent deterministic approach enables tackling parameterisation with hundreds of variables, such as in GRAPE, the existence of barren plateaus and the high cost of estimating gradients of expensive noisy black box functions make it less successful in VQAs. By resorting to ansätze with fewer parameters (typically less than 20), such as in CRAB, an exploration of the parameter space with gradient-free methods becomes viable. A wide variety of gradient-free methods exist, the best method to use depending on the formulation of the problem and the desired result. They can either be deterministic, with the example of Nelder-Mead downhill simplex method [110] or probabilistic, incorporating randomness into the optimisation process, often by introducing probabilistic sampling or perturbation techniques. Among them, the simulated annealing method described in Box. 21 is based on a non-greedy random search with a variable trust region and can still handle many parameters. Genetic, or evolutionary, algorithms draw inspiration from natural selection, evolving a population of candidate solutions through selection, crossover, and mutation operations to find the optimal solution, mimicking survival of the fittest and genetic recombination [111]. Particle swarm optimisation is a population-based technique mimicking bird flocks or fish schools, where candidate solutions, or particles, adjust their positions based on their own and neighbouring particles best-known positions, collectively converging towards optimal solutions by dynamically exploring promising regions of the parameter space [112]. Finally, surrogate-based techniques approximate expensive black box functions with faster-to-evaluate models [113], such as neural networks, random forest of decision trees [114] or also Gaussian processes. They are typically trained on a limited set of calls to the original function and can then provide fast predictions of the function values at unexplored points in the parameter space. With a fairly representative model, the optimum found should be close to the real one. However, the accuracy of the model is limited by several issues: the dimension of the input space, the complexity of the modelled function, the presence of data noise and the bias–variance trade-off [115]. The latter accounts in the observed discrepancy for two primary error sources: bias, resulting from flawed assumptions in an under-trained algorithm, and variance, originating from the excessive sensitivity to noise in an over-trained algorithm. The learning sample size must therefore be sufficient without being excessive. By sequentially adding carefully selected data to the model, we can drive it to approach C in regions the function is more likely to show a minimum. This process typically involves leveraging a score function constructed from the model parameters. The target regions should be the

surroundings of a spotted potential minimum, but also the unexplored areas that could host one. The Bayesian algorithm presented in the following section encompasses all of these strategies in order to quickly detect the minimum of an expensive black box function with as few calls to it as required.

3.3 - Bayesian algorithm as classical optimiser

Bayesian optimisation algorithm has two keywords, model and decide [116]. It encompasses both a statistical model that reconstructs the landscape of the target function C while providing an uncertainty on such reconstruction and a decision maker, the acquisition function a, which indicates where the next evaluation will be most likely to enhance optimisation. After a training phase where the model first fits few evaluations of C, as in surrogate-based methods, each new iteration will favour either exploitation of promising areas of the parameter space or exploration of regions where the high uncertainty leaves room for a potential minimum. The newly acquired measurement $C(\theta)$ updates the prior knowledge of the model using a Bayesian inference technique. In the following, we delve a bit more into statistical modelling with Gaussian processes and how the acquisition function helps making decisions.

3.3.1 - Statistical modelling with Gaussian processes

For the algorithm to perform efficiently, the modelling of the landscape needs to be expressive enough while staying as simple as possible to prevent undue costs. Random processes avoids the necessity of an explicit parameterisation and guarantee essential properties like continuity and differentiability. While in the following, we focus on Gaussian processes, it is possible to use other statistical surrogate models such as binomial distributions which can represent the uncertainty of quantum measurements in a fairer way [117].

Gaussian processes

A Gaussian process (GP) is a probabilistic model used for capturing uncertainty in functions. It is defined by a *mean* function μ and a *covariance* function (also known as *kernel*) k. Sampling from a GP at θ yields a set of values normally distributed with mean $\mu(\theta)$ and variance $k(\theta, \theta)$, while their correlations with other values sampled at θ' are governed by the positive semi-definite covariance matrix defined with $k(\theta, \theta')$. Selecting μ and k establishes a *prior* probability distribution $p(f)$ for the GP, encompassing the beliefs about C before observing any data.

Faithfully reproducing the target function C with a GP requires some initial assumptions, which, in the case of black-box functions, are usually quite poor. With no prior knowledge of C, no peculiar behaviour for the mean can be assumed and a standard procedure consists in arbitrarily setting the mean function μ to a constant μ_0. If C is said to show some parametric structure, μ can also be written as a sum of low-order polynomials in θ. The choice of k enables to choose to what degree closeness in the parameter space should lead to correlation in values with various examples given in Box. 16/

Box 16: Examples of kernel functions

Various kernels are available, based on the properties the models are assumed to exhibit, with among them [118]:

- the Gaussian, or Radial-basis, kernel given by:

$$k(\theta, \theta') = \exp\left(-\frac{||\theta - \theta'||^2}{2l^2}\right) \tag{3.14}$$

  where the length scale l accounts for how fast a surrogate model is believed to change with θ. This infinitely differentiable kernel is often used, due to its smooth and stationary behaviour.

- the Matern kernel, generalisation of the above, is written as:

$$k(\theta, \theta') = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{\sqrt{2\nu}}{l}||\theta - \theta'||\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}}{l}||\theta - \theta'||\right) \tag{3.15}$$

  with Γ, the gamma function, ν, the smoothness parameter, and K_ν, a modified Bessel function. As ν increases, the model should smoothen, with peculiar values at 1/2, 3/2 and 5/2 where it is respectively continuous, once differentiable and twice differentiable. A fairly used one is

$$k(\theta, \theta') = \sigma^2\left(1 + \frac{\sqrt{3}||\theta - \theta'||}{l}\right)\exp\left(-\frac{\sqrt{3}||\theta - \theta'||}{l}\right) \quad \text{for } \nu = 1.5 \tag{3.16}$$

  where σ is a normalisation term.

- the ExpSineSquared kernel is especially useful if the target has a known periodicity since

$$k(\theta, \theta') = \exp\left(-\frac{2\sin^2(\pi||\theta - \theta'||/p)}{l^2}\right) \tag{3.17}$$

  where p is the periodicity of the model.

- the Rational Quadratic kernel, which is a mixing of Gaussian kernels with different length scales

Training the prior model

In order for the prior model to reproduce the target, we need to train it on a few observed evaluations of C. Provided with a budget of n_r calls, C is surveyed according to an initial space-filling pattern. For small-sized spaces, the probing points, if enough, can be formed into a discrete grid efficiently covering the entire parameter space. However, as the number of dimensions grows, it becomes less resource consuming to utilise a more efficient space filling pattern such as Latin hypercube sampling. The latter ensures a covering of the entire range of each input variable while minimising

correlation between variables. The acquired training dataset $D = (\boldsymbol{\theta}, \boldsymbol{C}) = \{\theta^{(k)}, C(\theta^{(k)})\}_{k=1,\dots,n_r}$ will be used to prepare an initial model of C.

Tuning the hyperparameters

The training dataset is supposed to have been drawn randomly with probability $p(\boldsymbol{C}|f)$ by sampling a model function f obtained from the prior probability distribution $p(f)$. While f is entirely defined by μ and k, the latter include numerous hyperparameters η, for instance $\eta = \{\mu_0, \sigma, l\}$ in the case of constant mean and Matern kernel. One can also include a noise hyperparameter which will be used to fit the amount of noise estimated in the observations. Two methods exists for the hyperparameters tuning in order to best capture the underlying function [119].

- Maximisation of the log marginal likelihood. This likelihood represents the probability of observing the data given a particular set of hyperparameters and reads

$$p(D|\eta) = \int_f p(D|f)p(f|\eta)df \tag{3.18}$$

  The set of hyperparameters explaining the best the data are $\tilde{\eta} = \text{argmax}_\eta \log(p(D|\eta))$.

- Full Bayesian approach. By combining a prior distribution over the hyperparameters $p(\eta)$ with the above likelihood one can retrieve the posterior distribution using the Bayes rule

$$p(\eta|D) = \frac{p(\eta)p(D|\eta)}{p(D)} \tag{3.19}$$

  Note that in practice, this distribution can only be approximated, using computational methods such as Markov chain Monte Carlo. With this posterior distribution, one can make predictions about the optimal hyperparameters or estimate the uncertainty associated with each hyperparameter.

Following the hyperparameter tuning, one can update the prior assumption of μ and k to generate the posterior distribution $p(f|\boldsymbol{C})$ using the following conditioning operations

$$\mu'(\theta) = \boldsymbol{\kappa}^T \boldsymbol{K}^{-1}(\boldsymbol{C} - \mu(\boldsymbol{\theta})) + \mu(\theta) \tag{3.20}$$

$$k'(\theta, \theta) = k(\theta, \theta) - \boldsymbol{\kappa}^T \boldsymbol{K} \boldsymbol{\kappa} \tag{3.21}$$

where $\boldsymbol{\kappa} = k(\theta, \boldsymbol{\theta})$ encompasses the correlation between a new point and the precedently observed ones, while $\boldsymbol{K} = k(\boldsymbol{\theta}, \boldsymbol{\theta})$ constitutes the kernel matrix of the observed data. The model is then considered trained. If we sample the posterior distribution $p(f|\boldsymbol{C})$ multiple times across the parameter space, each generated function evaluated at a point $\theta^{(k)} \in \boldsymbol{\theta}$ will yield the exact value $C(\theta^{(k)})$ with zero variance. In the vicinity of $\theta^{(k)}$, the sampled functions will produce values close to $C(\theta^{(k)})$ with variance increasing as the distance from $\theta^{(k)}$ grows. However, further away from the already sampled points, the new ones will be randomly distributed with larger variance. This behaviour reflects the uncertainty captured by the posterior distribution across the parameter space. An example is given in Fig. 3.5(a) where a noisy $C(\theta)$ (red) is approximated by a Gaussian process, using a fitted Matern kernel, trained over 5 observations spaced with the latin hypercube sampling method. Regions devoid of observations exhibit large uncertainties and since the points are somehow distant, the noisy behaviour of the function has yet to be understand by the model.

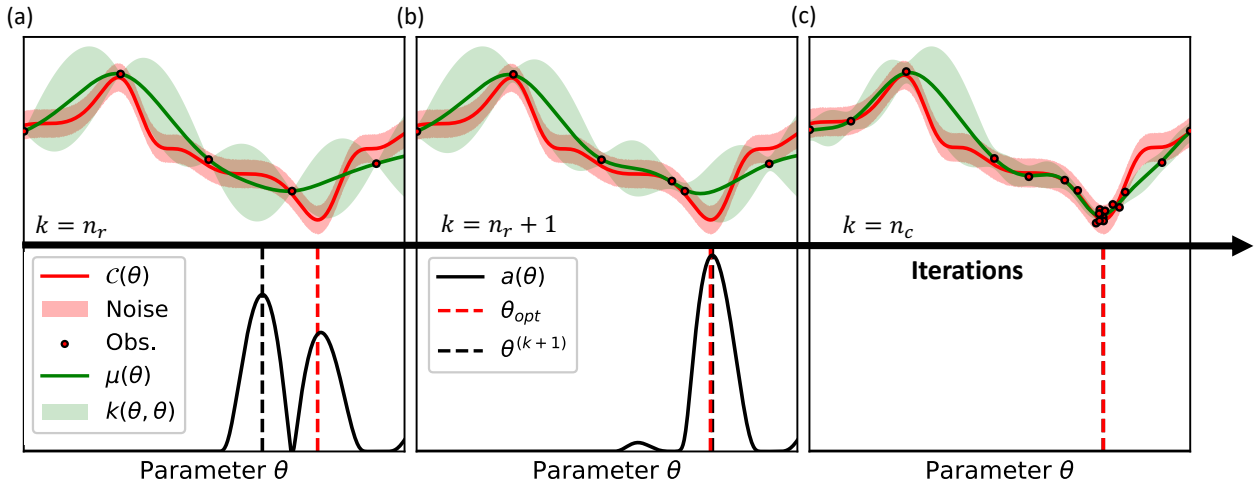


Figure 3.5: **Iterative procedure of a Bayesian optimisation.** (a) A noisy cost function C (red) is sampled 5 times to generate observation data (dots) used for training a Bayesian algorithm. The fitted GP model (green) approximates C with μ with relative uncertainty given by k. Below, the maximisation of the acquisition function a (black) outputs the next point to query $\theta^{(k+1)}$ (dashed black). (b) Updating the model with this new observation modifies μ, k and a iteratively. (c) After enough iterations, the model has reproduced interesting regions of C and locates the minimum θ_{opt} (dashed red).

3.3.2 - An iteratively updated decision maker

Navigating the parameter space with the acquisition function

The acquisition function $a(\theta)$ is defined using the trained surrogate model $f(\theta)$ provided by the Gaussian process regression. It needs to take into account both the objective value predicted by f and the uncertainty around this value and computes a score with a trade-off between exploitation and exploration: while the former focuses on points in the vicinity of the last optimum, the latter pushes the search towards unexplored regions with higher variance values. These regions where we have little knowledge and are often distant from our previous measures might host better points. Thus reducing their high uncertainty may be valuable. If exploitation is largely favoured over exploration, then the algorithm will certainly converge faster, but towards a solution that risks being simply local and not global. Otherwise, a solution will only be found after a slow sampling of the entire space. The balance between those two strategies largely impact the performance of the algorithm.

Types of acquisition function

Once the GP model has been trained, various acquisition functions can be considered to select the best point to probe next :

- The *lower confidence bound* (LCB) allows to directly balance exploration and exploitation with

$$a_{\text{LCB}}(\theta) = -\mu(\theta) + \alpha\sqrt{k(\theta,\theta)} \tag{3.22}$$

  where $\alpha > 0$.

- The *probability of improvement* (PI) uses the utility function:

$$u(\theta) = \begin{cases} 1 & \text{if } f(\theta) \leq c \\ 0 & \text{if } f(\theta) > c \end{cases} \quad \text{where } c = \min \boldsymbol{C} \tag{3.23}$$

  which outputs a reward if the probed θ turns out to improve the minimum. The simplicity of the reward, i.e. constant and not proportional to the improvement, constitutes a weakness of this approach and may lead to local optima rather than global ones. The resulting acquisition function is the expected utility:

$$a_{\text{PI}}(\theta) = \mathbb{E}[u(\theta)|\theta, \boldsymbol{C}] = \Phi[c\,; \mu(\theta), k(\theta,\theta)] \tag{3.24}$$

  with Φ, the probability distribution function. The next optimal point returned by the a_{PI} is the one bearing the highest probability of improvement, i.e. the maximal expected utility.

- The *expected improvement* (EI) builds on the above method and refines it: the reward is now defined as the value of the improvement, using the following utility function:

$$u(\theta) = \max(c - f(\theta), 0) \tag{3.25}$$

  The corresponding expected utility gives :

$$a_{\text{EI}}(\theta) = \mathbb{E}[u(\theta)|\theta, \boldsymbol{C}] = (c - \mu(\theta))\Phi[c\,; \mu(\theta), k(\theta,\theta)] + k(\theta,\theta)\phi[c\,; \mu(\theta), k(\theta,\theta)] \tag{3.26}$$

  with ϕ, the cumulative distribution function. The two components of a_{EI} encompass respectively the exploitation part and the exploration part.

Fig. 3.5(a) gives an example of the landscape of a_{EI} using the trained model.

Optimisation of a

Having a well-defined acquisition function a over the parameter space θ enables to switch from an expensive optimisation of C to a relatively cheaper one. In order to find where to evaluate the function next, one can solve $\theta^{(k+1)} = \text{argmax}\ a(\theta)$, using the quasi-Newton method L-BFGS-B [120] which exploits knowledge of the first and second derivatives of a. This gradient-based algorithm makes only an estimation of the Hessian matrix, usually hard to compute, which greatly speeds up the process. The optimised value $\theta^{(k+1)}$ represents where C should be probed next as conveyed by the additional observation in Fig. 3.5(b).

Updating the model

With this additional point, the model can be improved, following the two steps already mentioned, i.e. the hyperparameter tuning and the conditioning operations, using the updated dataset of observations. Fig. 3.5(b) gives the updated model when incorporating the point chosen by a at previous iteration. After only one iteration, the model has already located a point to query close to the true minimum. After $n_c - n_r$ additional iterations, the model has converged to θ_{opt} and no better improvement is expected as conveyed by Fig. 3.5(c). Another feature of this algorithm is the estimation of the noise level of the function it tries to optimise. It is worth noting that due to the noisy behaviour of C, sampling close points may return very different values; the model has thus modified the hyperparameter describing the noise to take that into account. This is highlighted by the non zero variance at already observed points.

3.3.3 - Convergence in presence of local minima

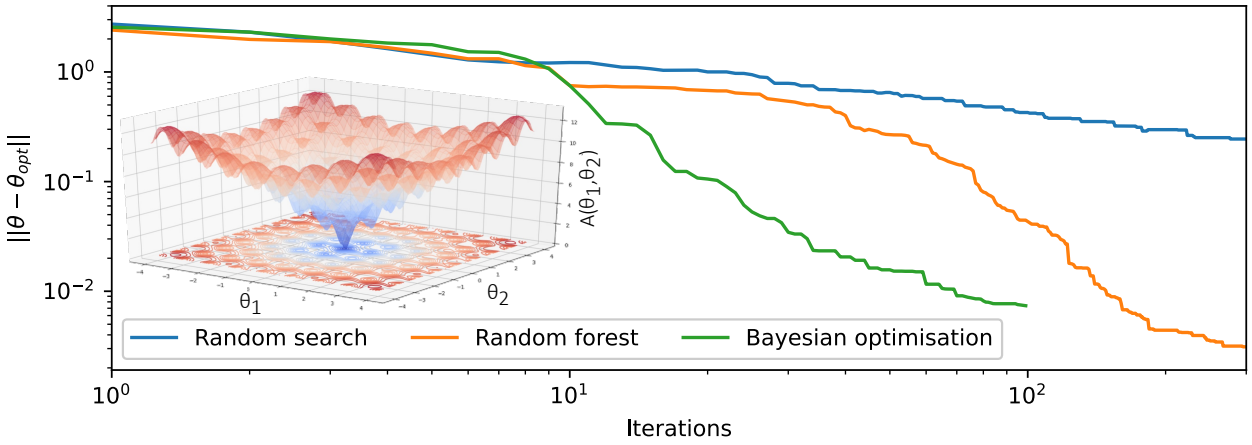


Figure 3.6: **Example of convergence for Bayesian optimisation**. The averaged convergences of a random search (blue), a random forest algorithm (orange) and a bayesian algorithm (green) are compared when optimising an instance of the Ackley function (inset). The performance is assessed with the distance to the global minimum.

The Bayesian algorithm is said to converge faster to more global optima than other gradient-free methods. In order to test this assumption and assess its relevance before including it in the VQAs of following chapters, we realise a convergence comparison between several algorithms. The 2D-Ackley function, displayed in the inset of Fig. 3.6,

$$\mathcal{A}_2(\theta_1,\theta_2) = e + \alpha[1 - \exp\left(-\beta\sqrt{\frac{\theta_1^2+\theta_2^2}{2}}\right)] - \exp\left(\frac{\cos(2\pi\theta_1)+\cos(2\pi\theta_2)}{2}\right), \tag{3.27}$$

with $\alpha = 20$ and $\beta = \alpha/100$, is taken as the expensive black box and the search space is the square $[-4,4]^2$. While exhibiting a large amount of local minimum, this function only has one global null minimum at the origin. We benchmark the Bayesian approach to a Random Forest algorithm [121],

also used in Chapter 8. An averaged over 20 runs for each makes the comparison more robust to lucky initial guesses.

As conveyed by Fig. 3.6, a random search struggles to approach the global minimum. The Bayesian approach clearly outperforms the random forest, quickly locating the region of overall minimum within 20 iterations on average, 5 times quicker than its counterpart. However, the algorithm is quite resource consuming as the models are updated and the observation dataset enlarged at each iteration. Hence, the use of the Bayesian algorithm seems justified to study unknown functions with many local minima as long as the cost evaluation remains large in front of the model updating duration (few seconds after 100 iterations). The algorithm is thus suited as a classical routine for VQAs. For cases where the cost evaluation is negligible, other gradient-free methods are preferable.

Part II

Quantum many-body dynamics for quantum simulation

4 - Quantum simulation of Ising magnets

Contents

Studying intriguing properties of condensed matter systems like superconductivity requires being able to vary the system parameters to evaluate their impact on observed features. Since some of them may not be tunable in reality, modelling systems from a microscopic viewpoint allows for the numerical examination of emerging macroscopic traits by adjusting the desired parameters of the effective model. Insulating magnetic materials can for instance be described by spin Hamiltonians as the main degrees of freedom of the system are the magnetic moment of electrons localised at the sites of a crystal lattice structure. The most extensively studied spin models include the isotropic Heisenberg Hamiltonian $\hat{H} = \sum_{\langle i,j \rangle} J_{ij} \hat{\vec{\sigma}}_i \cdot \hat{\vec{\sigma}}_j$ and simpler ones like the anisotropic XY coupling $J(\hat{\sigma}_i^x \hat{\sigma}_j^x + \hat{\sigma}_i^y \hat{\sigma}_j^y)$ (whose dipolar version is tackled in chapter 5) or the uni-axial $J\hat{\sigma}_i^z \hat{\sigma}_j^z$ introduced by Ernst Ising [122]. Despite the apparent simplicity of these Hamiltonians, they lead to a fascinating variety of collective phenomena referred to as quantum magnetism [123]. The motivation for quantum simulation then arises from a simple statement. The basic exponential cost of numerically emulating quantum systems led scientists like Richard Feynman to propose that for efficient computation of dynamics in large-scale instances, the simulator itself should be quantum-mechanical in nature [124]. Since then, countless examples of quantum simulation tasks have been presented across the various quantum technologies, including trapped ions [125, 126], superconducting circuits [127], quantum dots [128], atoms [129] and polar

molecules [130] in optical lattices, and atoms trapped in tweezers [74, 131].

In this chapter, we leverage the capability of Rydberg atoms to simulate an Ising-like Hamiltonian and explore the phase diagram of the antiferromagnetic 1D Ising model. Although this model has been extensively studied, we employ the formalism of variational quantum algorithms, introduced in Chapter 3, to adiabatically prepare the antiferromagnetic ground state and enhance preparation efficiency. Additionally, following a thorough benchmark of the limiting noise processes present during these adiabatic protocols, we implement a closed-loop system to improve results obtained on the QPU. Lastly, we briefly touch upon the many-body localisation problem [132], which necessitates initialisation in the antiferromagnetic phase, and demonstrate how it can be tackled using quenched dynamics.

4.1 - Probing the antiferromagnetic phase of the Ising model

Although relatively simple, the transverse field Ising (TFI) model yields intuitive insights into quantum magnets. Its nearest-neighbour 1D case was analytically studied in the thermodynamic limit 70 years ago [133], with its phase diagram calculated. As such, it serves as a valuable toy-model for benchmarking protocols such as adiabatic preparation ones.

4.1.1 - 1D Ising model with Rydberg atoms

Considering a 1D chain of N spin$-1/2$ $|\downarrow\rangle - |\uparrow\rangle$ placed in a magnetic field with longitudinal component $B_\parallel$ (along the quantisation axis of the spin) and transversal component $B_\perp$ (perpendicular to the quantisation axis of the spin), the Hamiltonian of the system reads

$$\hat{H}_{TFI} = B_\perp \sum_{i=1}^{N} \hat{\sigma}_i^x + \sum_{i=1}^{N} (B_\parallel + B_{\text{local}}^i)\hat{\sigma}_i^z + J \sum_{i=1}^{N} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \tag{4.1}$$

with the spins interacting with nearest neighbours coupling J and $\hat{\sigma}_i^\alpha$ the Pauli matrices. The different components of H each promote distinct spin arrangements. While $B_\perp$ favours spins aligned in the transverse direction $|\rightarrow\rangle = (|\downarrow\rangle + |\uparrow\rangle)/\sqrt{2}$, $B_\parallel$ favours spins aligned in the longitudinal direction, with the orientation depending on its sign. The spins structure themselves in a ferromagnetic order with $J < 0$ as neighbouring spins tend to align in the same direction in order to minimise the system energy, whereas they anti-align in an antiferromagnetic order with $J > 0$. A local longitudinal magnetic field B_{local}^i can also be incorporated to, for instance, study disordered chains. This system exhibits a quantum phase transition [134] as explained in the following.

This simple, yet rich, model can be "reproduced" in neutral atom platforms using the ground-Rydberg encoding of the qubits. Indeed a 1D chain of N regularly spaced qubits with basis states $|\downarrow\rangle = |g\rangle$ and $|\uparrow\rangle = |nS\rangle$ gives rise to the following Hamiltonian

$$\hat{H}_{\text{ryd}}[\Omega, \delta] = \frac{\hbar\Omega}{2} \sum_{i=1}^{N} \hat{\sigma}_i^x - \hbar\delta \sum_{i=1}^{N} \hat{n}_i + U \sum_{i<j} \frac{a^6}{r_{ij}^6} \hat{n}_i \hat{n}_j, \tag{4.2}$$

with $\hat{n}_i = (\hat{1} + \hat{\sigma}_i^z)/2$, a, the chain spacing, $U = C_6/a^6$ the nearest neighbours coupling and $r_{ij} = a|i-j|$. Eq. 4.2 fundamentally differs from $\hat{H}_{TFI}$ due to the power-law decay of its van der Waals interaction. However, for pairs separated from more than $r_b = (C_6/\Omega)^{1/6}$, the decay in $1/R^6$ enables to overlook

the pairwise terms as their strength quickly become negligible compared to Ω. In addition, $\hat{H}$ can be written such that $\hat{H}_{\text{ryd}} = \hat{H}_{TFI} + N(U/4 - \hbar\delta/2)\hat{1}$ taking $B_\perp = \hbar\Omega/2$, $B_\parallel = -\hbar\delta/2$, $U = 4J$ and $B^i_{\text{local}} = z_i U/2$, with z_i the number of neighbours of spin i. The additional term in the equality is diagonal and can be gauged out when evolved over time. For infinite-size and periodic systems ($z_i = 2$), the local fields B^i_{local} can be absorbed in the global term $B_\parallel$. It is thus possible to implement an Ising-like Hamiltonian with Rydberg atoms, up to some shifts and rescaling of the coefficients. We will therefore use this system to probe features of the Ising model, especially of its *antiferromagnetic* phase with $U > 0$.

4.1.2 - Characterising ground state of 1D Ising model

An essential inquiry in studying a model involves discerning the system's ground state relative to variations in the Hamiltonian parameters such as Ω and δ. This entails uncovering the system phase diagram, which delineates distinct regions, called *phases*, where the ground state exhibits unique characteristics, such as ordered spins.

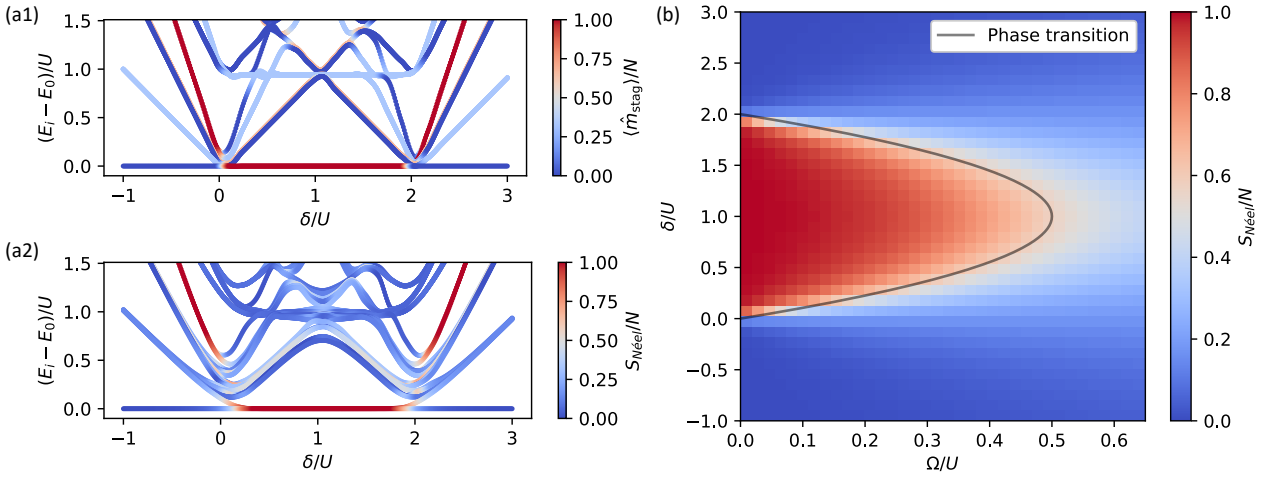


Figure 4.1: **Phase diagram of the 1D Ising model** for even-length periodic chain of $N = 6$. Energy spectrum of $\hat{H}_{\text{ryd}}$ for varying $\delta/U \in [-1, 3]$ and (a1) $\Omega/U = 0$ and (a2) $\Omega/U = 0.15$ with levels coloured by their degree of ordering, either with $\langle \hat{m}_{\text{stag}} \rangle$ for no transverse field or with $S_{\text{Néel}}$. (b) 2D phase diagram over the region $\delta/U \in [-1, 3]$ and $\Omega/U \in [0, 0.65]$. At each point, the structure factor is computed for the ground state of $\hat{H}_{\text{ryd}}$. The lowest levels of (a1,a2) are vertical slices of (b). The transition line calculated at thermodynamic limit is figured in black.

We consider a even-numbered chain of spins with $N = 2n$ with periodic boundaries as to avoid the edge-effects described in Box. 17. In the absence of transverse field, i.e. $\Omega = 0$, the Hamiltonian in Eq. 4.2 is diagonal in its computational basis. When $|\delta| \gg U$ the ground state of that system is *paramagnetic*, all the spins are aligned towards a favoured direction depending on the sign of δ. The system is fully magnetised as characterised by the overall magnetisation $\hat{m}^z = \sum_i \hat{\sigma}^z_i$ and the interactions play no role. To locate the range of values for which the ground state is in its antiferromagnetic phase, we can look at the energy of states in which this order appears. We denote $|\psi_{AF\downarrow}\rangle = |\downarrow\uparrow \ldots \downarrow\uparrow\rangle$ the product state for which spins are anti-aligned with the first one being down

and conversely $|\psi_{AF\uparrow}\rangle$ with the first one being up. The AF order of these product states can be characterised by their *staggered magnetisation*

$$\hat{m}_{\text{stag}} = \frac{1}{N}\sum_i (-1)^i \hat{\sigma}_i^z, \tag{4.3}$$

a one-body observable assessing the anti-alignment of neighbouring spins. Those two states have the same energy (since they are equivalent up to a rotation) which in the NN approximation reads $E_{AF}(\delta) = -n\delta$. Flipping one spin from $\uparrow$ to $\downarrow$ alters this energy by δ while flipping one from $\downarrow$ to $\uparrow$ modifies it by $-\delta + 2U$. For those two quantities to be positive and thus $E_{AF}(\delta)$ to be the minimal energy, one needs to have $0 \leq \delta/U \leq 2$. We can check this assumption even further by diagonalising $\hat{H}_{\text{ryd}}$ for various values δ and $\Omega = 0$, revealing in Fig. 4.1(a1) the AF zone where $|\psi_{AF\downarrow/\uparrow}\rangle$, characterised by $\langle \hat{m}_{\text{stag}} \rangle = \pm 1$, are degenerate ground states of the system. At the critical points $\delta = 0$ and $\delta = 2U$, many energy levels cross and there is no gap between the ground state and the first excited state.

Box 17: Edge effects in an open even chain

In the absence of transverse field, one of the ground states of a even-numbered $N = 2n$ chain in a NN Ising model would be $|\psi_{\text{AF}\downarrow}\rangle$ with energy $-n\delta$. However, unexpected edge effects arise even in this simplified model and states, such as $|\psi_{\text{pAF}}\rangle = |\uparrow\downarrow\downarrow\uparrow \ldots \downarrow\uparrow\rangle$ where the first two spins are inverted, also have energy $-n\delta$. Adding the tail of interactions lifts the degeneracy as the respective energies write:

$$E_{\text{AF}\uparrow} = -n\delta + U\sum_{i=1}^{n-1} \frac{n-i}{(2i)^6} \tag{4.4}$$

$$E_{\text{pAF}} = -n\delta + U\left(\sum_{i=1}^{n-1} \frac{n-i-1}{(2i)^6} + \sum_{i=1}^{n-1} \frac{1}{(2i+1)^6}\right) \tag{4.5}$$

The gap between those two levels reads

$$\Delta E_{\text{AF}} = E_{\text{AF}} - E_{\text{pAF}} = U\sum_{i=1}^{n-1}\left(\frac{1}{(2i)^6} - \frac{1}{(2i+1)^6}\right) \simeq 4.6\%U \tag{4.6}$$

Thus for an even number of atoms in a 1D chain, the ground state of the Ising Hamiltonian is no longer the AF state but a slightly perturbed version where an excitation can move to the edge to avoid interacting. In order to avoid these edge effects, the chain can be made periodic, with the atoms forming a circle for instance. In the case of a circle of radius R, the gap becomes:

$$\begin{aligned}\Delta E_{\text{AF}} &= U\left(\frac{a}{R}\right)^6 \sum_{i=1}^{n-1} (|2\sin(2i\pi/N)|^{-6} - |2\sin((2i+1)\pi/N)|^{-6}) \\ &\simeq -U\left(\frac{a}{R}\right)^6 |2\sin(\pi/N)|^{-6} < 0\end{aligned} \tag{4.7}$$

and the AF state becomes the ground state.

Adding a transverse field $\Omega > 0$ complicates the picture. In the limit where $\Omega/U \gg 1$, the spins

align in a paramagnetic phase along the y axis but there is a delimited region in phase space (Ω,δ) which exhibits AF order. This order can no longer be characterised by the same observable as the eigenstates are superposition of product states. For instance at $\delta/U = 1$ and $\Omega/U = 0.15$, the ground state is $|\psi_{\text{AF}}\rangle = (|\psi_{AF\downarrow}\rangle + |\psi_{AF\uparrow}\rangle)/\sqrt{2}$ with vanishing staggered magnetisation. Rather than looking at one-body observables to characterise this state, it is better to look at the correlations between spins. Indeed, in $|\psi_{\text{AF}}\rangle$, spins are perfectly anti-correlated. We introduce the Néel *structure factor* with general expression over a lattice given in Box. 18 which reads in the 1D case

$$S_{\text{Néel}} = \sum_{k\in\mathbb{Z}^*} (-1)^{|k|} g_c^{(2)}(k) \tag{4.8}$$

where $g_c^{(2)}$ is the two-body correlation function obtained from Eq. 4.9 with $\hat{\sigma} = \hat{n}$. We can quickly check that $S_{\text{Néel}}$ reaches its maximum value of N for the superposition of equally favoured AF product states, i.e. $|\psi_{\text{AF}}\rangle$, and equals 0 in the paramagnetic phase where the aligned spins are not correlated and for the classical AF product states. The 1D phase diagram presented in Fig. 4.1(a2) obtained with $\Omega/U = 0.15$ still displays an AF zone centred around $\delta/U = 1$ although a bit reduced compared to the case $\Omega = 0$. Having $\Omega > 0$ lifts the degeneracy around the borders and a gap appears. This gap is known to decrease as $1/N$ when the system grows in size meaning that the gap vanishes in the thermodynamic limit $N \to \infty$.

Box 18: Static structure factor

Given a $d-$dimensional lattice $L = \{\sum_l k_l \vec{u}_l \mid k_l \in \mathbb{Z}\}$ with lattice vectors $\vec{u}_l$ and a measurement operator $\hat{\sigma}$, one can compute the correlation function

$$g_c^{(2)}(\mathbf{k}) = \frac{1}{|\mathcal{P}(\mathbf{k})|} \sum_{(i,j)\in\mathcal{P}(\mathbf{k})} [\langle\hat{\sigma}_i\hat{\sigma}_j\rangle - \langle\hat{\sigma}_i\rangle\langle\hat{\sigma}_j\rangle] \tag{4.9}$$

from the *connected correlations* between sites separated by $\mathbf{k}$ where $\mathcal{P}(\mathbf{k}) = \left\{(i,j) \in [1,N]^2 \mid \vec{r}_i - \vec{r}_j = \sum_l k_l \vec{u}_l\right\}$. The structure factor arising from $g_c^{(2)}$ reads:

$$S(\mathbf{q}) = \sum_{\mathbf{k}\neq\vec{0}} e^{i\mathbf{q}\cdot\mathbf{k}} g_c^{(2)}(\mathbf{k}) \tag{4.10}$$

For a square lattice with spacing R, $\vec{u}_1 = R\,(1,\ 0)$, $\vec{u}_2 = R\,(0,\ 1)$, $\mathbf{q} = \pi(1,\ 1)$ and $\hat{\sigma} = \hat{n}$, one can compute the Néel structure factor S_{Neel} which is maximised for the perfectly correlated antiferromagnetic state $|\psi_{AF}\rangle$.

Finally, we diagonalise $\hat{H}_{\text{ryd}}$ for $N = 6$ over the phase space (Ω, δ) and compute the structure factor over the ground state obtained for $\delta/U \in [-1, 3]$ and $\Omega/U \in [0, 0.65]$. Fig. 4.1(b) highlights the shape of the AF phase characterised by $S_{\text{Néel}} > 0$ for this small system. In an infinitely large system, the two phases are delimited by a second-order quantum phase transition of the $(1+1)$D Ising universality class [134]. This phase transition is described by a line that forms a dome with critical point $\Omega/U = 0.5, \delta/U = 1$ and can be calculated analytically in the thermodynamic limit [133] or by extrapolating the position of the vanishing gap as N increases as done in [5].

In the following, we will prepare the antiferromagnetic ground state of the 1D Ising model using adiabatic protocols on the Rydberg quantum simulator and reconstruct this phase diagram.

4.2 - Adiabatic driving: standard and optimised schedules

Ground state preparation in quantum systems often involves using a schedule of external controls to guide the system from an initial, easily prepared state to the desired ground state. This approach relies on the adiabatic theorem, which in essence ensures that if a system evolves slowly compared to its characteristic energy scale, it follows the instantaneous eigenstate of the Hamiltonian.

4.2.1 - Adiabatic theorem: what slow means

Utilising the adiabatic theorem requires first to couple an initial Hamiltonian $\hat{H}_i$ to a final one $\hat{H}_f$ through a time-dependent schedule $s(t) \in [0, 1]$ such that

$$\hat{H}(s(t)) = (1 - s(t))\hat{H}_i + s(t)\hat{H}_f. \tag{4.11}$$

$\hat{H}_i$ is typically chosen such that its ground state is trivial such as $|\downarrow\rangle^{\otimes N}$ or $|+\rangle^{\otimes N}$. In our case, neutral atom systems are initialised in $|\downarrow\rangle^{\otimes N}$ and $\hat{H}_i = \hat{H}_{\text{ryd}}[\Omega = 0, \delta < 0]$, favouring all spins down.

Let $|\psi_0(s)\rangle, |\psi_1(s)\rangle$ be the instantaneous ground-state and first excited state of $\hat{H}(s)$, with respective energies $E_0(s)$ and $E_1(s)$. According to the adiabatic condition [135], a quantum system remains in its instantaneous ground-state with probability $1 - \epsilon$ if the following condition on the evolution duration T is satisfied:

$$T \gg \frac{1}{\sqrt{\epsilon}} \max_{s\in[0,1]} \frac{|\langle\psi_1(s)|\partial_s\hat{H}(s)|\psi_0(s)\rangle|}{\Delta(s)^2}. \tag{4.12}$$

where $\Delta(s) = |E_1(s) - E_0(s)|$ is the instantaneous gap between ground and excited states. The transition term $|\langle\psi_1(s)|\partial_s\hat{H}(s)|\psi_0(s)\rangle|$ relates to the difference in evolution of the ground and first excited levels in the spectrum. As the schedule is continuous, its derivatives are bounded and this term has no singular scaling with the system size N [136], the adiabatic condition of equation (4.12) can be replaced by the simpler form

$$T \gg \mathcal{O}(\max_{s\in[0,1]} ||\partial_s\hat{H}(s)||/\Delta^2_{\text{min}}), \tag{4.13}$$

where $\Delta_{\text{min}} = \min_{s\in[0,1]} \Delta(s)$. Therefore one requires non-degenerate states without level crossings, i.e. $\Delta_{\text{min}} > 0$, to perform adiabatic evolution within reasonable times. Even when levels do not cross in the presence of a drive, the *anti-crossing* phenomenon can create exponentially small gaps that make the evolution time impractically large. For instance, crossing the transition line depicted in Fig. 4.1(b) at its apex, i.e. at $(\Omega/U = 0.5, \delta/U = 1)$, induces a minimum gap scaling as $1/N$ meaning that the duration to remain truly adiabatic would scale as N^3. As the system scales up, it becomes increasingly challenging to maintain adiabaticity, leading to the introduction of diabatic errors when crossing the phase transition as emphasised by the Kibble–Zurek mechanism [137].

4.2.2 - Driving through the transition with the Rydberg Hamiltonian

In order to probe the phase space, we need to perform an adiabatic evolution between $\hat{H}_i$ and $\hat{H}_f = \hat{H}_{\text{ryd}}[\Omega > 0, \delta]$ using the controls $\Omega(t)$ and $\delta(t)$. A standard protocol [138] consists in starting by a linear increase of Ω towards $\Omega_{\max}$, keeping $\delta(t) = \delta_i < 0$ fixed. It lifts the degeneracy of the levels at $\delta = 0$ by a quantity $\Delta \propto \Omega_{\max}$. The detuning is then linearly ramped up towards its final value $0 < \delta_f < 2U$ while Ω remains fixed. Finally, $\Omega(t)$ is ramped down to 0 with δ fixed at its final value. Examples of such *ramp schedules* are depicted in black on Fig. 4.2(a) and for a given δ_f, the value of $\Omega_{\max}$ and the timings will determine where the transition is crossed. It is worth noting that while such a schedule may not strictly adhere to the formalism described in Eq. 4.11, it can still be connected to it by utilising two schedule functions, denoted as $s_{i/f}(t)$. In the following, we will fix the total duration of the schedules used to $T = 3$ μs for reasons explained later. We will prepare the ground state of a $N = 10$ circle of atoms.

4.2.3 - Standard protocol and constraints

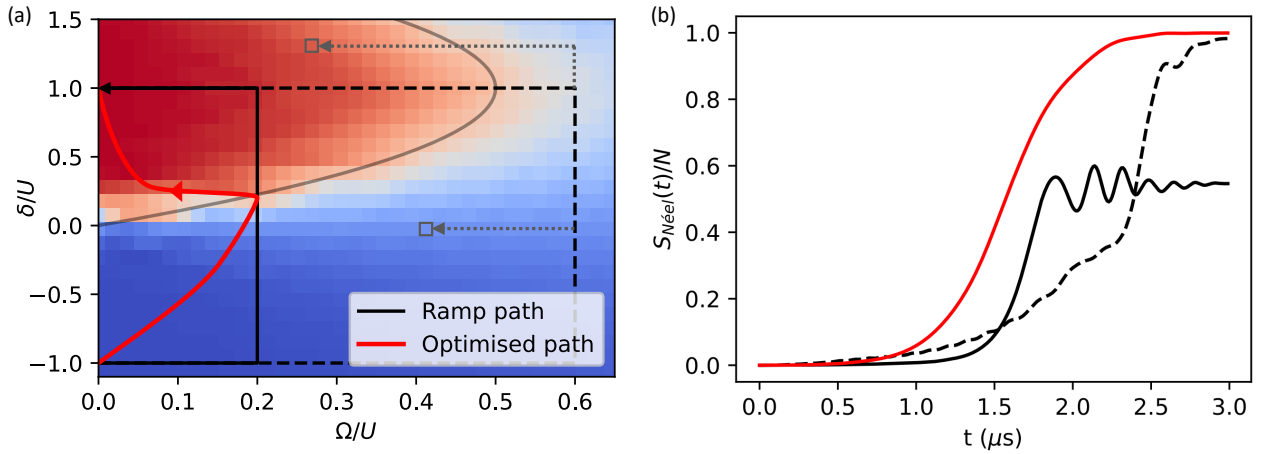


Figure 4.2: **Adiabatically probing the phase space of the 1D Ising model.** (a) The coloured phase diagram is constructed from adiabatically evolving a 6-spin system using ramp schedules (dashed grey) and measuring the structure factor (similar to Fig. 4.1(b)) of the prepared states. A ramp path crossing the transition at its apex is given (dashed black). Two implementable paths respecting the constraints on maximum amplitude and detuning of the device are also plotted: a ramp path (solid black) and an optimised path (red). (b) Evolution of the normalised structure factor $S_{\text{Néel}}(t)/N$ through time for paths given in (a) on system of 10 spins.

One possibility to probe the phase diagram is to enter the phase by its apex as performed with the ramp schedules depicted in dashed line in Fig. 4.2(a). As conveyed by Fig. 4.1(a), higher values of Ω make the gap wider, especially at $\delta/U = 1$ facilitating the adiabaticity of a protocol following such a path. The success of such protocol is shown in Fig. 4.2(b) where the structure factor slowly increases towards its maximum value, showcasing the efficient preparation of $|\psi_{\text{AF}}\rangle$. With similar starting ramp schedules, all going from $(-1, 0.)$ to $(\alpha, 0.6)$ via $(-1, 0.6)$ in the $\delta/U, \Omega/U$ phase space, and then stopping at (α, β) (figured by the marked squares in Fig. 4.2(a)), we can probe the ordering of the

system at any point and reconstruct with implementable protocols the phase diagram of Fig. 4.1(b). However, this implies reaching high values of Ω/U which for limited laser power, and thus limited amplitude Ω, can be hard to achieve. One possibility is to reduce U by spacing more the atoms for example at the expense of the global energy scale of the system and thus timescale of the dynamics. Another possibility is to cross the transition at another point by reducing $\Omega_{\max}$ like with the schedule given in Fig. 4.3(a). In fact, this protocol makes the most out of the possibilities of the Fresnel device used in this chapter. Atoms are spaced by $a = 5$ μm, leading to $U/2\pi = 8.9$ MHz and the laser power gives $\Omega_{\max}/2\pi = 1.8$ MHz and $|\delta|_{\max}/2\pi = 9$ MHz. Thus, any schedule would be constrained within the region delimited by this schedule. These limitations directly impact the efficiency of the preparation as conveyed by the solid black line of Fig. 4.2(b) where the structure factor only reaches around 55% of its maximum value. This decrease can be tracked back to diabatic errors as shown by the energy spectrum of Fig. 4.3(b). At the beginning of the dynamics, only the paramagnetic ground state (blue side of the bottom line) is populated but as the system evolves, more and more population jumps into excited states, as the appearing and thickening lines illustrate.

The probability of having populated the first excited state through the evolution can be obtained using the Landau-Zener formula

$$P_{\text{cross}}(t) = \exp\Big[-\pi\Delta(t)^2/\dot{\Omega}_{\text{eff}}(t)\Big]. \tag{4.14}$$

In essence, fast driving when the gap is small would induce population transfers whereas spectrum regions where the gap remains large can be quickly covered. Computing the evolution of the gap or more specifically its minimum value over the dynamics constitutes a resource consuming problem and can generally only be done for small systems. This crossing probability is shown in Fig. 4.3(c) and reaches almost 50% around the time where the system crosses the transition. At the end of the dynamics, the system remains in a equal superposition of AF ordered ground state (red side of the bottom line) and paramagnetic excited states (above blue lines). Although this straightforward protocol may lose efficiency due to hardware constraints, it is feasible to design an optimised pathway aimed at achieving improved ground state preparation within the same limitations.

4.2.4 - Optimising adiabatic driving towards antiferromagnetic phases

The optimal control algorithm benefits from the classical Bayesian optimiser described in Sec. 3.3 to shape the pulse schedule so as to maximise the ground state preparation.

We resort to the smooth parameterisation described in Fig. 4.4(a) obtained by interpolating the control fields with monotonous cubic splines. This enables to describe a constrained schedule with only $2m$ parameters $\boldsymbol{\theta} = \{\Omega(t_i), \delta(t_i)\}_{i=1\cdots m}$ with times t_i regularly spaced in $[0, T]$. The Bayesian optimiser can thus explore the parameter space defined by the hyper rectangle $[0, \Omega_{\max}]^m \times [\delta_i, \delta_f]^m$. We fix the values of the pulses at $t = 0$ and $t = T$ with $\Omega(0) = \Omega(T) = 0$ and $-\delta(0) = \delta(T) = U$ and use $m = 3$. The state $|\psi_{\boldsymbol{\theta}}(T)\rangle$ produced under the application of such parameterised schedule is obtained using Eq. 1.28. As a metric to optimise, one could settle with the infidelity $\mathcal{I}(\boldsymbol{\theta}) = 1 - |\langle\psi_{\text{AF}}|\psi_{\boldsymbol{\theta}}(T)\rangle|^2$. However, this leads to a rather flat landscape, especially for increasing N, as the probability of obtaining a non-zero projection approaches 0. While this may still work for this numerical study, this would not be viable for an implementation where the resolution of the probabilities is limited by the number of samples obtained as conveyed in the next section. The structure factor constitutes a metric more appropriate as it can capture emerging order and not only perfect arrangement of spins.

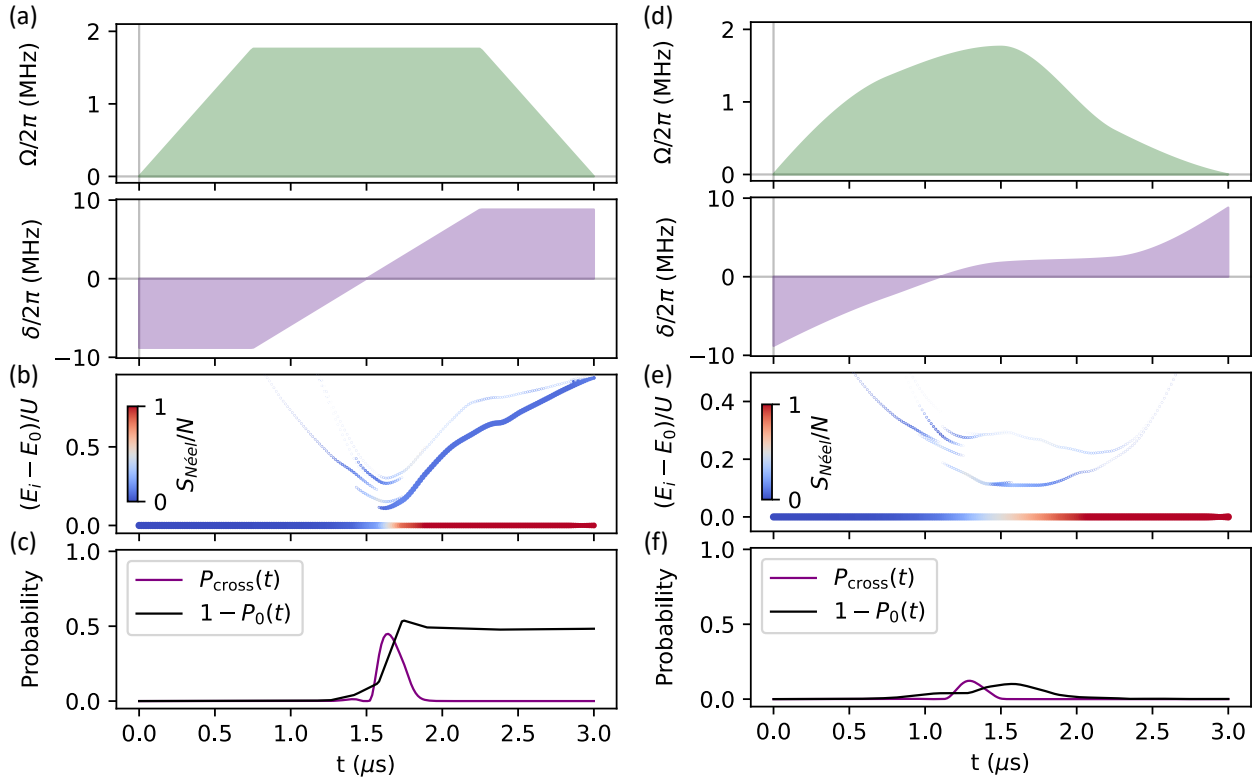


Figure 4.3: **Adiabaticity of control schedules for ground state preparation.** (a) Profiles of amplitude Ω and detuning δ for a ramp schedule. (b) Instantaneous energy spectrum of a 10-spin system with levels coloured by structure factor and thickened by population (levels are only appearing when populated). (c) Evolution of the probability from having jumped from ground to first excited state $P_{\text{cross}}(t)$ (purple) and the probability to having populated excited states $1-P_0(t)$ (black) through the schedule. (d,e,f) Same for optimised schedule.

The optimiser is thus tasked to minimise $1 - S_{\text{Néel}}/N$ within $n_c = 300$ iterations, using the first $n_r = 30$ iterations as initialisation. Its convergence is displayed in Fig. 4.4(b), reaching a maximum structure factor of $0.9994N$. While the best minimum found keeps on improving with the number of iterations (red line), the optimiser regularly tries to explore unseen regions of the parameter space as illustrated by the coloured points. When it finishes to exploit a known region (dark points) with good scores (such as around iteration 50 or 100) it then moves to new regions (brighter points) which may result in higher scores at first but ultimately leads to better values. As the optimisation carries on, the optimiser starts to converge towards interesting regions, not trying anymore to explore new ones (or having efficiently covered the search space) as shown by the disappearance of yellow spots. This example illustrates how the Bayesian optimiser avoids getting stuck in local minima and thus proves very useful when exploring unknown landscapes. The optimised path is shown in Fig. 4.2(a) as well as the corresponding evolution of the structure factor. It is worth noting that the optimiser has no prior knowledge on where the transition lies but adapts the pulse so as to move fast when in

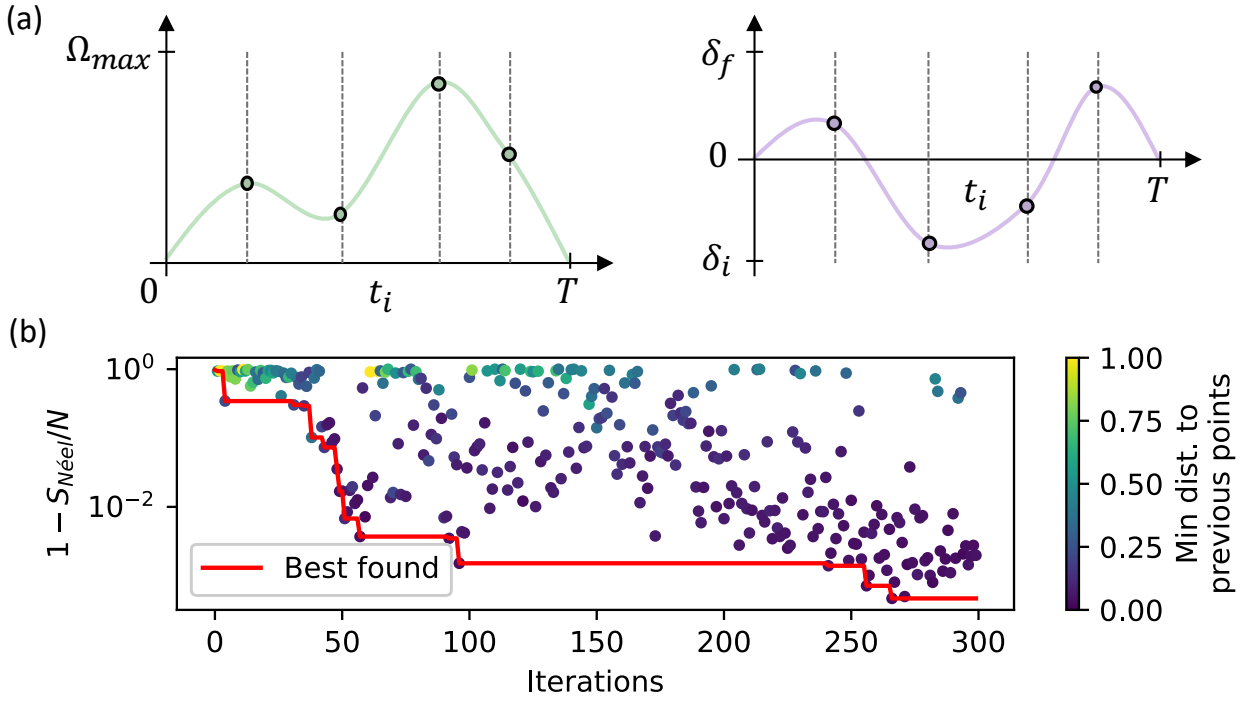


Figure 4.4: **Optimal control for ground state preparation using Bayesian optimisation.** (a) Parameterisation of the controls with bounds set by $\Omega_{\max}$ and $\delta_{i/f}$. (b) Convergence study of the optimiser over 300 iterations. Each point representing a tried schedule is coloured by its normalised minimal distance to previously tried points. The best value found is tracked down during the process (red).

the paramagnetic phase and then slow when crossing the transition. This results in fewer diabatic errors as conveyed by Fig. 4.3(e,f) with P_{cross} peaking at 5% and a bit of leakage in the excited states. Another interesting feature of the optimised pulse is the increased sweep rate while deep in the AF phase as conveyed by the increasing final slope of the detuning δ in Fig. 4.3(e). This seems to enable to repopulate the ground state as $P_0(t)$, after increasing a bit during the transition crossing, returns to 0 at later times. Overall, with the same constraints as for the ramp schedule we can find a schedule with improved preparation efficiency even when not staying fully adiabatic. Another possibility would be to increase T at a fixed size N as discussed below.

4.3 - Preparation on hardware with scaled duration

4.3.1 - Emulation on 18-site hexagon

As conveyed by Eq. 4.12, for a given schedule, it is possible to lower the diabatic errors by increasing the total duration T. To emphasise this effect, we turn to a system of $N = 18$ atoms placed in a 4-sites sided hexagon as displayed in the inset of Fig. 4.5. This choice of geometry stems from a punctual constraint of the device used where the trap pattern was fixed to a triangular lattice. This periodic system, although different from a perfect circle, exhibits a similar phase diagram as the next nearest neighbours terms are only 3% or 1.5% of U. The hexagon features 4 atoms per side so as not to favour one AF state over the other. Indeed, with $N = 12$ for instance and 3-sites sides, the alternate configuration with spins in $|\downarrow\rangle$ in the corners would have been slightly favoured due to the geometry. To drive this system towards its AF ground state we apply a schedule obtained by numerical optimisation with similar constraints as before. As the system size starts to become significant, we perform the optimisation using the Matrix Product State method described in Sec. 2.2. To keep emulation times down, we restrict the number of iterations to 50, leaving the optimiser little time to work through the parameter space.

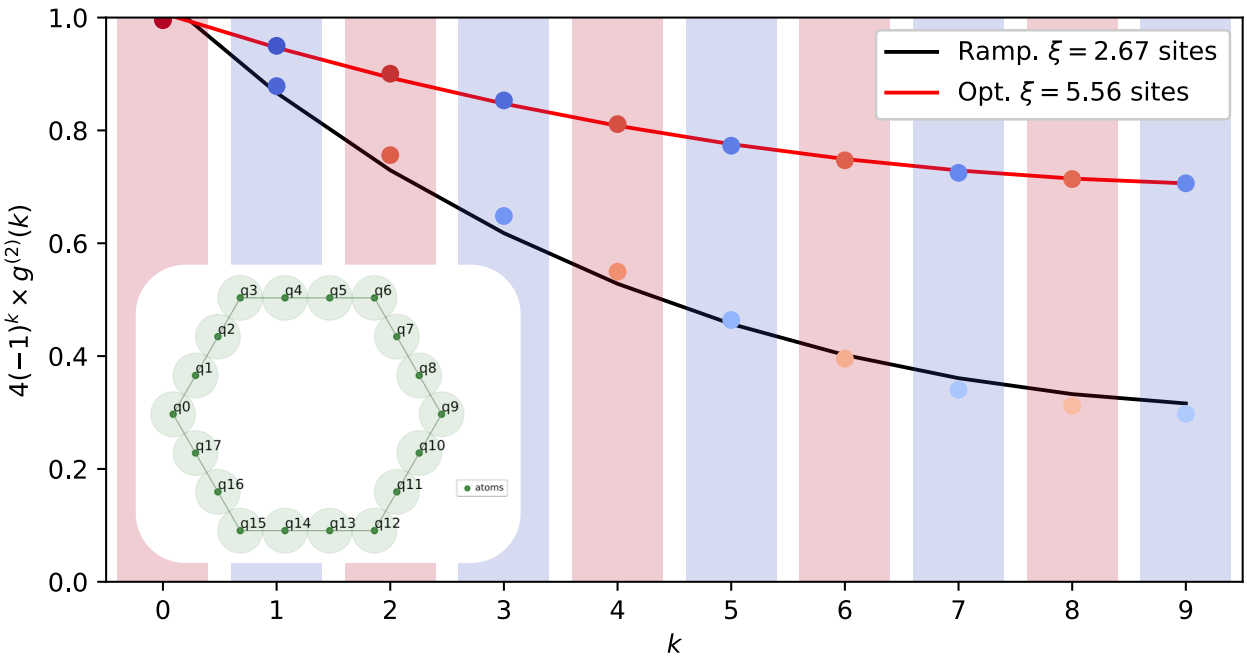


Figure 4.5: **Correlation functions of prepared ground states** for (black) ramp and (red) optimised schedules over a $N = 18$ hexagon. Points are coloured according to $g^{(2)}(k)$ (red being positive and blue negative) and exponential fits (solid lines) gives the correlation length obtained for each method.

Fig. 4.5 depicts the correlation function $g^{(2)}$ measured on the final states obtained either after a ramp protocol or the optimised protocol. The value of $g^{(2)}(k)$ describes how correlated are pairs of spins distant of k sites on the hexagon. While a perfect $|\psi_{\mathrm{AF}}\rangle$ state can be characterised by

$g^{(2)}(k) = (-1)^k/4$, both prepared states exhibit a decay of their correlations with the distance, meaning that more distant spins influence each other less. This decay can be fitted exponentially by $f(k) = A\exp(-2k/\xi)$ on an open chain and by $f(k) = A[\exp(-2k/\xi) + \exp(-2(N-k)/\xi)]$ on a periodic one with ξ being the *correlation length*. ξ is another order parameter for the quantum phase transition and should diverge in the AF phase [134]. We obtain a correlation length of 5.56 sites for the state obtained with the optimised protocol lasting $T = 3$ μs and around half as much with the ramp.

4.3.2 - Stretching a pulse for adiabaticity benchmark

We could boost this value by conducting further optimisation steps, or more simply by stretching the pulse to extend its duration as represented by the insets of Fig. 4.6. Emulation results (small squares) typically reveal that the correlation length $\xi(T)$ tends to increase with T within the range $T \in [1\mu\text{s}, 6\mu\text{s}]$. Doubling the total duration from 3μs to 6μs enables to also double ξ up to 11 sites, more than half the size of the system. We also see that correlation length and structure factor are strongly correlated as expected. Longer pulses seem to benefit the adiabaticity. However, when experimentally running the same protocol (circles), we notice a consistent reduction across all durations. More notably, we observe a plateau ($\xi \sim 2.3$) at intermediate values and even a slight decline at longer durations ($\xi \sim 2.15$). We first check that adding detection errors (squares) $\varepsilon = 1\%$ and $\varepsilon' = 8\%$ only scales down $\xi(T)$ without modifying the increasing trend. Better matching the decay of the experimental data thus requires adding decoherence into the picture.

4.3.3 - Noisy benchmark on a small hexagon

In the following, we provide a benchmark on a smaller hexagonal systems for which we can perform extensive noisy simulations and assess how low-level experimental errors can impact high-level observables such as the correlation length. We apply with the Fresnel device the previous schedule optimised for $N = 18$ on three small isolated hexagons with $N = 6$ atoms each spaced by 5 μm. This *stamping* method enables to acquire statistics faster by multiplying isolated copies of small enough systems when large trap layouts are available. We ensure that the hexagons are located at $d \sim 20$ μm from each other as to ensure they do not interact. We select a stretching $T = 4.3$ μs as this seems to be the optimal time before the decay of ξ in Fig. 4.6 and extend the schedule with constant δ for few hundreds of ns at start and finish.

Fig. 4.7 encompasses numerical results with and without noise taken into account as well as the experimental results averaged over 500×3 repetitions as to ensure limited shot noise. Perfect emulation of the system shows, for all considered observables, steady increase towards expected values : (a) $\langle\hat{n}\rangle(T) = 0.5$ since all 6 sites are symmetrically half-filled, (b) $S_{\text{Néel}}(T) = N$ and (c) diverging ξ as the system orders. The experimental data exhibits similar behaviour but with limited contrast : $\langle\hat{n}\rangle(T) = 0.45$, $S_{\text{Néel}}(T)/N = 60\%$ and $\xi/N = 48\%$. A slight delay in the dynamics was noted, yet it was verified that this was not caused by inadequate synchronisation of the pulses on the hardware. In order to explain this contrast loss, we introduce two noise models:

- a *real* two-level model incorporating $\varepsilon = 1\%$, $\varepsilon' = 8\%$, $\eta = 1\%$, $\sigma_\Omega = 3\%$, Rydberg decay, finite waists of the beam with $w_{\text{eff}} \approx 180$ μm as well as temperature $T_{\text{atom}} = 100\mu$K induced effects such as Doppler shifts and thermal motion. To mitigate the effects of the latter, an adiabatic ramping down of the traps was performed on the hardware, so that the atoms are slowed down but get spread out more, resulting in $\sigma_v(T_{\text{atom}}/5) \approx 44$ nm/μs and $\sigma_r(5T_{\text{atom}}) \approx 0.35$ μm. We

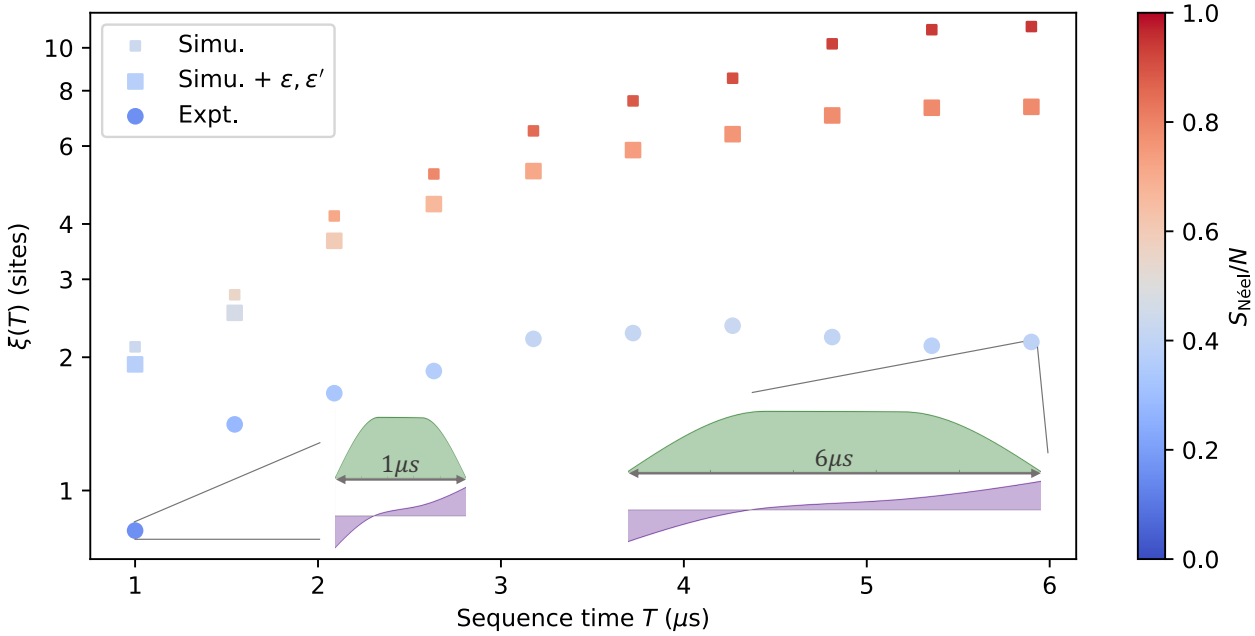


Figure 4.6: **Adiabaticity study by stretching the schedule duration.** The correlation length $\xi(T)$ obtained for stretching a pulse (insets) to reach a given total duration $T \in [1\mu s, 6\mu s]$ is computed numerically with (squares) and without (small squares) detection errors and is benchmarked to experimentally obtained data (circles). The points are coloured according to their normalised structure factor.

consider time-dependent interactions as the atoms are moving and the schedule duration is such that $\sigma_v T \sim \sigma_r$. Finally, we consider the laser phase noise and other low-contribution noises with an effective dephasing term $\Gamma/2\pi = 1/7.5$ MHz which is our only free parameter as the other values have been independently measured on the device.

- an *effective* two-level model incorporating $\varepsilon = 1\%$, $\varepsilon' = 8\%$, relaxation with $T_1 = 100$ μs and dephasing with $T_2 = 4.5$ μs measured with a Ramsey experiment.

This kind of adiabatic preparation is very robust under shot-to-shot fluctuations (including the Doppler shift, and fluctuations in position) as the latter only contributes slightly to the decay of the structure factor. The real noise model manages to explain the experimental data and even reproduce the delay in the dynamics of ξ as conveyed by overlapping curves in Fig. 4.7(c). The limited preparation of AF states can thus be explained by the physical processes happening in the device and adding incrementally the noise one by one enables to track down their individual contribution such as performed in Sec. 5.2.3 for the XY model. More practically, the effective model adeptly mirrors the behaviour of the real model, especially concerning the structure factor in Fig. 4.7(b), amalgamating preparation errors, temperature influences, time-dependent interactions, and laser phase noise into a single dephasing term T_2. This confirms that characterising the noisy operation of the device can be limited to measuring these 4 effective parameters for this kind of experiment and therefore for larger systems,

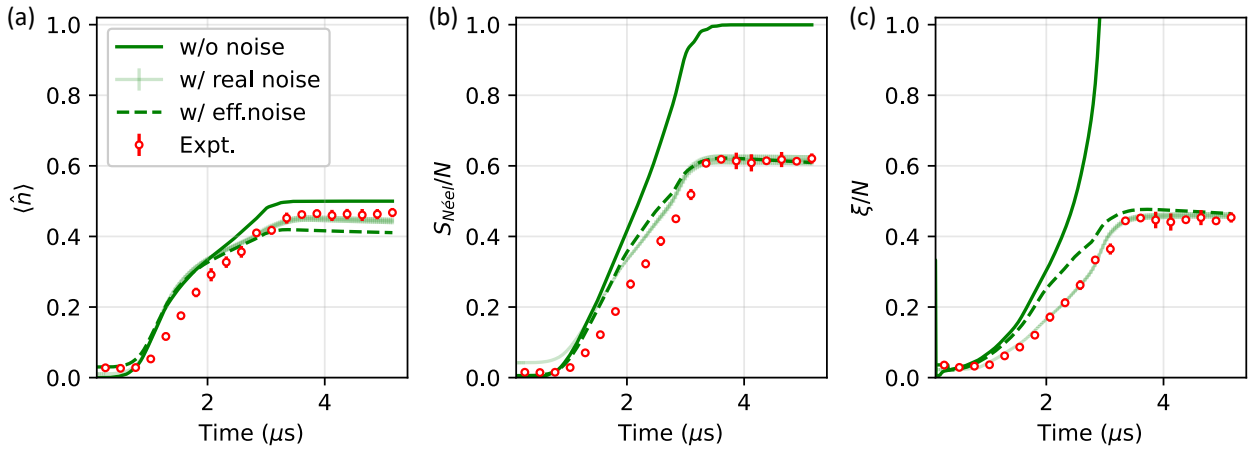


Figure 4.7: **Noisy benchmark of adiabatic preparation on Fresnel class device.** Temporal evolution of (a) the mean site-occupancy $\langle \hat{n} \rangle$ (b) the normalised structure factor $S_{\text{Néel}}/N$ and (c) the normalised correlation length ξ/N for numerical emulation without noise (solid green), with detailed noise model (light green), effective noise model (dashed green) and experimental data (red) obtained on Fresnel.

one can use the Monte Carlo Wave Function method [81] described in Sec. 2.3.1 to implement such a model.

4.4 - Closed-loop on hardware for improved correlation length

4.4.1 - How to improve hardware results

Three possibilities exists to improve the preparation efficiency on a device with assessed noise intensity. The first one entails identifying the primary contributions by incrementally introducing noise in the descriptive model, thus creating a roadmap for potential hardware enhancements. While this will be performed to improve the Fresnel device in a near future, other options are available if we choose to maintain these specifications. We can for instance perform a numerical schedule optimisation using the effective noise model. Such a process can be quite resource-intensive, particularly for large systems. We could consider running such optimisation on a noisy version of a tensor network algorithm, but with current simulation resources, this could potentially take several days. Finally, we can perform a closed loop optimisation directly on the device which would output a schedule fine tuned for the current noise level. This comes with two main limitations :

- At current repetition rate, probing a system of $N = 18$ atoms with finite resolution, for instance $n_{\text{shots}} \sim 1000$, takes around 20 minutes. For larger systems, since efficient quantum state estimation should theoretically require a number of measurements exponential with N and since the effective repetition rate also decreases with N as rearrangement fails more, probing a state can become quite time consuming. If probing one point lasts hours or days, drifts of parameters hitherto unconsidered could occur. The machine therefore requires regular automatic recalibration, extending further the duration of the implementation.

- In case of hardware miscalibration, such as $\delta \to \delta + 2\pi \times 1$ MHz, the landscape of optimisation could be altered significantly and an optimised schedule could not be used when the device is later properly recalibrated. Similarly, if the noise level improves or deteriorates, applying the optimal schedule might no longer yields the same efficiency as before.

4.4.2 - Building a closed-loop for state preparation

We construct a closed loop with the experiment to optimise the structure factor of the $N = 18$ hexagon presented in Fig. 4.5. The classical part is handled by a Bayesian optimiser trusted with $n_c = 180$ iterations and $n_r = 20$ dedicated to the initialisation. The parameterisation of the chosen schedules is done using Pulser and the communication with the QPU is managed through the cloud platform of Pasqal as described in Sec. 2.1.2. While the start and ending values of Ω are still fixed at 0, we let the optimiser locate the best starting and ending values of δ within $[-U, U]$. Furthermore, the total duration $T \in [1\mu\text{s}, 6\mu\text{s}]$ also becomes a parameter so that the optimiser can balance adiabaticity and decoherence on its own. Slicing the duration in 4, this gives a total of $3+5+1 = 9$ parameters to optimise. We ask for $n_{\text{shots}} = 1250$, so as to retrieve around $n_{\text{shots}} \times 80\% = 1000$ statistics on average, accounting for failed rearrangement attempts. The loop thus takes around 63 hours to complete without counting the recalibration jobs regularly interleaved and without any break in communication between the classical and quantum parts at any time, highlighting the robustness of the experimental device.

4.4.3 - Experimental results: convergence and outlook

The schedule obtained via the closed loop protocol (see Fig. 4.8(a)) enables to improve the maximum ξ obtained from 2.3 to 3 as shown in Fig. 4.8(b). As expected after the stretch study of Fig. 4.6, the optimised schedule only lasts 4.53 μs as longer dynamics would be detrimental. More surprisingly, it prefers to wait deep in the paramagnetic phase up to 1 μs before starting to ramp up the amplitude and apply a more dense profile of Ω than the schedule of Fig. 4.3(e), not slowing down at the phase transition. Either the resulting diabatic errors are not the limiting factor of the preparation and the optimiser understood it or when trying to mitigate a decoherence-inducing phenomenon, the optimiser fails to see that smoother schedules could improve the picture. Overall, the optimiser still converges quite fast towards preparing ordered states as visible bands of alternating correlation function are appearing up to $k = N/2$ after only 50 iterations. We manage to improve the correlation length after 95 iterations and the structure factor goes up to 48% of N at iteration 175. Moreover, using Bayesian optimisation, we can extract the parameters the figure of merit depends most on, i.e. T and $\delta(3T/4)$. The latter should thus be prioritised when aiming to build on this optimised schedule, either with additional optimisation steps only on those two parameters, or when building a similar loop for another instance. Another way forward would be to parameterise the schedule with twice as many points: half would be fixed at the optimal values obtained here while the new ones would be optimised to check whether an intermediate profile other than a cubic spline might be beneficial. Finally, another possibility would be to distribute the time intervals irregularly, for instance by giving the optimiser more freedom around the phase transition, while constraining it more at the start. This is supported by our observation that a change in $\delta(T/4)$ does not seem to have any impact on the score.

Overall, the optimised schedule output by the closed-loop scheme makes the most out of the noisy

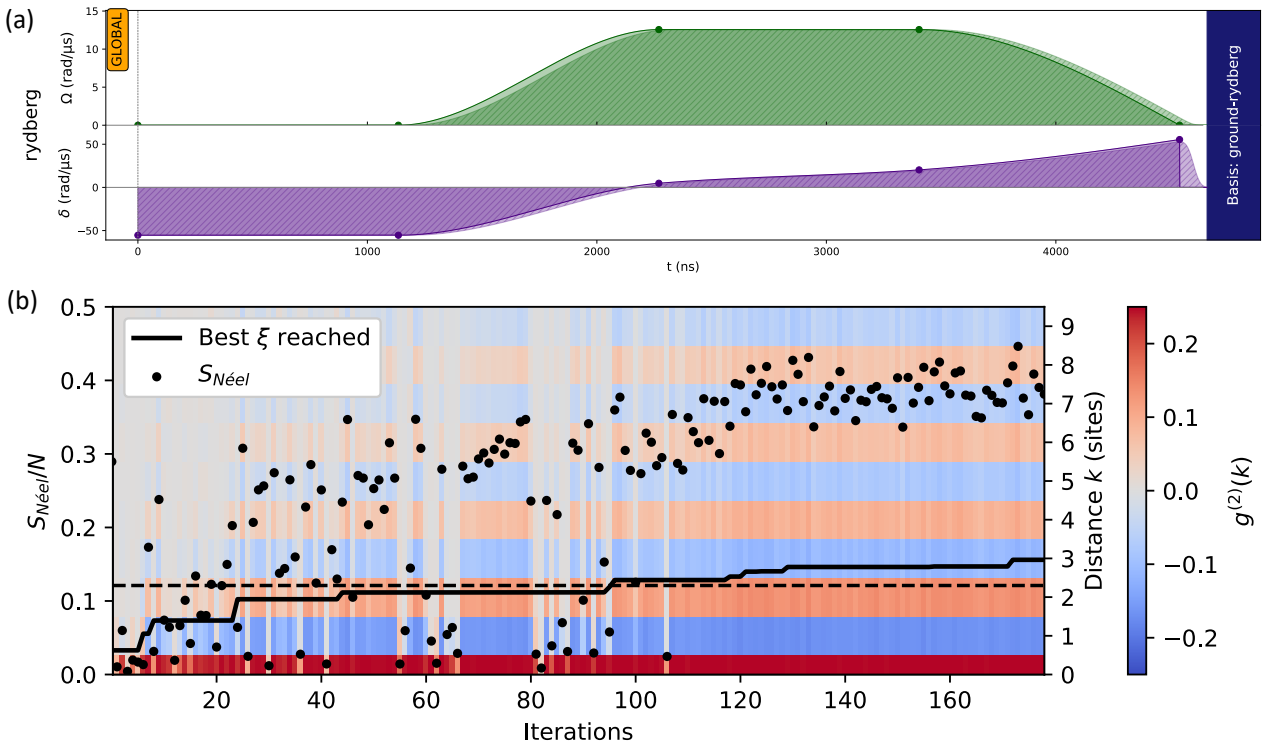


Figure 4.8: **Experimental results for an optimisation closed loop for AF state preparation.** (a) The optimised amplitude (green) and detuning (purple) of the schedule to apply. (b) Evolution of the structure factor with the number of iterations performed (dots) as well as best correlation length found (solid line) compared to the best one obtained by stretching (dashed). The correlation function $g^{(2)}(k)$ of the state prepared at each iteration is also plotted.

neutral atom QPU in order to prepare as best as possible an antiferromagnetic state, regardless of miscalibrations or noise processes that are hard to emulate. The discrepancies between this protocol and the ones optimised on emulation give a better understanding of the limitations of hardware. Either the latter remains unchanged and this precise protocol can then be used as a building block for experiments requiring initialisation in an antiferromagnetic state, or the hardware is recalibrated to allow the use of theoretically obtained pulses.

4.5 - Many-body localisation in 1D Ising model

Previous sections solely focused on the properties of the ground state of the 1D Ising model and additional excitations above this ground state were considered as diabatic errors introduced by imperfect adiabatic protocols. However, with sudden variations of the control parameters, or *quenches*, one can also probe the intriguing phenomena occurring in quantum many-body systems out of equilibrium such as quantum scars [139]. When quenching a quantum system originally in its ground state, the instantaneous state is no longer an eigenstate of the Hamiltonian and non-trivial dynamics make it difficult to predict whether the system will relax towards its equilibrium or not. The eigenstate thermalisation hypothesis posits that individual eigenstates of a quantum system with many degrees of freedom exhibit thermal behaviour, similar to what is expected from the micro-canonical ensemble of statistical mechanics [140]. *Ergodic* systems will thus thermalise reaching a state where macroscopic observables converge towards their thermal values. In contrast, many-body localisation (MBL) occurs when disorder or strong interactions prevent a quantum system from thermalising, even in the presence of interactions [141]. In an MBL system, the dynamics are effectively frozen, and the system retains memory of its initial conditions over long timescales. This lack of thermalisation arises from the presence of strong localisation effects, which inhibit the transport of energy and prevent the system from reaching thermal equilibrium. In this section, we explore the possibility of experimentally observing a MBL transition in the out-of-equilibrium dynamics of the 1D Ising model when a source of local disorder is incorporated.

4.5.1 - Preparing antiferromagnetic disordered chains

Similarly to [142, 143], we consider an excitation transport experiment. The protocol consists in quenching a disordered chain of N atoms that has been initialised in one of its antiferromagnetic ground states using a global pulse with constant parameters Ω and δ. The initial state, for instance $|\psi_0\rangle = |\psi_{\mathrm{AF}\uparrow}\rangle$, can be characterised by its maximum staggered magnetisation defined in Eq. 4.3. By monitoring this one-body observable during the dynamics, we can check if the introduced excitations localise and assess the degree of disorder a system requires to retain memory of its initial state instead of thermalising.

Introducing disorder with random positioning/local light shifts

We introduce disorder in the interaction terms of the Hamiltonian by adding a random component in the positions of the atoms of the chain of spacing a as described in Fig. 4.9(a). The nearest-neighbour interactions $U_{i,i+1}$ are uniformly spread [142] such that

$$U_{i,i+1} = U(1 + w_i) \quad \text{, with} \quad w_i \sim \mathcal{U}(-W, W), \tag{4.15}$$

where $U = C_6/a^6$ is the regular coupling and W is the disorder strength with $0 < W < 1$. Weakly disordered chains ($U_{i,i+1} \approx U$) can be obtained for $W \approx 0$. The positions of the N atoms are obtained sequentially by positioning the atom i, drawing a value of w_i, converting $U_{i,i+1}$ into a distance $r_{i,i+1}$ and placing the atom $i+1$ such that the pair is separated by $r_{i,i+1}$. This straightforward method only works in 1D and subsequent works in 2D would require a mapping technique in order to find from a pairwise distance matrix the adequate positions in the plane (see Box. 27). Fig. 4.9(a) displays the atom positions probability densities leading to uniformly random couplings $U_{i,i+1}$. Another possibility

to introduce disorder in the system consists in randomly light shifting each site of the chain, using a local addressing scheme with an additional SLM. This amounts in adding to the Hamiltonian random local detunings terms δ_i (see Box. 12).

In both cases, generating random Hamiltonian requires changing an SLM pattern between each instance, either for setting the traps at random positions or generating the local light shifts. Both methods yield the same cost.

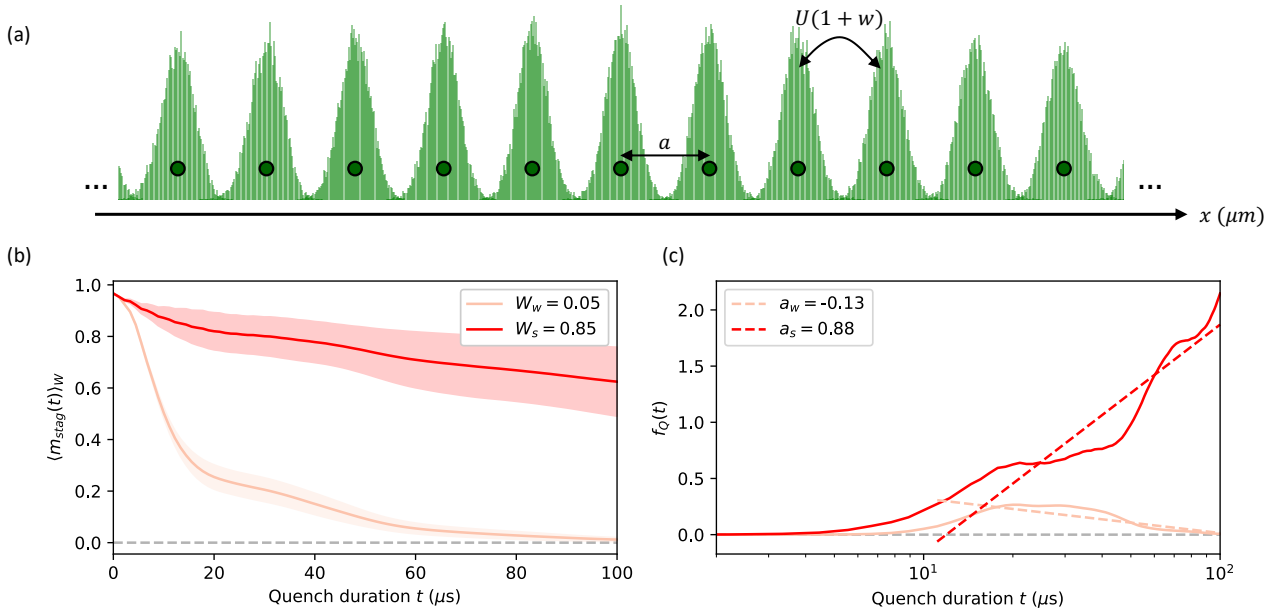


Figure 4.9: **Many-body localisation example on a 1D disordered chain of atoms.** (a) The atoms of a regularly-spaced chain (green dots) can be randomly perturbed in positions (green histograms) in order to introduce uniform disorder in the interaction part of the Hamiltonian. (b) Temporal evolution of the staggered magnetisation $\langle m_{\text{stag}}(t)\rangle_W$ averaged over 50 instances of disordered chains for (beige) weak or (red) strong disorder. (c) Temporal evolution of the quantum Fisher information $f_Q(t)$ for (beige) weak or (red) strong disorder. A logarithmic fit (dashed) is performed to characterise the speed of the growth/decay in both cases.

Initial staggered state preparation

On the contrary to the previous section in which the target state was $|\psi_{\text{AF}}\rangle$, the equal superposition of the two antiferromagnetic states $|\psi_{\text{AF}\uparrow}\rangle$ and $|\psi_{\text{AF}\downarrow}\rangle$, we need to initialise the system in $|\psi_{\text{AF}\downarrow}\rangle$ only. Two distinct methods can be used:

- an adiabatic driving towards $|\psi_{\text{AF}\downarrow}\rangle$ using an odd-numbered chain. In this case, the system exhibits a very similar phase diagram as the one displayed in Fig. 4.1(b) but with a phase characterised by $\hat{m}_{\text{stag}}$ where $|\psi_{\text{AF}\downarrow}\rangle$ is the ground state. Thus after preparing the latter on a long enough chain, we can asymmetrically truncate the sides so as to remain with an even-numbered of atoms, and avoid edge effects in the meantime.

- a local addressing scheme utilising an additional SLM to light shift all even sites. Applying a global resonant π pulse such that $\Omega \gg U_{i,i+2} \approx U/64$ will only flip the spins located at odd sites. This procedure displays the key advantage of being faster than the previous one, as it does not rely on adiabaticity and hence can mitigate decoherence effects. However it requires to be able to change the light shift weights on each site in between the preparation and the quench. Such a scheme is in fact used in Chapter 5 for preparing checker-board patterns of alternating spins in the XY model.

In order to highlight the MBL effects we need to perform the state preparation + quench protocol on a significant number of randomly drawn instances of disordered Hamiltonians. With only global addressing available, we can resort to the previous optimisation scheme to efficiently prepare the ground states of many disordered chains of atoms. While for weakly disordered chains, the schedule optimised for no disorder still yields good efficiency on all instances, chains strongly disordered might cause some issues, as pairs of atoms can become isolated and modify the overall ground state. Optimising a schedule for each instance proves to be extremely costly. Choosing an intermediate approach, one can rather try to optimise the same pulse on a subset of instances and then use it on the rest. A device equipped with local addressing, however, will only need a global π pulse for the preparation on each instance, potentially saving a lot of resources.

4.5.2 - Many-body localisation properties of the quenched dynamics

Previous work [143–145] suggest that one should observe an ergodic to MBL transition for the system under consideration.

Memory of the initial state in the staggered magnetisation

We perform a numerical study of a 1D system of $N = 30$ atoms in the case of weak and strong disorder, using $W_w = 0.05$ and $W_s = 0.85$ respectively. The dynamics of these relatively large systems can be emulated using the time-evolving block-decimation approach described in Fig. 2.5(a). For both cases, we generate 50 chains disordered in positions and initialise them perfectly in $|\psi_{\mathrm{AF}\downarrow}\rangle$. We apply a global control pulse lasting 100 μs with parameters $\Omega/U = 0.17$ and $\delta/U = 2$, effectively quenching the system outside of the antiferromagnetic phase. In the case of weak disorder, the system should be ergodic and the staggered magnetisation averaged over the realisations, noted $\langle m_{\mathrm{stag}}\rangle_{W_w}$, should return to 0. On the contrary, for strong disorder, a memory phenomenon should appear and $\langle m_{\mathrm{stag}}\rangle_{W_s}$ should deviate from its initial maximal value with longer timescales, showing MBL properties.

Fig. 4.9(b) showcases such behaviours. On one hand, $\langle m_{\mathrm{stag}}(t)\rangle_{W_w}$ quickly deviates from 1 and reaches 0 after 60 μs. The imbalance between odd and even sites thus quickly disappears in the case of weak disorder. On the other hand, $\langle m_{\mathrm{stag}}(t)\rangle_{W_s}$ exhibits a rather flat evolution, reaching only 0.7 after 100 μs. This macroscopic observable remains finite even at long times in the strong disorder case.

Growth of entanglement

Another well-known measure to characterise the MBL property of a 1D system is the logarithmic growth of entanglement entropy between the two halves of the chain [132]. Since this quantity is not an experimentally-accessible observable we focus instead, following [143], on the quantum Fisher information f_Q associated with the staggered magnetisation and defined by

$$f_Q(t) = \frac{4}{N}\left(\langle m_{\text{stag}}(t)^2\rangle_W - \langle m_{\text{stag}}(t)\rangle_W^2\right). \tag{4.16}$$

This quantity is a lower bound on the entanglement contained in the system, while presenting the advantage of only requiring measurements of two-body observables in the z-basis. Fig. 4.9(c) displays the logarithmic increase of f_Q in time for the strong disorder case while in the weak case, the value decays back to 0. We can fit the growth/decay with a logarithmic function of the form $a_W \log(t) + b_W$ and get $a_w = -0.13$ and $a_s = 0.88$, highlighting the distinct behaviour between the two cases at long times.

Varying the disorder strength

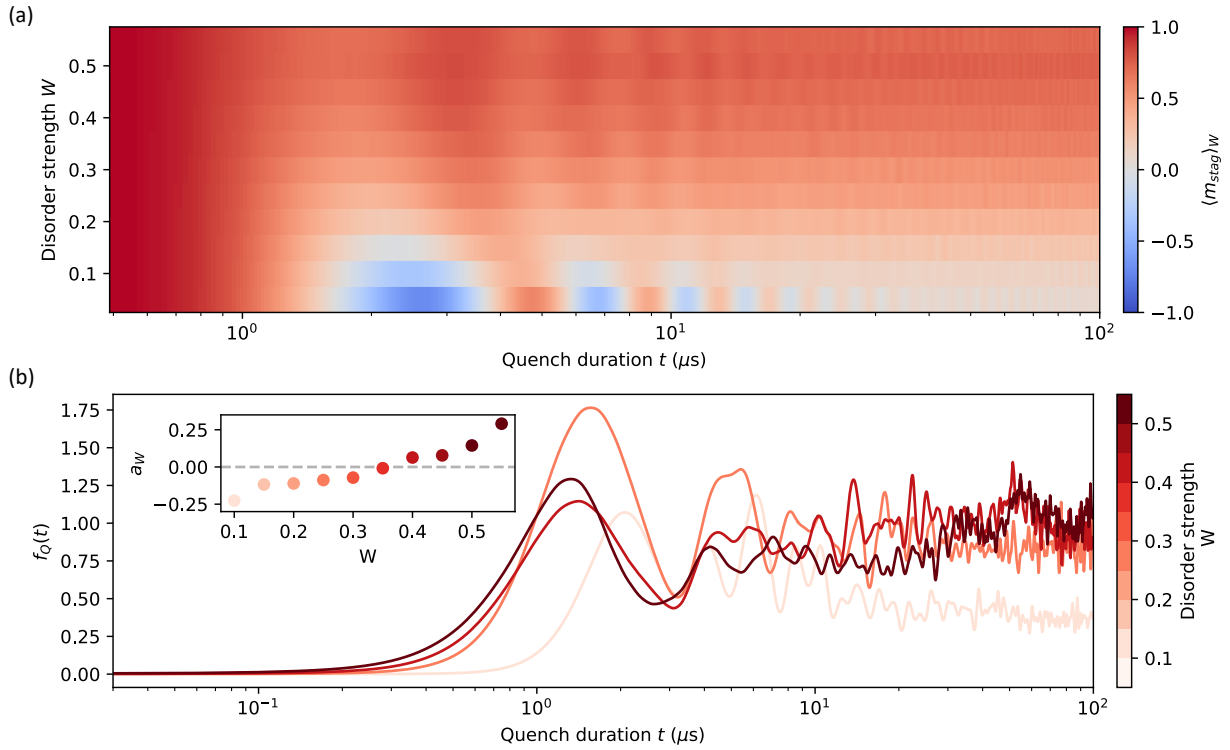


Figure 4.10: **Sweeping over the disorder strength.** (a) Color map of the temporal evolution of the staggered magnetisation $\langle m_{\text{stag}}(t)\rangle_W$ averaged over 50 instances of disordered chains for various values of disorder strength between 0.05 and 0.55. (b) Temporal evolution of the quantum Fisher information $f_Q(t)$ for various values of disorder strength between 0.05 and 0.55. (inset) Logarithmic growth constant a_W obtained by fitting each corresponding curve between 30 μs and 100 μs.

In order to observe the evolution of the growth with the disorder, we repeat the procedure given above for a smaller system of $N = 10$ atoms, varying W between 0.05 and 0.55. This smaller system, although bound to suffer from more significant edge effects, can be quickly emulated using Pulser. Fig. 4.10(a) shows the evolution of the staggered magnetisation during the quench. Starting at its maximum value for all disorder, it quickly starts exhibiting damped oscillations. As disorder levels increase, the maximum contrast of the oscillations diminishes, and the ultimate value towards which they converge increases linearly with W. The greater the disorder within the system, the more memory of its initial state it preserves. Fig. 4.10(b) features the evolution of the Fisher information for several values of W. While the small size of the system perturbs the behaviour a bit compared to the larger case, it still illustrates the discrepancies between weakly and strongly disordered cases. As depicted in the inset displaying the fitted logarithmic growths a_W, below $W = 0.35$, $f_Q(t)$ decreases more rapidly as the system becomes more ordered. However, above $W = 0.35$, it exhibits faster growth with increasing values of W.

This experiment requires several features, such as the ability to study quenched dynamics over long timescales, fast updating of the SLM pattern, and efficient preparation of antiferromagnetic states. Given the current state of neutral atom QPUs, performing such a task would demand a colossal allocation of resources. The Rydberg lifetime ~ 100 μs would not be negligible and even if an isolated system should remain staggered in theory, decoherence would strongly impact the decay of m_{stag}. A further limitation arises from the available field of view, constraining the maximum radius at which the atoms can be placed. For few atoms, a linear chain remains suitable, but it becomes necessary to bend it in more complex shapes as the number of atoms increases so as to fit inside the field of view of the objective. For the strongly disordered case, this might cause unexpected interactions between supposedly distant atoms in the chain. Finally, obtaining a curve similar as in the inset of Fig. 4.10(b) would require $n_W \times n_{\text{instances}} \times n_{\text{timesteps}} \times n_{\text{shots}}$ shots. Assuming that the SLM pattern can be changed within a few minutes between each instance, and having a repetition rate of 1 Hz, this would take of the order of 100 hours to compute. While this small numerical study is limited to relatively small 1D systems, it can be interesting with a QPU being able to reach higher number of qubits to tackle similar problems in 2D [141, 146] as it should exhibit more complex phase diagrams.

5 - Order and frustration in dipolar XY model

Contents

The quantum XY model plays a significant role in condensed matter physics, providing valuable insights into the behaviours of frustrated magnetic systems. Unlike its Ising counterpart, the XY model has a more pronounced quantum nature, with no classical configurations for its ground states. Understanding its phase diagram, and associated phase transitions across various geometric configurations, is essential for understanding exotic phases of matter, notably spin liquids [147], where frustrated spin interactions are key elements. Additionally, exploring quantum many-body systems with long-range interactions offers opportunities to uncover richer physics compared to those with short-range interactions. Systems with dipolar interactions, in particular, have been linked to stabilising various intriguing phases, including supersolids [148, 149], the Haldane phase [150], and spin ice [151]. Understanding the precise interplay of competing instabilities within these systems and how they lead to order and entanglement remains an active research field [152].

In this chapter, we operate the programmable Rydberg platform Chadoq as a dipolar XY quantum simulator. We first probe the ferro- and antiferromagnetic ordered phases on a two-dimensional square lattice, elucidating the spontaneous breaking of the continuous rotational symmetry inherent in the model. In a second part, we enrich the capabilities of the simulator by incorporating arbitrary local controls in the Rydberg-Rydberg encoding. This enhancement enables us to measure multi-body observables and thus perform quantum state tomography of correlated states. Additionally, we explore the dynamics of frustrated geometries, exemplified by a pair of triangular plaquettes.

The following chapter draws from the following articles to which I contributed:

- [37] C. Chen et al., "Continuous symmetry breaking in a two-dimensional Rydberg array," Nature, vol. 616, pp. 691–695, 2023.
- [58] G. Bornet et al., "Enhancing a Many-body Dipolar Rydberg Tweezer Array with Arbitrary Local Controls." arXiv, 2024.

5.1 - Ordered phases from continuous symmetry breaking

Spontaneous symmetry breaking underlies much of the classification of phases of matter and their associated transitions [153–155]. The nature of the underlying symmetry being broken determines many of the qualitative properties of the phase. For instance, in contrast to the discrete case, the breaking of a continuous symmetry leads to the appearance of gapless Goldstone modes, influencing thermodynamic stability [156, 157]. Various constraints can prevent symmetries from breaking in many-body systems. For instance, long-wavelength fluctuations rule out the breaking of continuous symmetries in low-dimensional systems with short-range interactions [158–162]. However, long-range interactions can either stabilise finite-temperature orders, which would otherwise be forbidden [163–166], or lead to frustration, where interactions compete with one another, preventing the formation of order [167–171]. They can thus alter the picture [172] in a non trival way, and generate orders which can be fundamentally distinct, regarding the dispersion of excitations or the decay of correlation functions [165, 173, 174].

Quantum simulators are ideally suited to study these features and while ultra-cold atoms in optical lattices have already investigated continuous symmetry breaking with contact interaction [175], dipolar molecules in lattices [176–178] and trapped ions [179–182] constitutes promising platforms to realise the long-range case. With the help of a Rydberg quantum simulator with up to 100 atoms, we can study a long-range interacting, two-dimensional XY spin system exhibiting a continuous spin-rotational symmetry, with either ferromagnetic (FM) or antiferromagnetic (AFM) couplings. The exploration of the many-body physics of XY interactions on this system complements recent works utilising the Rydberg-blockade mechanism to realise Ising-type interactions exhibiting discrete spin rotation symmetry [74, 183–185].

5.1.1 - Characterising continuous symmetry breaking

Considering a collection of spins 1/2 $|\uparrow\rangle - |\downarrow\rangle$ arranged in a square lattice as depicted in Fig. 5.1(a), the dipolar XY model reads

$$H_{\mathrm{XY}} = -J/2 \sum_{i<j} \frac{a^3}{r_{ij}^3} (\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y) \tag{5.1}$$

where σ_i^α are Pauli matrices, J/h the dipolar interaction strength, a the lattice spacing and r_{ij} the distance between spins i and j. This Hamiltonian possesses the following continuous symmetry:

$$U_z(\theta) H_{\mathrm{XY}} U_z(-\theta) = H_{\mathrm{XY}} \tag{5.2}$$

with $U_z(\theta) = \exp(-i\theta M^z/2)$. This operator is generated by the total magnetisation, $M^z = \sum_i \sigma_i^z$, thus a conserved quantity of H_{XY}, and represents the Lie group $U(1)(\cong SO(2))$. $U(1)$ corresponds to the circle group, consisting of all complex numbers with absolute value 1, under multiplication.

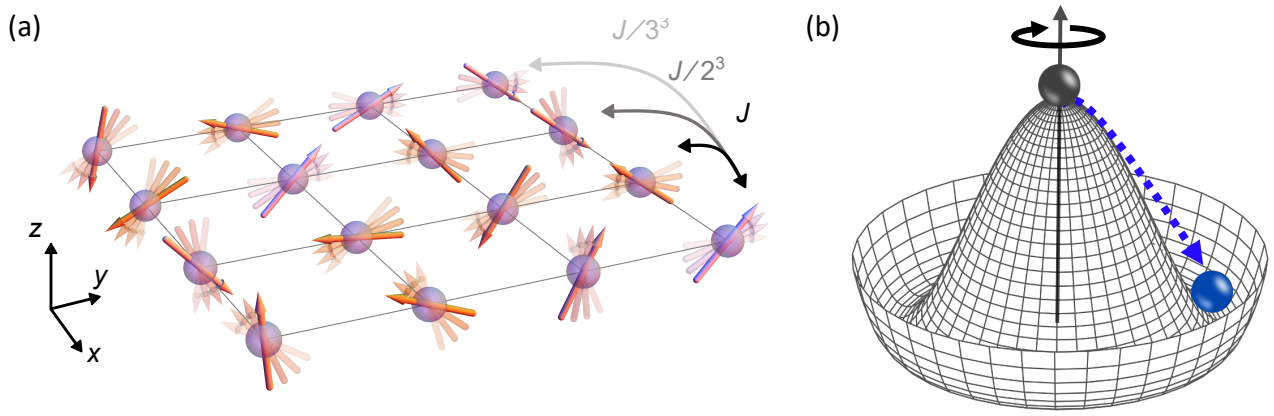


Figure 5.1: **Symmetry breaking in spins lattice.** (a) Schematic depicting the long-range dipolar XY model. An effective spin is encoded in a pair of Rydberg states which exhibit dipolar flip-flop interactions. (b) A Goldstone "sombrero" potential exhibits a $U(1)$ rotational symmetry. The system comprising the potential and a ball located at its top (black ball) remains symmetric. However, a lower energy state where the ball has rolled down (blue) spontaneously breaks the symmetry as the system is no longer invariant when rotating around the centre axis.

Additionally, H_{XY} is invariant under the $\mathbb{Z}_2$ Ising symmetry, $\alpha_2 : (\sigma^x, \sigma^y, \sigma^z) \to (\sigma^x, -\sigma^y, -\sigma^z)$, as well as any spatial symmetries of the lattice, such as translation or rotation. This model is also time-reversal-symmetric, as represented by the anti-unitary operator $\mathcal{T} = \mathcal{C}$, where $\mathcal{C}$ applies complex conjugation. Here $\mathcal{T}$ differs from the usual $SU(2)$ time-reversal symmetry, which applies the unitary spin rotation $U_y(\pi) = \exp(-i\pi M^y/2)$ in addition to $\mathcal{C}$. Our atypical choice of $\mathcal{T} = \mathcal{C}$ allows it to remain a symmetry in the presence of a on-site perturbation such as H_{Z} given in Eq. 5.6. In a finite closed quantum system, all eigenstates of H_{XY} can be chosen to be simultaneous eigenstates of all of these symmetry operators. In particular, they are eigenstates of the total magnetisation M^z and so can be collected into magnetisation sectors, conventionally labelled by $S^z = M^z/2$. As a consequence, all M^z-non-conserving operators such as σ_i^x and σ_i^y have identically vanishing expectation values, $\langle \sigma_i^x \rangle = \langle \sigma_i^y \rangle = 0$, in any energy eigenstate or in any superposition of eigenstates within the same magnetisation sector.

The many-body ground state of H_{XY}, in either the FM ($J > 0$) or AFM ($J < 0$) case, exhibits *off-diagonal long-range order* (LRO), which is indicative of a continuous symmetry breaking phase [186]. Put simply, the system spontaneously favours a configuration which no longer respects the overall symmetry as illustrated by a simpler example in Fig. 5.1(b). This LRO property can be quantified by the behaviour at long distance of the *connected* correlator,

$$C_{ij}^x = \langle \sigma_i^x \sigma_j^x \rangle - \langle \sigma_i^x \rangle \langle \sigma_j^x \rangle \tag{5.3}$$

In the special case of M^z eigenstates with $\langle \sigma^x \rangle = 0$, it becomes $C_{ij}^x = \langle \sigma_i^x \sigma_j^x \rangle$. If $|C_{ij}^x|$ approaches a constant $C_\infty^x > 0$ for distantly separated spins i, j, then the corresponding state is said to possess LRO. Such long-distance plateau of correlations is the defining feature of continuous symmetry breaking in

finite quantum systems. An equally good order parameter to use as a smoking gun for $U(1)$ symmetry breaking is the in-plane magnetisation squared

$$m^2_{\mathrm{FM/AFM}} = \frac{1}{N^2}\sum_{i,j}(\pm 1)^{i+j} C^x_{ij} \tag{5.4}$$

In the thermodynamic limit $N \to \infty$, any state with a correlation plateau $C^x_\infty \neq 0$ will also have a finite magnetisation $m^2_{\mathrm{FM/AFM}}$, and vice versa [187].

When continuous symmetry breaking occurs in the thermodynamic limit, then at finite size the lowest energy state in each S^z sector will be approximately,

$$\left|\Gamma_s^{\mathrm{FM/AFM}}\right\rangle = \frac{1}{\mathcal{N}_s}\int_0^{2\pi}\frac{\mathrm{d}\theta}{2\pi} e^{is\theta}\left|\theta^{\mathrm{FM/AFM}}\right\rangle \tag{5.5}$$

where $\left|\theta^{\mathrm{FM/AFM}}\right\rangle$ is the classical, symmetry-breaking product state where each spin points at angle θ or $-\theta$ in the xy-plane, s is an integer specifying the S^z sector, and $\mathcal{N}_s$ is a normalisation factor. Known either as the Anderson tower or Dicke states, $|\Gamma_s\rangle$ are angular momentum eigenstates of an emergent rigid rotor degree of freedom describing the collective orientation of all the spins in the system [187–190]. The true ground states in each S^z sector are also dressed by quantum spin wave fluctuations, which weaken the magnetic order [188]. For the ideal case of a uniform superposition over fully spin-polarised states $\left|\theta^{\mathrm{FM/AFM}}\right\rangle$, the correlations in $|\Gamma_0\rangle$ lead to $C^x_\infty = m^2 = 0.5$, plus $1/N$ corrections. The effective in-plane magnetisation of a $U(1)$-symmetric state should thus be identified as $m_{\mathrm{eff}} \equiv \sqrt{2C^x_\infty}$. That is, if one were to add a small symmetry-breaking field, then the corresponding non-symmetric ground state would have an average magnetisation $\langle\sigma^x\rangle = m_{\mathrm{eff}}$.

For the dipolar XY FM, theory predicts that the continuous symmetry breaking order persists in the presence of thermal fluctuations [164, 165] in apparent violation of the Mermin-Wagner theorem [191]. On the contrary, dipolar interactions are insufficient to stabilise finite temperature long-range order in the antiferromagnet [162]. Rather, one expects power-law decaying, algebraic long-range order due to Berezinskii-Kosterlitz-Thouless physics [192–196]. It is therefore of interest to prepare those ordered states and probe their phase diagrams with respect to thermal and quantum fluctuations.

5.1.2 - Preparing XY ground state through adiabatic protocols

The ground state of H_{XY} can be efficiently prepared by creating a connection to the spectrum of states of a simpler Hamiltonian and smoothly transitioning the spins from one state to another through adiabatic evolution. We utilise the Hamiltonian

$$H_Z(t) = \hbar\delta(t)\sum_{i\in B}(1+\sigma_i^z)/2 \tag{5.6}$$

where only a checkerboard sublattice B is addressed, applying a time-dependent staggered field $\delta(t)$ on half the spins as shown in Fig. 5.2(a). In the energy sector of null total magnetisation, i.e. $M^z = 0$, the classical Néel spin configuration is either the lowest energy state of H_Z ($\delta > 0$) or the highest energy state ($\delta < 0$). This staggered arrangement $|\psi_{N\acute{e}el}\rangle$ of $|\downarrow\rangle \in B$ and $|\uparrow\rangle \in A$ remains a good approximation of the ground state/highest energy state of the total Hamiltonian $H(t) = H_Z(t) + H_{\mathrm{XY}}$ when $\hbar|\delta| \gg J$. Adiabatically ramping down $|\delta(t)|$ towards 0 enables to connect $H(0) = H_Z$ where

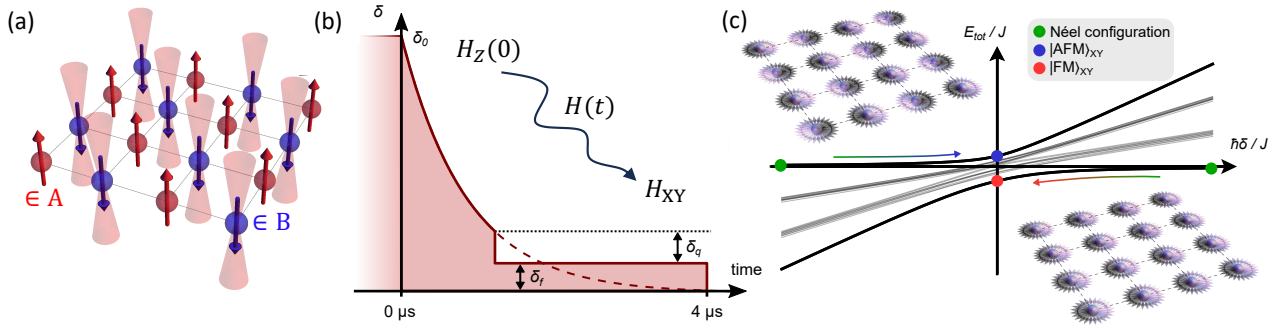


Figure 5.2: **Adiabatic driving towards ground state.** (a) A spatially dependent light-shift is used to prepare the system in a Néel spin configuration. (b) The amplitude δ of the light-shift is decreased as a function of time from δ_0 to a final value, δ_f, coupling H_{XY} to H_Z through a path $H(t)$. To study the robustness of the magnetic order with respect to an excess energy, we also introduce a diabatic quench of magnitude δ_q. This protocol enables to connect H_Z at $t = 0$ starting from the Néel state (green dot) to H_{XY}. (c) Energy spectrum of H as a function of δ, for $N = 2 \times 3$ spins. When starting in its ground state for $\hbar\delta/J \gg 1$, the system is adiabatically ramped to the ferromagnetic XY state (red dot), pictured by the coloured fluctuating arrows correlated in directions. When starting in the highest excited state for $\hbar\delta/J \ll -1$, the system is adiabatically ramped to the antiferromagnetic XY state (blue dot), portrayed by the anticorrelated fluctuating arrows.

the staggered field prevails to $H(t_f) = H_{\mathrm{XY}}$ where the dipolar interactions prevail. We use in the following an exponential ramp profile, $\delta(t) \approx \delta_0 e^{-t/\tau}$ as displayed in Fig. 5.2(b). Choosing such a protocol enables to start in the sector of null magnetisation and works with rather large gaps during the dynamics, preventing potential diabatic errors. Compared to just starting from $|\downarrow \cdots \downarrow\rangle$ and adding excitations with MW, this method proves more reliable.

The adiabatic ramp can also be quenched by an amount δ_q to study diabatic effects on the dynamics. For $\delta(t) > 0$, the ramp connects $|\psi_{N\acute{e}el}\rangle$ to low-temperature ferromagnetic states of H_{XY}, as shown in Fig. 5.2(b). Meanwhile, for $\delta(t) < 0$, the adiabatic ramp prepares *negative* temperature states of H_{XY} or equivalently, low-temperature antiferromagnetic states of $-H_{\mathrm{XY}}$ [197]. In the thermodynamic limit of both cases, a quantum phase transition is expected to occur at some critical $\delta(t) = \delta_c^{\mathrm{FM/AFM}}$, between the Néel configuration and the XY order (Methods of [37]).

5.1.3 - Description of the experimental implementation

The experimental setup consists of a two-dimensional square lattice of ^{87}Rb atoms trapped in an optical tweezer array as shown in Fig. 5.3(a). We encode an effective spin 1/2 in a pair of opposite-parity Rydberg states, $|\uparrow\rangle = |60S\rangle$ and $|\downarrow\rangle = |60P\rangle$. We manipulate them using resonant MW at 16.7 GHz. Resonant dipole-dipole interactions between the spins naturally realise the dipolar XY model as explained in Sec. 1.3.2. The lattice has fixed spacing $a = 12.5$ μm leading to $J/h = 0.77$ MHz. The quantisation axis is defined by an external magnetic field perpendicular to the lattice plane, which ensures that the dipolar interactions are isotropic.

The addressing laser pattern generated the staggered field uses a 1013-nm laser beam detuned from the transition between the intermediate state $|e\rangle$ and $|\uparrow\rangle$ as shown in Fig. 5.3(b). The sign of

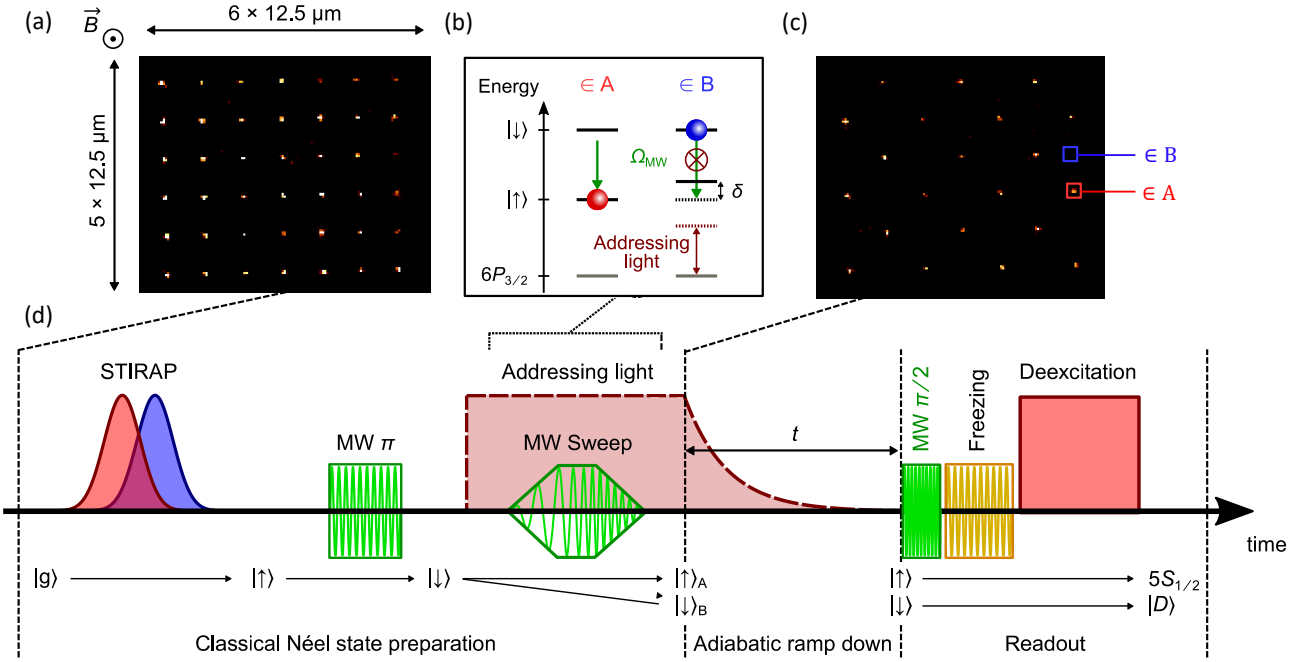


Figure 5.3: **Experimental procedure and pulse sequence.** (a) Fluorescence image of the atoms in a fully assembled 6×7 array. (b) Scheme for the preparation of the initial staggered state using the addressing light. (c) Detected staggered state, corresponding to the situation for which all the atoms in sublattice A are in $|\uparrow\rangle$, and all the atoms in sublattice B are in $|\downarrow\rangle$. (d) Experimental sequence described in Sec. 5.1.3.

the detuning sets the one of the light-shift δ: the frequency of the addressing laser is tuned below or above the resonance by ~ 250 MHz.

A dedicated SLM produces the desired pattern of addressing beams. Each beam is focused on a $1/e^2$ radius of about 1.5 μm, for a typical power of 60 mW. The light-shift for each addressed atom is measured by microwave spectroscopy on the $|\uparrow\rangle - |\downarrow\rangle$ transition and the values are dictated by available laser power. On average, for a 42(100)-atom array with 21(50) addressed atoms, $|\delta_0|/2\pi = 15(9)$ MHz, confirming the ability to start in the $\hbar|\delta| \gg J$ regime. The dispersion of δ across the addressing beams is 2.4%.

The experimental sequence is shown in Fig. 5.3(d). After assembling the array [198] we use Raman sideband cooling along the radial directions of the tweezers, and reach a temperature of 10 μK. We then optically pump the atoms in $|g\rangle$ before adiabatically ramping down the tweezer depth by a factor ~ 40. Following this, we switch off the tweezers, and excite the atoms to $|\uparrow\rangle$ using a two-photon stimulated Raman adiabatic passage (STIRAP) with 421-nm and 1013-nm lasers (~ 2 μs duration). To generate the classical Néel configuration along z, we first transfer all the atoms from $|\uparrow\rangle$ to $|\downarrow\rangle$ using a 54 ns microwave π-pulse. Subsequently, the addressing beams are applied to the atoms in sublattice B. We then transfer the atoms A from $|\downarrow\rangle$ back to $|\uparrow\rangle$ by an adiabatic microwave sweep while the atoms B remain in $|\downarrow\rangle$. In this procedure, exciting first the atoms in $|\downarrow\rangle$ has the advantage of minimising the depumping of the $|\uparrow\rangle$ atoms by the addressing light. An example of perfect Néel

configuration obtained at the end of the preparation is shown in Fig. 5.3(c).

At the end of the sequence, we read out the state of each atom in the natural z-basis. To do so, we deexcite the atoms from $|\uparrow\rangle$ to the $5S_{1/2}$ manifold where they are recaptured in the tweezers and imaged. Thus, the $|\uparrow\rangle$ (resp. $|\downarrow\rangle$) state is mapped to the presence (resp. absence) of the corresponding atom. The subsequent deexcitation is performed by applying a 2.5 μs light pulse resonant with the transition between $|\uparrow\rangle$ and the short-lived intermediate state $6P_{3/2}$ from which the atoms decay back to $5S_{1/2}$.

Additionally, when we want to measure the spins along x we rotate them by applying a 27 ns microwave $\pi/2$-pulse on the $|\uparrow\rangle - |\downarrow\rangle$ transition prior to the detection. However, this procedure is efficient only for light-shifts $|\delta(t)|$ much smaller than the microwave Rabi frequency, i.e. for times larger than ~ 0.5 μs during an adiabatic preparation.

The experimental sequence is repeated typically over 1000 defect-free assembled arrays. This allows us to calculate the magnetisation and the spin correlations by averaging over these realisations.

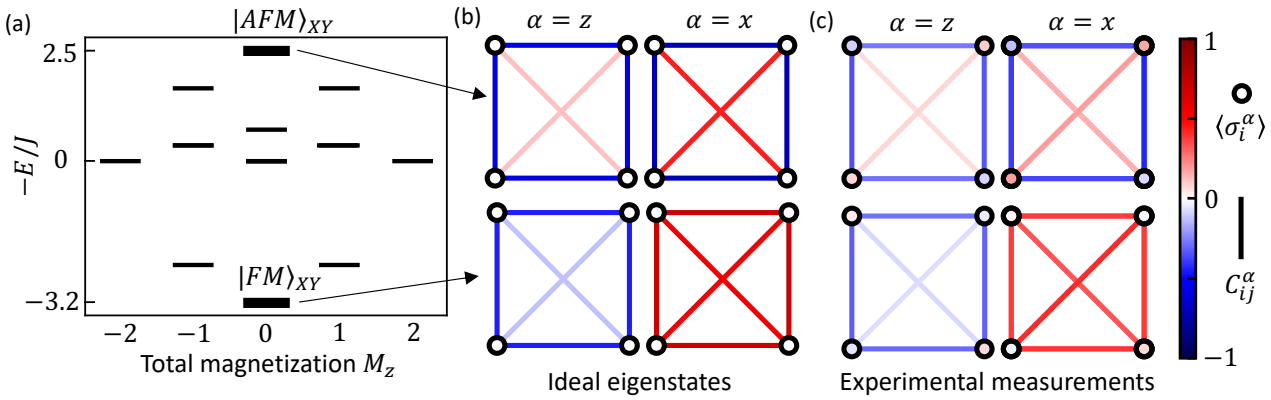


Figure 5.4: **Theory and experiment on a plaquette.** (a) Energy spectrum of H_{XY} on a 2×2 plaquette arranged by magnetisation sectors. Magnetisation and correlations of (b) ideal eigenstates and (c) experimentally prepared states (at $t = 2$ μs). The colour of each site (bound) represents the value of the magnetisation (correlation) at that site (bound).

5.1.4 - Benchmark on a plaquette and noise model

As a first experiment, we apply the procedure described above on a square plaquette of 2×2 atoms with $\delta_0/2\pi = 15$ MHz and $\tau = 0.3$ μs. The spectrum of such a system is displayed in Fig. 5.4(a). After preparing the system in the zero magnetisation classical state $|\psi_{N\acute{e}el}\rangle$, we start ramping down the light-shift applied on the atoms in sublattice B. After a time $t = 2\,\mu$s (end of the ramp $\delta(t) \approx 0$) we measure the magnetisation $\langle\sigma_i^\alpha\rangle$ and the connected correlations $C_{ij}^\alpha = \langle\sigma_i^\alpha\sigma_j^\alpha\rangle - \langle\sigma_i^\alpha\rangle\langle\sigma_j^\alpha\rangle$ on the plaquette. Figure 5.4(b-c) shows a comparison between the ideal and experimentally produced states. Qualitatively, the observations are in agreement with the ideal expectations for the FM or AFM state, in particular for the sign of the correlations between nearest neighbour (NN) and next nearest-neighbours (NNN): $C^x_{\mathrm{NN}} > 0$ and $C^x_{\mathrm{NNN}} > 0$ for the FM case, and $C^x_{\mathrm{NN}} < 0$ and $C^x_{\mathrm{NNN}} > 0$ for the

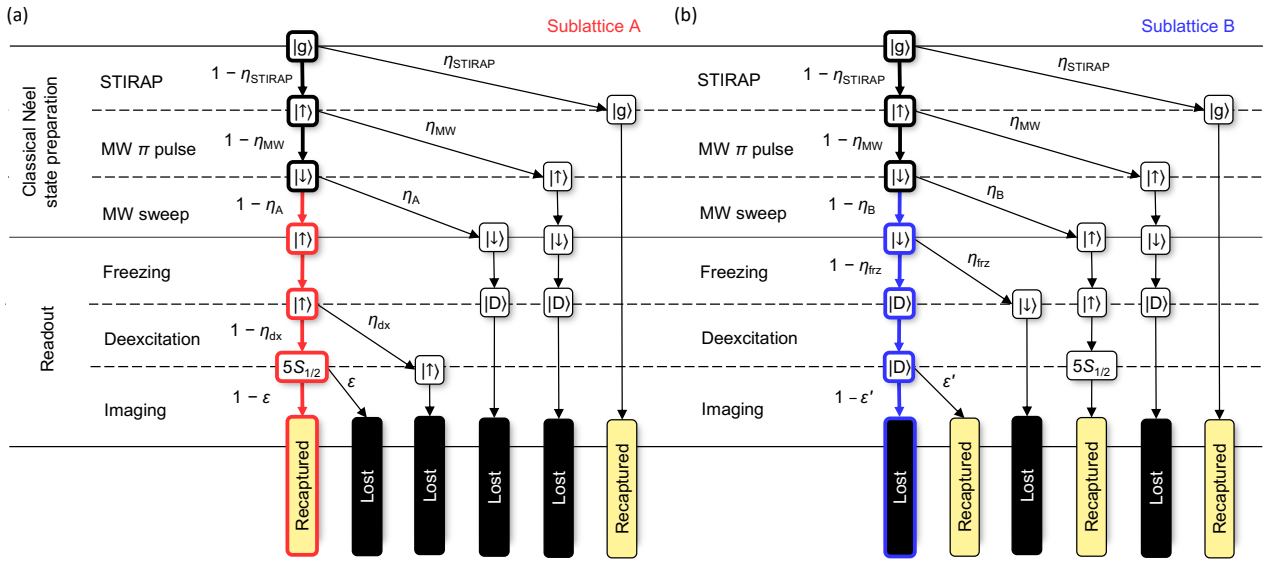


Figure 5.5: **Simplified error tree associated to the preparation and measurement of the initial Néel state,** for (a) the atoms in sublattice A (non-addressed), and (b) in sublattice B (addressed). For simplicity, the events with a probability of order 2 or higher in the η_i, ε, ε' are disregarded.

AFM case. However, residual non-zero magnetisation on the sites and loss of contrast on the bounds are still visible for both x and z. The protocol is therefore not perfectly preparing the desired states.

Numerical simulation of dynamics

For this small system, we can benchmark quantitatively the experimental results to numerical simulations of the dynamics of this 2×2 system.

We first consider the various experimental imperfections occurring at the preparation and measurement stages of the sequence, the so-called SPAM errors. In order to estimate the latter, we break down the sequence depicted in Fig. 5.3(d) into a series of steps i, each having a small but finite failure probability η_i. In the following, we keep only the contributions of imperfections to first order in the η_i's. As an example, we show in Fig. 5.5 the discretised sequence corresponding to the preparation and measurement of the classical Néel configuration. Table 5.1 summarises the corresponding values of the probabilities η_i, that are either inferred from a series of dedicated experiments, or estimated from numerical simulations. For atoms in sublattice A (non-addressed), the error tree leads to the probability to recapture the atoms at the end of the sequence, which reads (to first order):

$$P_z^{\mathrm{A}} \approx 1 - \eta_{\mathrm{MW}} - \eta_{\mathrm{A}} - \eta_{\mathrm{dx}} - \varepsilon_A \tag{5.7}$$

Similarly, the calculation for sublattice B (addressed atoms) yields:

$$P_z^{\mathrm{B}} \approx \eta_{\mathrm{STIRAP}} + \eta_{\mathrm{B}} + \varepsilon' \tag{5.8}$$

The error tree allows us to infer the probability of successful initial preparation per spin: we find $1 - \eta_{\mathrm{STIRAP}} - \eta_{\mathrm{MW}} - \eta_{\mathrm{A/B}} = 0.89/0.88$ for the atoms in sublattice A/B. These values are very similar

Stage	Symbol	Value	Main physical origin
$\|\psi_{Néel}\rangle$ prep.	η_{STIRAP}	5%	Imperfect optical pumping, Laser phase noise, Spontaneous emission from $6P_{3/2}$ [199]
	η_{MW}	2%	Effect of H_{XY} during pulse
	$\eta_{\mathrm{A}}, \eta_{\mathrm{B}}$	4%, 5%	Effect of H_{XY} and finite value of $\|\delta_0\|$
Readout	η_{frz}	$< 1\%$	Effect of H_{XY} during pulse
	η_{dx}	3%	Mechanical effect of deexcitation beam
	$\epsilon_{A/B}$	1%/3%	Background gas collisions [199]/Addressing kick
	ϵ'	5%	Rydberg state radiative lifetime [199]

Table 5.1: **Summary of the experimental errors defined in Fig. 5.5, together with their main physical origin.**

to the experimentally obtained ones when correcting for the detection errors, i.e. 0.87/0.92 for A/B, indicating that this experiment is dominated by preparation errors.

In addition to carefully estimating the SPAMs, our model incorporates van der Waals interactions (see Box 9) and considers the thermal motion of atoms, resulting in time-dependent interactions. Notably, the addressing beam imparts an additional velocity kick on atoms in region B due to the repulsive ponderomotive force. Consequently, addressed atoms moves faster, altering further the interactions and increasing susceptibility to losses, necessitating the definition of $\varepsilon_B > \varepsilon_A$.

Simulation results are obtained through Monte Carlo sampling. In each run, an imperfect initial state is generated using the inferred preparation efficiencies per site, along with sampling initial positions and velocities for each atom. The simulation tracks dynamics from $t = 0$ to $t = 8.2$ μs and detection errors are also incorporated. For both the FM and the AFM, we measure spin observables, i.e. magnetisation averaged over sublattices A and B and correlations averaged over pairs of NN/NNN as a function of the ramping time t as shown in Fig. 5.6. In both cases, we observe a rapid growth of the correlations along x as we sweep down the lightshifts, showing the transition from a classical uncorrelated phase to a quantum correlated phase. The final signs of the experimentally measured correlations are compatible with the ones given by the ideal eigenstates, and their values are lower by approximately a factor 2. However, these values are well reproduced when including the errors in the simulation, indicating that we understand the main residual limitations of the setup.

The evolution of the magnetisation along x reveals a residual oscillation, which indicates that the prepared state is not stationary, and thus not an eigenstate of H_{XY}. Considering the spectrum shown in Fig. 5.4(a), the oscillation frequency corresponds to the energy difference between the lowest (resp. highest) energy states and the ones directly above (resp. below) for the AFM (resp. FM) case: this indicates that the origin of the oscillations is the imperfect preparation of the ground state of H_Z at $t = 0$ leading to non zero population in higher energy states. Further discrepancies can be explained by increased depumping on the addressed atoms in the FM case resulting in constant shifts in magnetisations and by the Rydberg finite lifetime resulting in a slight decay of the correlations along z at long times. Those two effects are further studied in the larger instances.

In addition to this benchmark, we can start to assess the role of the next nearest neighbour coupling

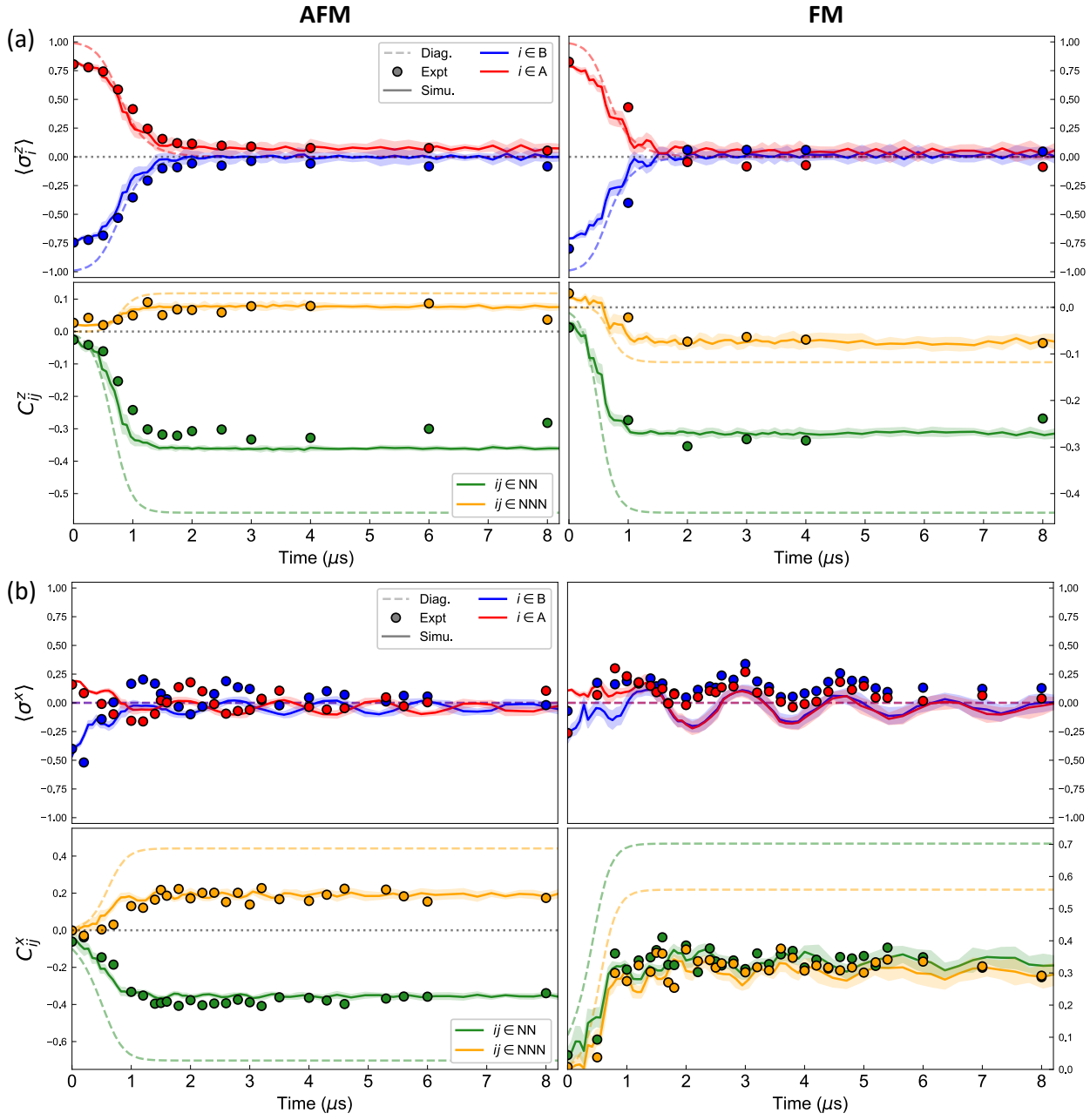


Figure 5.6: **Dynamics of the adiabatic preparation for the 4-atom system**. Experimental measurements (dots) are compared to perfect diagonalisation of instantaneous Hamiltonian $H(t)$ (dashed) and numerical simulations including experimental imperfections (solid). For the AFM (left column) and for the FM (right column), the mean magnetisation over sublattices A (red) and B (blue) and the mean spin-spin correlations overall nearest-neighbour (green) pairs of atoms and overall next-nearest-neighbour (orange) pairs are measured both along (a) z and (b) x.

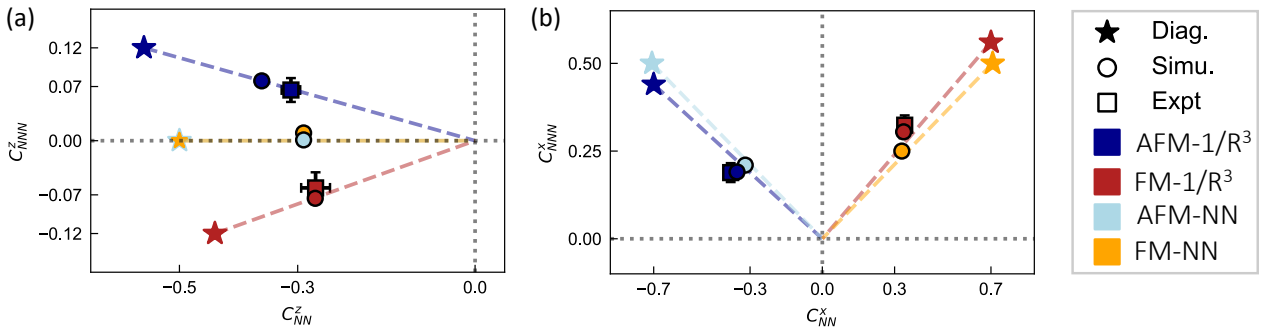


Figure 5.7: **Summary of connected correlations in 4-atom plaquette case with and without diagonal coupling.** Panel (a) (resp. b) exhibits C^x (C^z) for exact calculation (dot), the noisy simulation described in the text (star) and experimental realisations (cross). For both exact and noisy cases, NN (x-axis) and NNN (y-axis) correlations have been computed with the $1/R^3$ dipolar interaction for AFM (blue) and FM (red) case. In addition, the same quantities have also been computed when removing the diagonal coupling, in AFM (lightblue) and FM (orange) case. For both noisy simulation and experimental data, the values are the average between 2 μs and 8 μs after the beginning of the ramping down. Between those times, the lightshift is negligible and the system is believed to remain in the AFM/FM XY state.

in the evolution and final values of the correlations. This can be achieved by removing the coupling along the diagonal of the plaquette in the simulation. The results, shown in Fig. 5.7, highlight the significance of considering the $1/r^3$ dependence of the dipolar interaction. In the NN case, besides the sign of C^x_{NN}, there is no difference between FM and AFM cases.

5.1.5 - Probing the dynamics on larger systems

To investigate larger instances of the XY ferromagnet, we apply the protocol on a 6×7 lattice with $\delta_0 = 2\pi \times 15$ MHz and $\tau = 0.3$ μs. As depicted in Fig. 5.8(a), for both sublattices, the on-site z-magnetisation, obtained by averaging over many realisations of the experiment, decreases toward zero, with a residual late-time offset arising from experimental imperfections. This is consistent with the XY ferromagnet, which orders in the equatorial plane, but by itself, is insufficient to diagnose the phase. Indeed, quenching the staggered light-shifts (in less than 100 ns) leads to a near infinite temperature state, which also exhibits a magnetisation that rapidly relaxes to zero (lighter curves, Fig. 5.8(a)).

The key characteristic of the XY ferromagnet is only revealed upon measuring the correlation function, C^x_{ij} introduced in Eq. 5.3. For the quenched state, the correlation functions remain near zero for all times, consistent with high-temperature behaviour (lighter curves, Fig. 5.8(b)). The dynamics of the adiabatic protocol are markedly distinct – both nearest-neighbour and next-nearest-neighbour correlations grow to a stable non-zero value at late times, indicative of order [186]. By switching the sign of δ_0, we also investigate the XY antiferromagnet. Both the z-magnetisation shown in Fig. 5.9(a) and the correlation functions shown in Fig. 5.9(b) exhibit qualitatively similar dynamics as the ferromagnetic case. One notable difference is that $C^x < 0$ for nearest-neighbour correlations, indicating that neighbouring spins have anti-aligned.

We then first increase the system size to a 10×10 lattice and perform the analogous adiabatic preparation protocols. We find the same behaviour for all observables (insets, Fig. 5.8/5.9(a)), indicating that our results are robust to finite-size effects [200]. Second, to explore the adiabaticity of our protocol, we vary the time-constant of the exponential ramp. As shown in the insets of Fig. 5.8/5.9(b), the dynamics of the correlation function agree between $\tau = 0.15$ μs and $\tau = 0.3$ μs, indicating that potential diabatic errors appearing at larger sizes are not a limiting factor. Third, while the long-range tail of the dipolar interaction reinforces the XY FM order, it is weakly frustrating for the AFM [172]. As a consequence, the phase transition between the Néel configuration and the XY AFM is expected to occur at a smaller value of the staggered light-shift as compared to the XY FM, i.e. $|\delta_c^{\mathrm{AFM}}| < |\delta_c^{\mathrm{FM}}|$. This is indeed borne out by the data where we observe that the magnetisation decays to zero faster as a function of δ for the FM case than for the AFM.

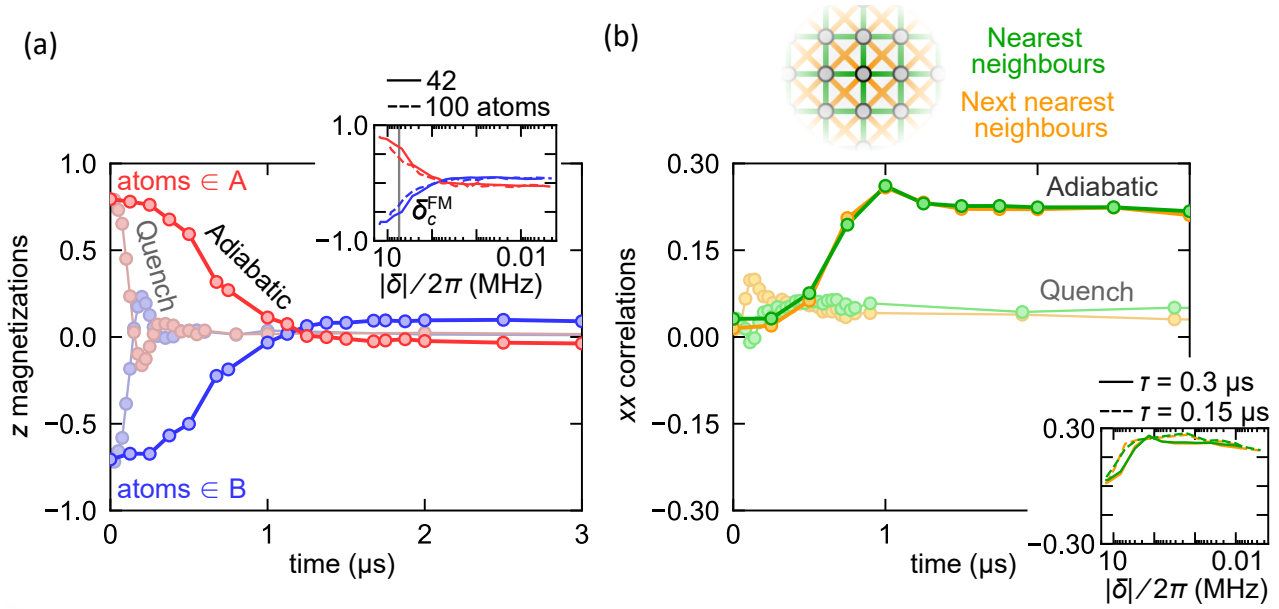


Figure 5.8: **Adiabatic preparation of dipolar XY ferromagnet.** (a) Sublattice-resolved magnetisation $\langle \sigma_i^z \rangle$ as the staggered field $\delta(t)$ is reduced. At $t = 0$, the state is prepared in a classical Néel state along the z-axis, as indicated by the opposing magnetisation of atoms in the A (red) and B (blue) sublattices. As the staggered field $\delta(t)$ is turned off, either adiabatically or via a sudden quench, the Néel magnetisation decays towards zero. (inset) Comparison of the z-magnetisations decay as a function of δ for a 6×7 versus a 10×10 lattice. The gray vertical line indicates the value δ_c^{FM} where the phase transition occurs, inferred from the theory (Methods of [37]).(b) The formation of a low-energy XY-ferromagnet is detected via the in-plane two-point correlation function, $C_{i,j}^x$. Data is shown for i, j averaged over either nearest or next-nearest pairs. The sudden quench produces additional energy which destroys the XY order and leads to correlations near zero. (inset) Nearest and next-nearest correlations for two different adiabatic ramp rates.

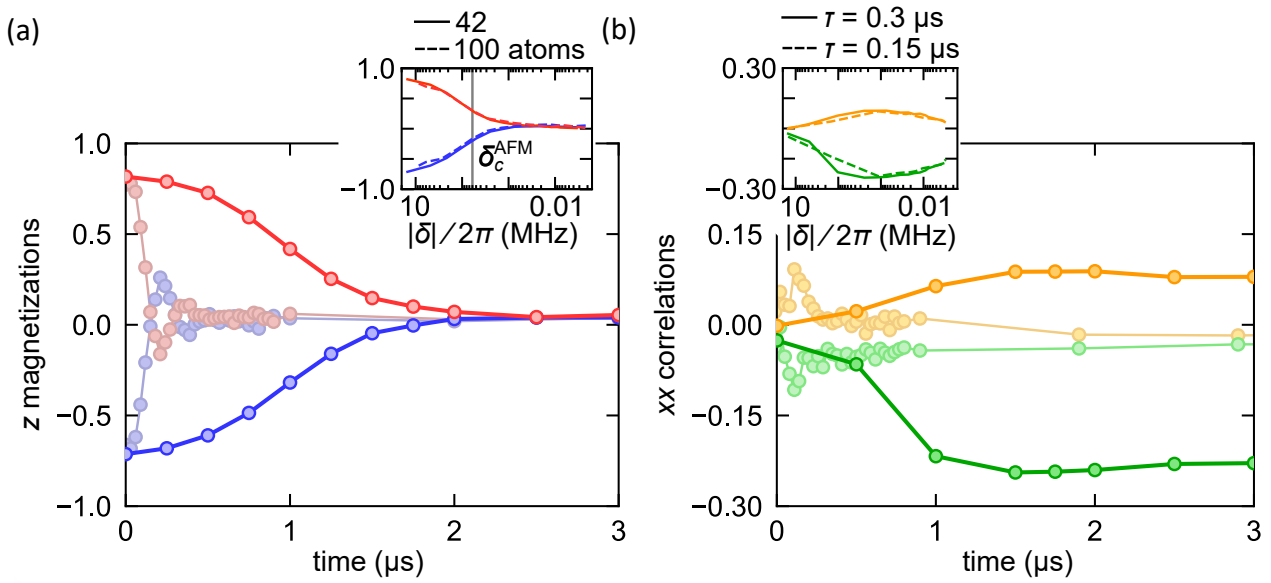


Figure 5.9: **Adiabatic preparation of dipolar XY antiferromagnet.** Similar results than for the antiferromagnetic case depicted in Fig. 5.8

5.1.6 - Measuring long range order

Our measurements of the local correlations suggest we have dynamically prepared low-temperature states of the XY FM and AFM – but are these states truly long-range ordered? To investigate this, we measure the long-distance spin-spin correlations of the 10×10 spins systems after adiabatic preparation. In Fig. 5.10(a) we show the correlations as a function of the displacement $\vec{d}$, averaging over initial positions: $C^x(\vec{d}) \equiv \langle C^x_{\vec{r},\vec{r}+\vec{d}} \rangle_{\vec{r}}$. The FM correlations are of constant sign and appear to plateau at long distances, indicative of long-range order, while the AFM correlations are staggered and exhibit a decay. For a more quantitative assessment, we plot $C^x(d)$, averaging over displacements of the same distance $d = |\vec{d}|$. In the XY AFM, correlations decay to zero at large distances, indicating the absence of long-range order. By contrast, the XY FM indeed exhibits a plateau, $C^x_\infty \sim 0.13$, which establishes it as a magnetically ordered state with an effective magnetisation density $m_{\text{eff}} \equiv \sqrt{2C^x_\infty} = 0.51$.

For additional insight, in Fig. 5.10(b), we plot the exact ground-state prediction obtained from DMRG calculations [201] using the MPS framework implemented in the TeNPy software library [202]. We can compute $C^x(d)$ up to 10×10 arrays for the FM case $H^{\text{FM}}_{\text{XY}}$ ($J > 0$), for the AFM case $H^{\text{AFM}}_{\text{XY}}$ ($J < 0$) and for the NN version H_{NN} where only NN terms are kept in Eq. 5.1. AS already conveyed by the 2×2 case, H_{NN} exhibits similar behaviour in the FM and AFM case. Overall, the ground state of the $H^{\text{FM}}_{\text{XY}}$ is clearly XY LRO, while for $H^{\text{AFM}}_{\text{XY}}$ and H_{NN}, the states exhibit stronger finite-size effects. Given that H_{NN} is rigorously known to be LRO in the thermodynamic limit, the similar behaviour observed for $H^{\text{AFM}}_{\text{XY}}$ is a strong indication that it is as well. While the qualitative structure of the measured $C^x(d)$ (e.g. alternating sign structure in the AFM case) is consistent with theory, the experimental correlations are weaker. A number of effects could contribute to this. For example, the finite fidelity of the initial Néel state introduces an entropy density (i.e. an effective finite temperature). This is especially destructive to the AFM, for which finite temperature long-range

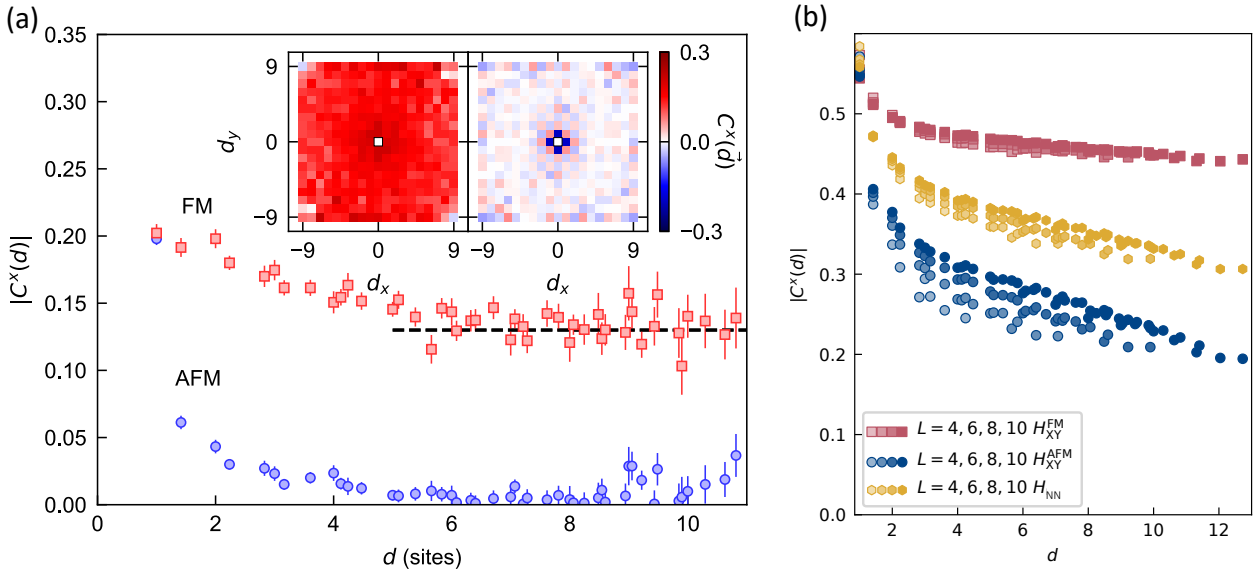


Figure 5.10: **Observing long-range XY order** (a) Correlations along x averaged over displacements of the same distance, $C^x(d)$ for the 10×10 lattice. The XY ferromagnet exhibits a plateau consistent with long-range order, while the XY antiferromagnet exhibits a decay to zero. (inset) Spatial correlations as a function of displacement, measured at time $t = 1$ μs. (b) Theoretical correlation profile $|C^x(d)|$ on $L \times L$ square clusters with open boundary conditions for $H_{\mathrm{XY}}^{\mathrm{FM/AFM}}$ (red/blue) and H_{NN}.

order is forbidden [162, 172], in agreement with our observation. Including other already discussed experimental imperfections, e.g. readout errors, leads to excellent agreement with the data for the 6×7 lattice (see Methods of [37]). However, we also observe that running the adiabatic preparation protocol to longer timescales leads to additional decoherence which adversely affects the ferromagnetic magnetisation plateau in a non-trivial fashion; in particular, correlations at the largest distances begin to decay before their shorter-distance counterparts as showcased by Fig. 5.11.

5.1.7 - Noise effects happening at large sizes

We discuss the influence of identified noises on the results obtained at large sizes. The initial magnetisations along z in Fig. 5.8/5.9(a) are a bit reduced due to preparation errors. We can check using the error tree in Fig. 5.5 that their measured values are in accordance with their expected ones, taking into account the finite efficiency of each step: $\sigma_z^{\mathrm{A/B}} = 2P_z^{\mathrm{A/B}} - 1 = 0.8/-0.70$. In addition, in the FM case, the z-magnetisations of sublattices A and B do not vanish at late times, but reach a constant finite value of a few percent. In contrast, this does not occur in the AFM case. We qualitatively explain this effect by the following observations. First, due to off-resonant scattering by the addressing beam, atoms in $|\uparrow\rangle$ are slowly depumped to the ground state $|g\rangle$; we have measured the effective lifetime of an addressed $|\uparrow\rangle$ atom to be ~ 4 μs, whether the light-shift is $2\pi \times 15$ or $-2\pi \times 15$ MHz (so that this alone, cannot explain the difference between the FM and AFM cases). However, during our adiabatic ramp down of light-shift $\delta(t)$, the addressed atoms are initially in $|\downarrow\rangle$ (and thus cannot be depumped). Depumping sets in only when the system enters the ordered phase,

where an addressed atom has a significant probability to be in $|\uparrow\rangle$. Since $\delta_c^{\mathrm{AFM}} < \delta_c^{\mathrm{FM}}$, the addressing beam intensity (and thus the depumping rate) is at this stage much smaller for the AFM case than for the FM case, and thus has a negligible effect in the former case.

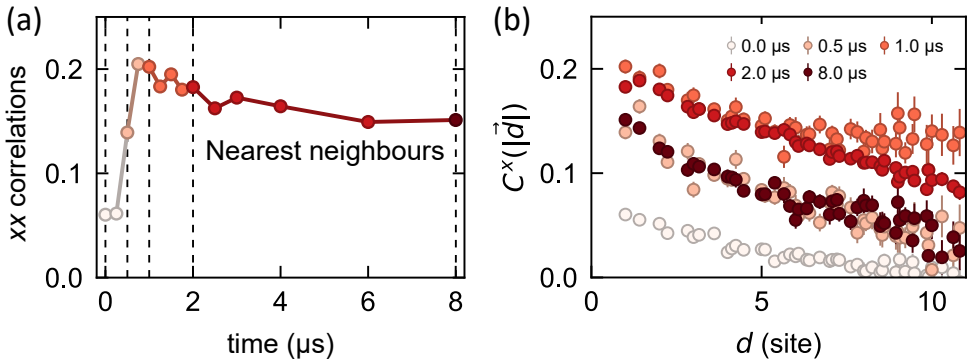


Figure 5.11: **Time dependence of the correlations along x in the FM case for a 10×10 lattice.** (a), Time evolution of the nearest-neighbour correlations along x (different colours correspond to different times). (b), Spatial correlations as a function of distance, measured at different times $t = \{0.0,\ 0.5,\ 1.0,\ 2.0,\ 8.0\}$ μs indicated by dashed lines in (a).

We also investigate the role of decoherence on the appearance of long-range order along x in the FM case, for the 10×10 array. Fig. 5.11(a) shows the time evolution of the nearest-neighbour correlations as we ramp down the light-shift, all the way up to 8 μs (in contrast with Fig. 5.8(b) where the evolution is shown only up to 3 μs, and for 42 atoms). Two timescales appear: first, correlations build up until $t \simeq 1$ μs as the FM state is adiabatically prepared ; then, they slowly decay and lose 25 % of their value in 7 μs. This decay is not expected, since the system should be ideally in steady state once it has reached the ferromagnetic phase. We conjecture that the experimental system is affected by decoherence arising from a combination of the residual atomic motion and spontaneous emission from the Rydberg states. To further analyse the evolution of the ferromagnetic order, we probe the full spatial structure of the correlations at different times. Figure 5.11(b) summarises the results. We observe that for a given distance d all the correlations feature a similar time evolution: a fast increase followed by a slow decay, with a turning point around 1 μs. For this particular point, the data reveal a plateau for distances of more than 6 sites – the signature of the long range order – that disappears for $t \gtrsim 2$ μs. This suggests that despite the decoherence present in the system, we are able to observe the long range ordering expected from the dipolar interactions over a substantial time window.

As a final characterisation of the prepared states, we investigate whether each realisation of the experiment produces a *classical* magnet pointing in a random direction θ in the xy-plane or a genuinely *quantum* many-body state described by Eq. 5.5. To do so, we analyse the statistical distribution of M^z, which is conserved during the adiabatic ramp. For a classical FM or AFM, each spin, aligned or anti-aligned along θ, is an equal superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$, so that M^z follows a binomial distribution. By contrast, the ground state of H_{XY} is an eigenstate of M^z, and its variance should be zero.

Figure 5.12(a,b) shows experimental histograms of the z-magnetisationat $t = 2\,\mu$s for the FM and AFM. Figure 5.12(c) presents the variance for various times t. We observe that the states exhibit variances smaller than those expected from a binomial distribution, suggesting that classical magnets are not being prepared. Notably, the non-zero variances measured can be entirely accounted for by the SPAM errors described in Fig. 5.5. Furthermore, we have verified the rotation invariance of the state about the z-axis by measuring magnetisation along the y-axis, yielding consistent results with measurements along the x-axis. Altogether, our measurements suggest a state which is a coherent quantum superposition over a continuous family of classical configurations. For such a state, the defining signature of order is a long-distance plateau in the correlation function $C^x(d)$ – as we observed in the XY FM [157].

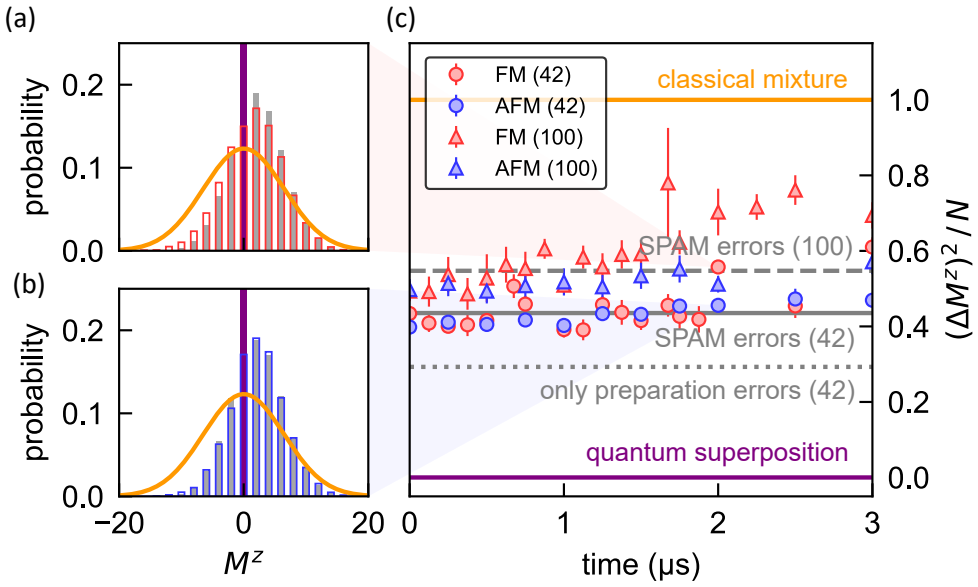


Figure 5.12: **Analysis of the z-magnetisation during the adiabatic ramp.** Experimental histograms of the z-magnetisation M^z ($N = 42$) for (a), the FM and (b), AFM case together with the ideal case (purple), and the expected distribution including state preparation and measurement errors (grey bars). The orange line is the binomial distribution corresponding to a classical magnet (see text). (c) Normalised variance $(\Delta M^z)^2/N$ as a function of time during the ramp, for the experiment (circles for $N = 42$, triangles for $N = 100$), the classical magnet (orange line) and a perfect XY-magnet (purple line). Grey continuous and dashed lines: ideal case including state preparation and measurement errors. Dotted line: ideal case ($N = 42$) including only state preparation errors.

5.1.8 - Exploration of the phase diagrams

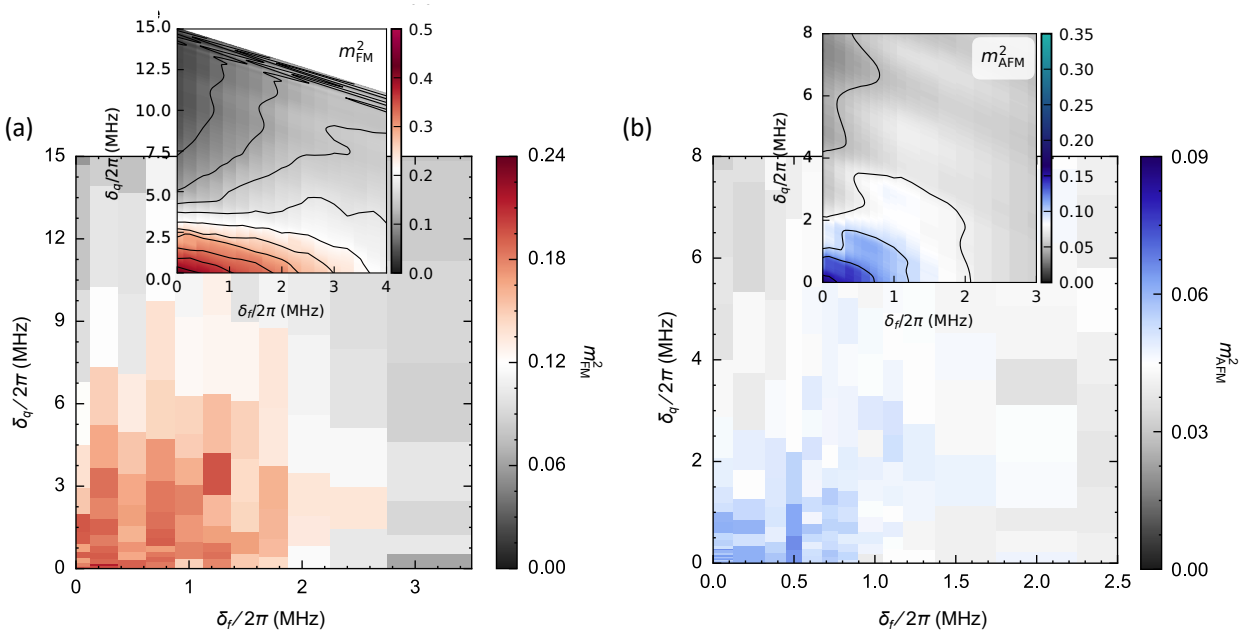


Figure 5.13: **Experimentally measured and theoretical XY phase diagrams.** (a) Ferromagnetic phase diagram depicting the magnetisation squared as a function of the final staggered field strength, δ_f and the diabatic quench magnitude, δ_q. Symmetry breaking is expected in a lobe about $(\delta_f = 0, \delta_q = 0)$ and is destroyed by either quantum (δ_f) or thermal (δ_q) fluctuations. On a 6×7 system, a crossover between ordered and disordered behaviour is observed. (b) Analogous phase diagram for the antiferromagnet. Note that at finite temperature, only algebraic long-range order is expected. (insets) Theoretical phase diagrams obtained with Minimally Entangled Typical Thermal States (METTS) algorithm [203] method.

As mentioned earlier, the long-range order observed in the FM case should persist at finite temperature. We therefore investigate the stability of the prepared magnetic orders as a function of an effective temperature. To do so, we insert a partial quench of amplitude δ_q into the ramp, followed by an equilibration time of at least 1 μs at a final value δ_f of the staggered field (Fig. 5.2(b)): the variable quench introduces excess energy into the system, and we observe a relaxation of the magnetisation and correlations during the equilibration time. We will use the amplitude of the quench, δ_q, as a proxy for the final effective temperature.

After each $\{\delta_f, \delta_q\}$ ramp, we measure the in-plane magnetisation squared $m^2_{\mathrm{FM/AFM}}$ defined in Eq. 5.4 and construct the phase diagram shown in Figs. 5.13. Starting with the ferromagnet (a), for small values of δ_f and δ_q (corresponding to low effective temperatures), the magnetisation per site is of $\mathcal{O}(1)$, consistent with the ordered phase. As either δ_f or δ_q increases, the magnetisation density decreases toward zero indicating melting into a disordered phase. This is consistent with theoretical expectations, where δ_q drives the transition via thermal fluctuations [164], while δ_f tunes across the quantum phase transition. We perform the same analysis for the antiferromagnet (b). Compared to the XY ferromagnet, we find that a much smaller region of the $\{\delta_f, \delta_q\}$ phase space exhibits

significant AFM correlations, consistent with the frustration induced by the long-range interactions which destabilises the phase. Theoretical expectations for those phase diagrams as well as precise thermometry of the system with correspondence between δ_q and an effective temperature are more deeply explained in Methods of [37] and in [152].

5.1.9 - Outlook

It would be interesting to investigate the nature of the phase transition between the disordered and XY-ordered phases; this will require overcoming a number of technical challenges including scaling to larger system sizes. Second, the ability to directly prepare low-temperature states in different M^z magnetisation sectors due to state preparation errors suggests the possibility of directly observing the so-called Anderson tower of states, which underlies continuous symmetry breaking in finite quantum systems [187–190]; the structure of these states has led to recent predictions for scalable spin squeezing by quenching in the ferromagnetic XY phase [204, 205]. Finally, combining optical tweezer geometries which exhibit frustration (i.e. triangular or Kagome lattices) with antiferromagnetic interactions leads to a rich landscape for exploring frustrated magnetism as showcased in the next section for a small-size system.

5.2 - Observing frustrated magnetism with multi-basis measurements

Enhanced level of control in quantum simulators enables the preparation of broader classes of initial states [206, 207], the measurement of multi-basis observables [208], and even mid-evolution gates [209]. These advances enabled the integration of novel quantum information protocols with quantum simulators [210–213]. In neutral atom systems, combining ground-state Raman manipulations [214, 215] with the ability to address individual atoms has already allowed for the demonstration of local rotations in such systems [216–218]. However, in the Rydberg-Rydberg encoding, no analogous procedure has been realised. Here, we address this challenge by demonstrating a general protocol implementing nearly-arbitrary local control in a dipolar Rydberg atom array and we adiabatically prepare and measure states exhibiting both ferromagnetic and antiferromagnetic (six-body) chiral-chiral correlation functions.

5.2.1 - Multi-basis measurement protocol

Our protocol to perform multi-basis measurements relies on the combination of microwave pulses and local light-shifts. The microwaves, tuned to the $|\uparrow\rangle - |\downarrow\rangle$ transition (at $\omega_0/(2\pi) \sim 16.7$ GHz), only allow for global rotations. To perform local rotations, we apply light-shifts on specific atoms using addressing beams generated by reflecting a 1013 nm laser on an SLM.

Generating the addressing pattern

The addressing laser beams are generated by an external cavity, 1013 nm diode laser seeding an amplifier outputting up to 8 W. The light is blue-detuned from the $(6P_{3/2}, m_J = 3/2)$ to $|\uparrow\rangle$ transition by $\Delta/2\pi \sim 400$ MHz resulting in a light-shift $\delta \sim \Omega_{1013}^2/(4\Delta)$ for a Rabi frequency Ω_{1013} on an addressed atom. We use a dedicated SLM to produce the pattern of addressing beams, superimposed onto the tweezer array pattern. Each beam is focused on a $1/e^2$ radius of 1.5 μm and induces a light-shift of either 1δ or two 2δ. A power of ≈ 300 mW on one atom results into a light-shift $\delta/(2\pi) \approx 23$ MHz. The average measured light-shifts for the 1δ (and 2δ) atoms are $2\pi \times 22.82$ MHz ($2\pi \times 45.46$ MHz) with a typical dispersion of $2\pi \times 0.2$ MHz ($2\pi \times 0.4$ MHz) between the different atoms. A current limitation of this scheme is that the only characteristic of the addressing pattern that can be dynamically modified during the $\sim 10\,\mu$s timescale of an experiment is the overall amplitude of the light-shift; however, one could envision to circumvent this limitation in the future, using e.g. several SLMs in a multiplexing configuration.

As illustrated in Fig. 5.14(a), the atoms are addressed with different intensities to produce different values of light-shifts, realising the Hamiltonian $H_z = \hbar \sum_i \delta_i (1 + \sigma_i^z)/2$, with $\delta_i = 0\delta$ (atoms not addressed), 1δ or 2δ (with $\delta/2\pi \approx 23$ MHz). From now on, we refer to these classes of atoms as the $0\delta, 1\delta$ and 2δ atoms.

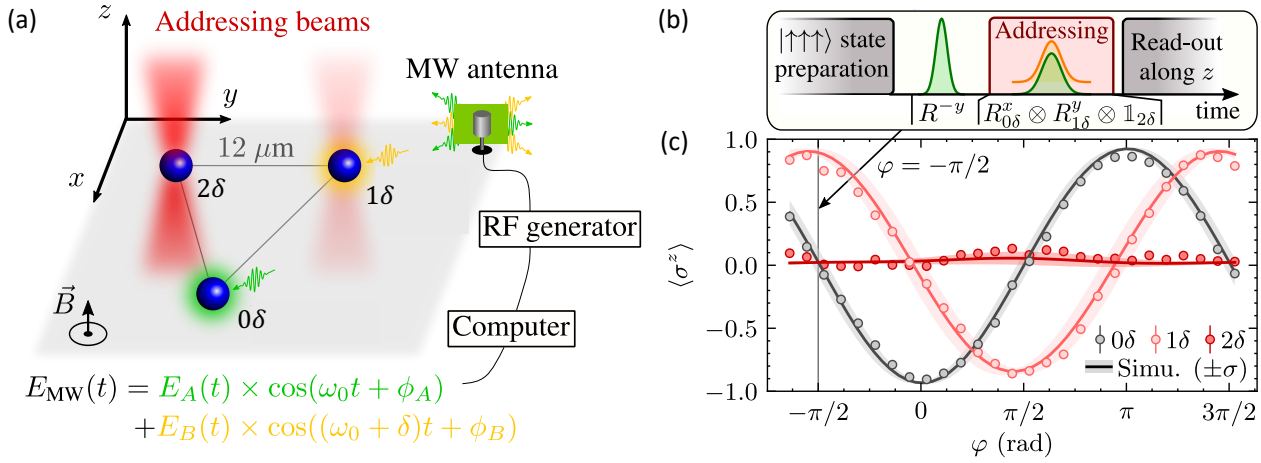


Figure 5.14: **Multi-basis measurement on a triangle.** (a) Experimental set-up. The microwaves at a frequency ω_0 and $\omega_0 + \delta$ (green/yellow) are on resonance with the 0δ and 1δ atom transitions (no colour/light red) and off-resonant with the others (red). (b) Experimental sequence to measure the state of three atoms in the y, z and x basis. A global MW pulse (green) is followed by a combination of addressing (red) and MW pulses (green/yellow). (c) Average magnetisation of each class of atoms during a Ramsey experiment. Experimental data (dots) are benchmarked against simulations including experimental imperfections (solid). The shaded areas represent the standard deviation in MC simulations.

Applying local rotations

To perform local rotations on these three classes, we apply the addressing beams and send simultaneously two microwave pulses with frequencies ω_0 and $\omega_0 + \delta$, resonant with the 0δ and 1δ atoms. This allows for arbitrary qubit rotations of the 0δ and 1δ atoms while the 2δ atoms remain unaffected. By applying a global rotation prior to the local ones, as detailed below, we can now perform measurements in arbitrary bases on three classes of atoms at the same time: the choice of the measurement basis is set by the duration and phase of each microwave pulse with respect to a local oscillator at ω_0.

As an example, Fig. 5.14(b) shows the experimental sequence used to measure the 0δ, 1δ and 2δ atoms along the y, z and x axis. The first microwave pulse applies a global $\pi/2$ rotation along the $-y$. We call R^{-y} the corresponding rotation operator. Then, combining two microwave frequencies with the addressing, we apply the following local rotations $R^x_{0\delta} \otimes R^y_{1\delta} \otimes \mathbb{1}_{2\delta}$ with $R^{\mathbf{u}}_{n\delta}$ the operators corresponding to a $\pi/2$ rotation of the $n\delta$ atoms around the $\mathbf{u}$ axis. This full sequence is thus equivalent to the rotations $(R^x_{0\delta} \cdot R^{-y}_{0\delta}) \otimes (R^y_{1\delta} \cdot R^{-y}_{1\delta}) \otimes R^{-y}_{2\delta}$. As, $R^x_{1\delta} \cdot R^{-y}_{1\delta} = R^z_{1\delta} \cdot R^x_{1\delta}$, and as we measure in the z-basis, the z rotation has no effect on the measured probabilities. The sequence thus amounts to the rotation $R^x_{0\delta} \otimes \mathbb{1}_{1\delta} \otimes R^{-y}_{2\delta}$ Another, more natural, experimental protocol would have been to apply 3 microwave frequencies tuned on the 0δ, 1δ and 2δ-atoms. However residual spatial inhomogeneities on the 2δ light-shifts degraded the fidelities of the rotation in an early attempt.

Building the noise model

In this section, we review the sources of errors decreasing the preparation and detection fidelities. We estimate the fidelity of the Rydberg excitation $\eta_{\text{STIRAP}} = 98 \pm 0.3\%$ [37, 205], thus a bit improved compared to previous section. For the detection errors, a set of independent experiments leads to first order to $\epsilon_\uparrow = (1 - \eta_{\text{deexc}}) + \epsilon = 1.5\% + 1.2\% = 2.7 \pm 0.3\%$ and $\epsilon_\downarrow = 1.5 \pm 0.3\%$.

We also identified six sources of imperfections occurring during the local rotations sequence. (1) The finite value of light-shits compared to the MW Rabi frequency $\Omega_{\max}/2\pi = 5.43$ MHz induces crosstalk with the off resonant microwave leading to imperfect rotations of the atoms. (2) The XY interactions can not be turned off during the ~ 100 ns of local rotations. This rotation time is optimised experimentally as to find a balance between shorter duration with increased $\Omega_{\max}$ while keeping $\delta \gg \Omega_{\max}$ to maximise the rotation efficiencies. (3) Due to the spontaneous emission induced by off-resonant coupling to the short-lived intermediate state $6P_{3/2}$, the addressed atoms in $|\uparrow\rangle$ are slowly depumped to the ground state $5S_{1/2}$. For $\Delta/2\pi \sim 400$ MHz and $\delta/2\pi \approx 23$ MHz we experimentally measure effective lifetimes of ~ 2.3 μs and ~ 1.1 μs for the 1δ and 2δ atoms in the $|\uparrow\rangle$ state. (4) The tightly focused addressing beams apply a ponderomotive force on the addressed atoms, pushing them away from their trap center, thus preventing them from being recaptured before readout. Experimentally, for $\delta/2\pi \approx 23$ MHz, we measure losses of $0.3 \pm 0.3\%$ and $1.3 \pm 0.3\%$ for the 1δ and 2δ atoms when sending a 80 ns addressing pulse. (5) As mentioned above, the light-shifts applied to the atoms are not perfectly homogeneous. We measured a dispersion on the order of 1% after calibration that can drift up to 3% after one day without further calibrations. When the addressing is on, the dispersion results into a variation of the phase accumulation of the 1δ-atoms across the array. This leads to a spread of the angle of rotation of the qubits when sending the microwave pulses for local rotations. (6) Finally, we measure an electronic jitter of ± 2 ns between the addressing and the microwave pulses. It has a

similar effect to that of light shift inhomogeneities. Shot-to-shot, the jitter induces an uncertainty in the angle of rotation of the 1δ atoms leading to imperfect microwave rotations.

Calibrating through Ramsey sequence

We illustrate and benchmark the protocol above by performing a Ramsey experiment: starting from all atoms in $|\uparrow\rangle$, we apply a first global rotation $R^{x\cos\varphi+y\sin\varphi}$, followed by the local rotations $R^x_{0\delta}\otimes R^y_{1\delta}\otimes\mathbb{1}_{2\delta}$ and finally read-out the states for various φ. Each experimental sequence is repeated ~ 500 times to compute the average magnetisations. We expect oscillations of the 0δ and 1δ-atom magnetisation that are out of phase by $\pi/2$. The 2δ-atom magnetisation should remain constant at 0. Fig. 5.14(c) shows the experimental results. We attribute the finite contrast of the oscillations to the experimental imperfections described above. To confirm this, we perform a Monte Carlo simulation including SPAMs, finite Rydberg lifetime, interactions between atoms, jitter of the addressing and depumping and losses induced by the addressing. Taking into account all these experimentally calibrated mechanisms in the numerics yields good agreement with the data.

5.2.2 - Quantum state tomography on entangled states

Basis (0δ,1δ,2δ)	Rotations	$\phi_{\text{all}}(^\circ)$	$\phi_{0\delta}(^\circ)$	$\phi_{1\delta}(^\circ)$
xxx	R^{-y}	-90	–	–
xxy	$(R^{-y}_{0\delta}\otimes R^{-y}_{1\delta})\cdot R^x$	0	-90	-90
xxz	$R^{-y}_{0\delta}\otimes R^{-y}_{1\delta}$	–	-90	-90
xyx	$R^x_{1\delta}\cdot R^{-y}$	-90	–	0
xyy	$R^{-y}_{0\delta}\cdot R^x$	0	-90	–
xyz	$R^{-y}_{0\delta}\otimes R^x_{1\delta}$	–	-90	0
xzx	$R^y_{1\delta}\cdot R^{-y}$	-90	–	90
xzy	$(R^{-y}_{0\delta}\otimes R^{-x}_{1\delta})\cdot R^x$	0	-90	180
xzz	$R^{-y}_{0\delta}$	–	-90	–
yxx	$R^x_{0\delta}\cdot R^{-y}$	-90	0	–
yxy	$R^{-y}_{1\delta}\cdot R^x$	0	–	-90
yxz	$R^x_{0\delta}\otimes R^{-y}_{1\delta}$	–	0	-90
yyx	$(R^x_{0\delta}\otimes R^x_{1\delta})\cdot R^{-y}$	-90	0	0
yyy	R^x	0	–	–
yyz	$R^x_{0\delta}\otimes R^x_{1\delta}$	–	0	0
yzx	$(R^x_{0\delta}\otimes R^y_{1\delta})\cdot R^{-y}$	-90	0	90
yzy	$R^{-x}_{1\delta}\cdot R^x$	0	–	180
yzz	$R^x_{0\delta}$	–	0	–
zxx	$R^y_{0\delta}\cdot R^{-y}$	-90	90	–
zxy	$(R^{-x}_{0\delta}\otimes R^{-y}_{1\delta})\cdot R^x$	0	180	-90
zxz	$R^{-y}_{1\delta}$	–	–	-90
zyx	$(R^y_{0\delta}\otimes R^x_{1\delta})\cdot R^{-y}$	-90	90	0
zyy	$R^{-x}_{0\delta}\cdot R^x$	0	180	–
zyz	$R^x_{1\delta}$	–	–	0
zzx	$(R^y_{0\delta}\otimes R^y_{1\delta})\cdot R^{-y}$	-90	90	90
zzy	$(R^{-x}_{0\delta}\otimes R^{-x}_{1\delta})\cdot R^x$	0	180	180
zzz	$\mathbb{1}$	–	–	–

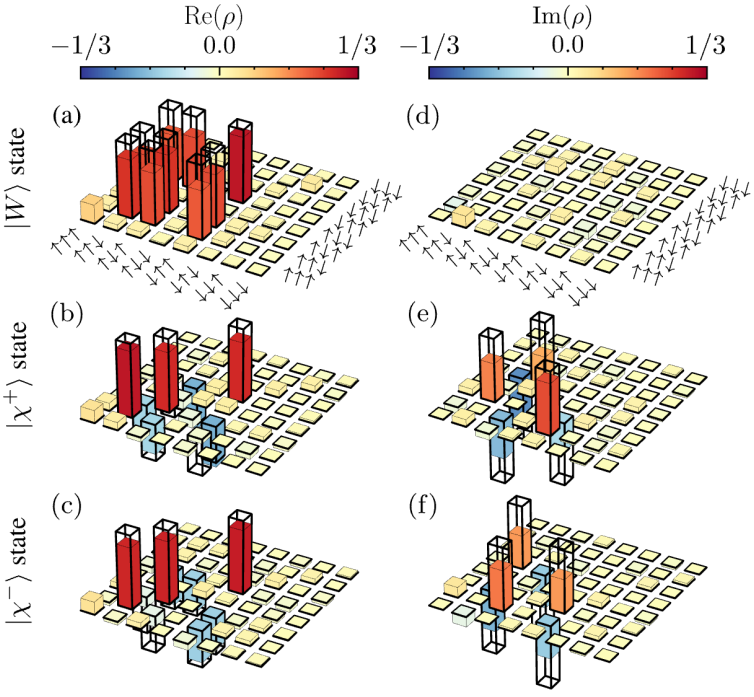


Figure 5.15: **State tomography.** (Table) First column: measurement basis for the 0δ, 1δ and 2δ atoms. Second column: applied rotations. Three last columns: relative phases of the microwave pulses used implement the corresponding rotations (ϕ_{all} refers to the global rotation and $\phi_{0,1\delta}$ to the local ones). The $-$ symbol indicates that the corresponding pulse is off for this sequence. Real (a,b,c) and imaginary (d,e,f) parts of the reconstructed density matrix for the $|W\rangle$, $|\chi^+\rangle$ and $|\chi^-\rangle$ states. The transparent bars represent the expectation values for perfect states.

We now demonstrate how the local control introduced in our work enables the preparation and

detection of complex, correlated states. In particular, we investigate the entangled states of three atoms placed in an equilateral triangle and interaction via H_{XY}. In this configuration, the interaction lifts the degeneracy between $|\uparrow\uparrow\downarrow\rangle$, $|\uparrow\downarrow\uparrow\rangle$ and $|\downarrow\uparrow\uparrow\rangle$, leading to three eigenstates $|W\rangle = (|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle)/\sqrt{3}$ and $|\chi^{\pm}\rangle = (|\uparrow\uparrow\downarrow\rangle + e^{\pm i\frac{2\pi}{3}}|\uparrow\downarrow\uparrow\rangle + e^{\pm i\frac{4\pi}{3}}|\downarrow\uparrow\uparrow\rangle)/\sqrt{3}$ separated in frequency by $3J/\hbar$.

In order to prepare these states, we proceed as described in the preparation phase of Fig. 5.16(a). Starting from all atoms in $|\uparrow\rangle$, we apply a Gaussian microwave pulse at frequency $\omega_0 + 2J/\hbar$ to drive a direct transition from $|\uparrow\uparrow\uparrow\rangle$ to $|W\rangle$. This pulse is performed at a relatively reduced Rabi frequency (much smaller than $3J$) compared to the values used for the measurement rotations (in which $\Omega_{\mathrm{MW}} \gg J$). The Rabi frequency is collectively enhanced by a factor of $\sqrt{3}$, compared to the one measured for single atom Rabi oscillation experiment. Finally, we turn on the addressing light for a duration t_{phase} to imprint a phase 0ϕ, 1ϕ and 2ϕ on the 0δ, 1δ and 2δ atoms, with $\phi(t_{\mathrm{phase}}) = \int_0^{t_{\mathrm{phase}}} \delta(t)\,\mathrm{d}t$, thus preparing $|\chi(\phi)\rangle = \left(|\uparrow\uparrow\downarrow\rangle + e^{i\phi}|\uparrow\downarrow\uparrow\rangle + e^{i2\phi}|\downarrow\uparrow\uparrow\rangle\right)/\sqrt{3}$.

We now exploit our ability to apply arbitrary local rotations to perform quantum state tomography of $|W\rangle$ and $|\chi^{\pm}\rangle$ and reconstruct their density matrix. To do so, we measure the state of each class of atoms in the x, y and z bases, corresponding to $3^3 = 27$ different measurements detailed in Fig. 5.15(table), from which we compute the relevant correlation functions, as well as extract the density matrix (see Box. 19).

Box 19: Density matrices reconstruction

In order to reconstruct the density matrices ρ of quantum states using the measurements in x ,y and z basis, we perform a maximum-likelihood estimation to constrain ρ to be physical. We follow the method described in the Supplemental Materiel of [219]. Any density matrix can be written as $\rho(T) = T^{\dagger}T/\mathrm{Tr}(T^{\dagger}T)$ with T being a complex 8×8 lower triangular matrix with real diagonal elements. Thus T has 64 independent real parameters $(t_1, t_2, \cdots, t_{64})$ that minimise the following cost function:

$$C(T) = \sum_{\alpha\in\{x,y,z\}^3} \sum_{\beta\in\{\uparrow,\downarrow\}^3} \left(\langle\beta|R_{\alpha}^{\dagger}\rho(T)R_{\alpha}|\beta\rangle - P_{\alpha}^{\beta}\right)^2.$$

Here, α is the basis in which we measure each atom, β is an experimental outcome, P_{α}^{β} the probability to measure β in the α basis and R_{α} the set of applied rotations to measure in α. For example, when measuring in the xyz-basis, $R_{xyz} = R_{0\delta}^{-y}\otimes R_{1\delta}^{x}\otimes \mathbb{1}_{2\delta}$. The minimisation process is numerically achieved using a L-BFGS-B algorithm and the density matrix is initialised as the identity.

Figure 5.15 shows, for one triangle, the real and imaginary parts of the density matrices ρ of the three states $|W\rangle$, $|\chi^{+}\rangle$ and $|\chi^{-}\rangle$. From them, we compute fidelities $F = \langle\psi|\rho|\psi\rangle$ of 0.74(1), 0.71(1) and 0.68(1) without correction and 0.80(1), 0.78(1) and 0.74(1) when correcting for detection errors. They are all above 2/3, revealing genuine three-partite entanglement [220–222]. In addition, the produced W-state violates the Mermin-Bell inequality: $S = |\langle\sigma_{0\delta}^{z}\sigma_{1\delta}^{z}\sigma_{2\delta}^{z}\rangle - \langle\sigma_{0\delta}^{x}\sigma_{1\delta}^{x}\sigma_{2\delta}^{z}\rangle - \langle\sigma_{0\delta}^{z}\sigma_{1\delta}^{x}\sigma_{2\delta}^{z}\rangle - \langle\sigma_{0\delta}^{z}\sigma_{1\delta}^{z}\sigma_{2\delta}^{x}\rangle| \leq 2$ as we measure $S_{\mathrm{exp}} = 2.083(26)$ [223]. Much like in the more conventional Bell-state case, this violation rules out a hidden-variable model for the measured correlations.

5.2.3 - Measuring the chirality

Box 20: Intuition for the chirality operator

The classical interpretation of the chirality operator $\chi = \boldsymbol{\sigma}_a \cdot (\boldsymbol{\sigma}_b \times \boldsymbol{\sigma}_c)$ highlights its connection to the handedness of three spins, or the signed volume of the formed parallelepiped by the Bloch vectors. It's odd under time reversal symmetry and invariant under spin rotations. However, it doesn't explain why the maximal value of chirality can't be achieved classically. An intuitive understanding emerges when considering it as a form of angular momentum of spins hopping the three sites forming the triangle of interest.
This picture becomes more transparent when writing the chirality operator in terms of the raising and lower spin operators:

$$\begin{aligned}\vec{\sigma_a} \cdot (\vec{\sigma_b} \times \vec{\sigma_c}) &= \epsilon^{\alpha\beta\gamma}\sigma_a^\alpha \sigma_b^\beta \sigma_c^\gamma \\ &= \sum_{\text{perm } \{a,b,c\}} \left(\sigma_x^a \sigma_y^b - \sigma_y^a \sigma_x^b\right) \sigma_z^c \\ &= \sum_{\text{perm } \{a,b,c\}} \underbrace{2i\left(\sigma_+^a \sigma_-^b - \sigma_-^a \sigma_+^b\right)}_{j_{ba}} \sigma_z^c\end{aligned} \tag{5.9}$$

where j_{ba} is the spin current from spin b into spin a. This means that the state $|\chi^+\rangle$ that maximises the chirality is a state hosting a persistent spin current in a particular direction, or alternatively, the states with maximal angular momentum. Under the action of the chirality operator, each term of the superposition leads to a contribution: $2i\frac{e^{i2\pi/3}-e^{-2\pi/3}}{3} = \frac{4}{3}\sin 2\pi/3 = \frac{2}{\sqrt{3}}$.
The angular momentum is maximised for permutation invariant state; classically, such a state would necessarily exhibit zero chirality since all the spins would be collinear. As a result, tripartite quantum entanglement is necessary to maximise the angular momentum. From the above calculation, an equal superposition of all the rotationally related terms of the wavefunction is needed for the terms to coherently add together. Since the magnetisation is globally conserved, this requires a W-like superposition, which cannot be described using a classical separable state.

Despite $|W\rangle$, $|\chi^+\rangle$ and $|\chi^-\rangle$ exhibiting homogeneous magnetisation and two-point correlation functions, they can be distinguished through their *chirality*. The chirality χ is a spin rotationally symmetric observable that breaks time reversal symmetry and is defined for three spins i, j and k by $\langle\chi_{ijk}\rangle = \langle(\boldsymbol{\sigma}_i \times \boldsymbol{\sigma}_j) \cdot \boldsymbol{\sigma}_k\rangle$, with $\boldsymbol{\sigma}_i = \sigma_i^x \boldsymbol{x} + \sigma_i^y \boldsymbol{y} + \sigma_i^z \boldsymbol{z}$ [224]. For a classical product state, $\langle\chi\rangle$ is bounded by ± 1 and this limit can be overcome for entangled states. $\langle\chi^\pm|\hat{\chi}|\chi^\pm\rangle$ reaches a maximum value of $\pm 2\sqrt{3}$ as explained in Box. 20.

To measure the chirality, we first note that it can be written as the sum of six terms corresponding to the different permutations of $\{x, y, z\}$: $\langle\chi_{0\delta,1\delta,2\delta}\rangle = \langle\sigma_{0\delta}^x\sigma_{1\delta}^y\sigma_{2\delta}^z\rangle + \langle\sigma_{0\delta}^y\sigma_{1\delta}^z\sigma_{2\delta}^x\rangle + \langle\sigma_{0\delta}^z\sigma_{1\delta}^x\sigma_{2\delta}^y\rangle - \langle\sigma_{0\delta}^y\sigma_{1\delta}^x\sigma_{2\delta}^z\rangle - \langle\sigma_{0\delta}^x\sigma_{1\delta}^z\sigma_{2\delta}^y\rangle - \langle\sigma_{0\delta}^z\sigma_{1\delta}^y\sigma_{2\delta}^x\rangle$. For each value of ϕ, we measure each set of bases to compute the total chirality of $|\chi(\phi)\rangle$, similarly to previous work using superconducting qubits [208].

Fig. 5.16(b) shows the results (purple circles) as a function of ϕ, together with the theoretical expectations. The amplitude being reduced due to experimental imperfections, we perform a detailed

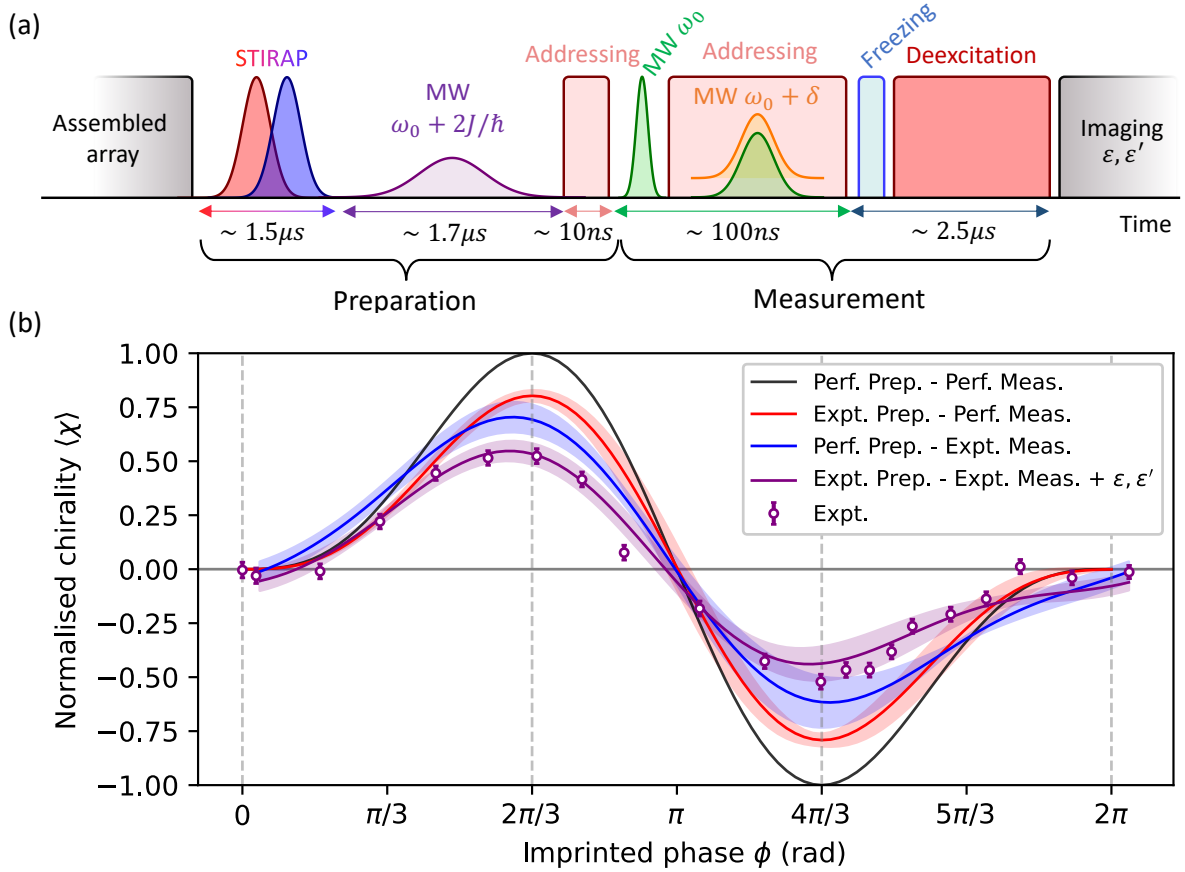


Figure 5.16: **Measuring the chirality.** (a) Experimental sequence used to prepare a $|\chi(\phi)\rangle$ state, from the Rydberg excitation step to the readout. (b) Chirality of $|\chi(\phi)\rangle$ as a function of ϕ. The theoretical amplitude (black) is reduced in experimental implementation (purple dots). Simulating with MC sampling either only imperfect preparation phase (red) or imperfect measurement phase (blue) gives the separate contributions of each phase. Simulating the whole procedure and accounting for detection errors (purple) enables to explain the data. The shaded areas represent the standard deviation.

numerical simulation of the experiment including all the sources of noise described above. We simulate the dynamics of this 5^3-level system ($|g\rangle, |e\rangle, |\uparrow\rangle, |\downarrow\rangle, |D\rangle$) using the package *qutip*, starting from $|g\rangle^{\otimes 3}$ and solving the Lindblad equation for each step of the sequence and each set of experimental parameters.

We account for the shot-to-shot fluctuations in the inter atomic distances and timing jitters by a Monte-Carlo sampling of the experimental parameters with the estimated bounds given in the previous sections. We also include inhomogeneities of the light-shifts across the array. As solving the Master equation becomes quickly resource-consuming ($\sim$ 1 hour/shot), we average the simulation results over only 20 shots, which appears to be enough to reproduce the experimental data with reasonable agreement. At the end of the Monte Carlo simulation, we apply to the results the addressing-induced

atom losses and readout errors following the procedure described in previous sections. The simulation predicts a STIRAP efficiency for one atom of 98.3%, in very good agreement with the experiment. The phases of the microwave pulses are optimised so as to reproduce the data in Fig. 5.14(c), and are then fixed to these values.

To simulate the chirality curves shown in Fig. 5.16(b) and assess the impact of the various imperfections, we first calculate the preparation fidelity of the $|W\rangle$. We obtain $\eta_W = 88.4 \pm 0.4\%$, with contributions to the infidelity of 4.0% from the STIRAP finite efficiency, 6.0% from the Rydberg lifetime, $0.5 \pm 0.2\%$ from positional disorder and $1.0 \pm 0.2\%$ from the microwave pulse finite amplitude. We experimentally measure a fidelity of $80 \pm 1\%$, corrected for detection error, with the tomography performed above. We attribute the remaining 8% of fidelity difference between experiment and simulation to various imperfections occurring during the rotations. We then include the phase imprinting step to prepare the state $|\chi(\phi)\rangle$. Measuring exactly the normalised chirality $\langle\chi\rangle$ of this imperfectly prepared state (by dividing by $2\sqrt{3}$) yields a maximum (minimum) value of $0.8(-0.79)$ (red curve in Fig. 5.16(b)). Applying the experimental measurement sequence for each of the 6 components of the chirality and starting from a perfect $|\chi(\phi)\rangle$ leads to normalised peak values of $0.7(-0.62)$ (blue curve). Finally, we simulate the preparation and the measurement phase, adding the detection errors, and obtain maximum (minimum) values $0.55(-0.44)$ (purple curve), in good agreement with the data. We thus find that the main limitations of the chirality measurement are the imperfections during the measurement phase.

5.2.4 - Frustration on triangular plaquettes

Having leveraged our local control to prepare and probe entangled states, we now demonstrate the power of this tool set in a quantum simulation experiment.

Using a frustrated geometry consisting of a pair of triangular plaquettes [Fig. 5.17(a1)], we attempt to adiabatically prepare low-energy states of the antiferromagnetic dipolar XY model. Strictly speaking, since in Eq. (5.1) we have ferromagnetic couplings $J < 0$, we explore the low-energy properties of the antiferromagnetic model ($J > 0$) by preparing the highest-energy state(s) of our quasi-isolated system. Owing to time-reversal symmetry, all states in the spectrum exhibit zero chirality, $\langle\chi\rangle = 0$. However, exact diagonalisation demonstrates that, the two lowest-energy states exhibit large, but opposite, chiral-chiral correlations. To illustrate this feature, we prepare both the ground and first excited states, by carefully choosing an appropriate pattern of local light shifts.

Our protocol proceeds similarly as in the first section: after initialising all the atoms in $|\downarrow\rangle$, we turn on a pattern of local light shifts [Fig. 5.17(a)]; we then apply a microwave pulse to rotate the non-addressed atoms to $|\uparrow\rangle$. This prepares a product state which is the lowest energy state of H_Z. Starting with $\delta \gg |J|$, we then reduce the light-shift as $\delta(t) = \delta_0 e^{-t/\tau}$ [with $\tau = 0.55$ μs and $\delta_0/2\pi = 23(46)$ MHz for the $1\delta(2\delta)$-atoms], thus quasi-adiabatically connecting the initial Hamiltonian ($\approx H_Z$) to the final one H_{XY}.

In general, such an adiabatic protocol is expected to prepare the ground state of the final Hamiltonian, regardless of the details of the ramp. This expectation fails when the system's ground state exhibits a level crossing, which requires either fine-tuning or some underlying symmetry. Utilising our ability to shape the addressing light, we thus consider two different patterns exhibiting distinct symmetries: pattern 1 respects a mirror symmetry M_y along the y-axis [Fig. 5.17(a1)], while pattern 2 respects inversion symmetry $\mathcal{I}$ [Fig. 5.17(a2)]. For the first pattern, both initial and ground states

live in the same symmetry sector of M_y and thus are adiabatically connected [Fig. 5.17(b1)]. We thus expect to prepare the ground state, leading to the observation of anti-ferromagnetic chiral-chiral correlations. By contrast, for the second pattern, the initial and ground states live in *different* symmetry sectors of $\mathcal{I}$ and thus cannot be adiabatically connected [Fig. 5.17(b2)]. We thus expect to prepare the first excited state, which exhibit ferromagnetic chiral-chiral correlations.

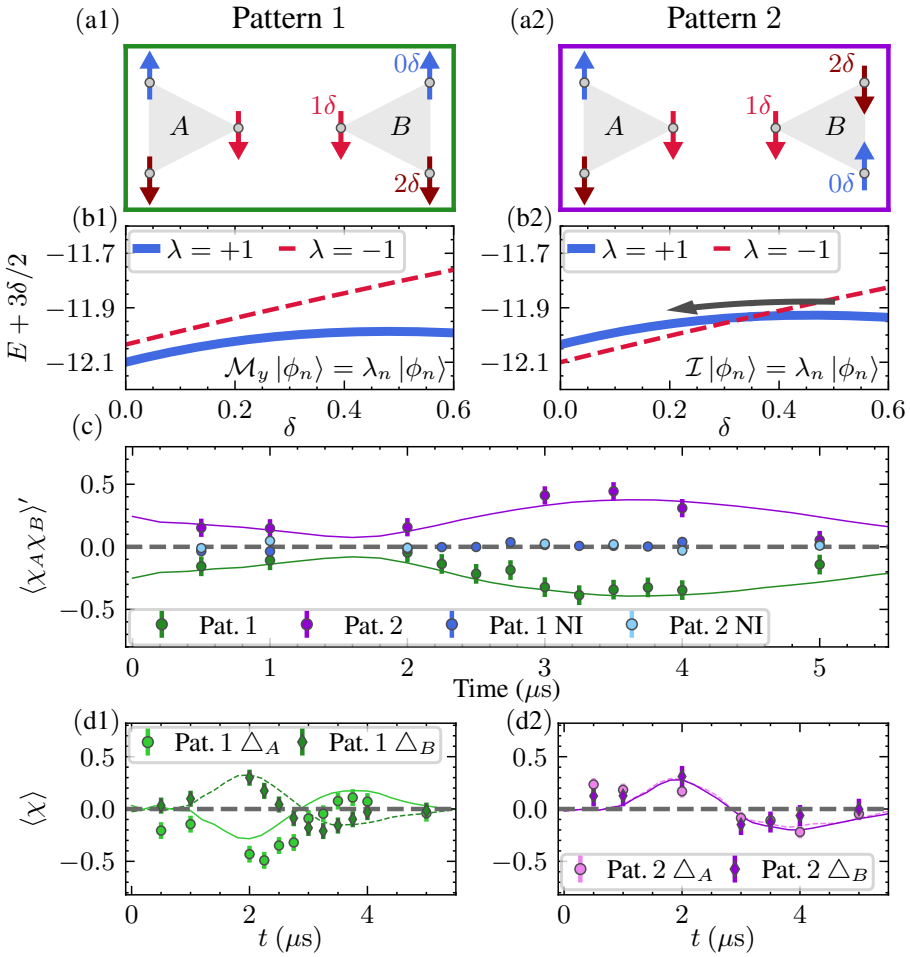


Figure 5.17: **Adiabatic preparation of low energy states** (a) Light shift patterns $(0, 1\delta, 2\delta)$ and associated initial states. Pattern 1 respects mirror symmetry along the y-direction M_y whereas pattern 2 respects inversion symmetry $\mathcal{I}$. (b) Energy spectrum during the adiabatic ramp for pattern 1[2]. (c) Connected chiral-chiral correlations $\langle \chi_A \chi_B \rangle'$ for pattern 1 (green), pattern 2 (purple) and non interacting triangles (blue). The early time ($t \lesssim 1.5\mu s$) observation of non-zero $\langle \chi_A \chi_B \rangle'$ is due to a necessary waiting period before measurement, during which the system undergoes additional dynamics.). (d) Chirality for triangles A and B under the two patterns.

We experimentally explore this difference using the multi-basis measurement protocol described above. It not only allows measuring the chirality $\langle \chi(t) \rangle$ of a single triangle, but it also enables the measurement of six-body correlation functions, from which the two-triangle chiral-chiral correlations, $\langle \chi_A \chi_B \rangle$, can be extracted.

In principle, the full reconstruction of $\langle \chi_A \chi_B \rangle$ requires the measurement of 36 different terms. However, a smaller subset of six terms is sufficient to faithfully capture the system's correlations as detailed in Methods of [58]. We can, therefore, use the same addressing pattern for both the adiabatic ramp and the multi-basis measurement. More specifically, we measure:

$$\langle \chi_A \chi_B \rangle' = \eta \sum_{a,b,c \in \mathrm{perm}(x,y,z)} \begin{array}{l} \langle \sigma^a_{0\delta} \sigma^b_{1\delta} \sigma^c_{2\delta} \tilde{\sigma}^a_{0\delta} \tilde{\sigma}^b_{1\delta} \tilde{\sigma}^c_{2\delta} \rangle \\ -\langle \sigma^a_{0\delta} \sigma^b_{1\delta} \sigma^c_{2\delta} \rangle \langle \tilde{\sigma}^a_{0\delta} \tilde{\sigma}^b_{1\delta} \tilde{\sigma}^c_{2\delta} \rangle \end{array} \tag{5.10}$$

where $\sigma[\tilde{\sigma}]$ refers to spins in triangle A[B] and $\eta = \pm 1$ is set by the relative handedness of the two three-spin measurement patterns: $\eta = -1$ for pattern 1 and $\eta = 1$ for pattern 2.

We begin by studying the quasi-adiabatic ramp using the pattern depicted in Fig. 5.17(a1). Focusing on the connected chiral-chiral correlation, we observe the development of strong anti-ferromagnetic $\langle \chi_A \chi_B \rangle'$ correlations that persist to late times [Fig. 5.17(c), green]. This observation is consistent with a preparation yielding more than 50% population in the ground state. By contrast, when considering the second pattern [Fig. 5.17(a2)], the dynamics exhibit similar features but with *opposite* sign. The presence of equally strong ferromagnetic $\langle \chi_A \chi_B \rangle'$ correlations is consistent with an equally large population in the first excited state of the system. To demonstrate that our observations indeed arise from the dipolar interactions between the two triangles, we also measure $\langle \chi_A \chi_B \rangle'$ for non-interacting triangles separated by ≈ 72 μm. In this case, neither patterns lead to significant correlations [Fig. 5.17(c), blue].

Finally, we discuss two important sources of imperfections in our protocol. First, although τ was chosen to be much longer than the timescale of the system $(1/J)$, residual diabatic errors manifest themselves in a small chirality value $\langle \chi(t) \rangle$ [Fig. 5.17(d)]. Second, there are fluctuations in the positions of the atoms owing to their initial position and velocity uncertainty upon the release from the tweezers. As a result, for each repetition of the experiment, the atoms experience slightly different time-dependent interactions, that ultimately lead to the damping of the chirality oscillation and to the decay of the connected chiral-chiral correlations.

5.2.5 - Conclusion

In conclusion, we have demonstrated a new tool combining global microwaves and local light-shifts to enable local control of qubits encoded in Rydberg levels. Our protocol is generic and can be extended to an arbitrary number of classes of atoms. The agreement between experiments and simulations highlights our good understanding of error in our system—a crucial ingredient for further improvements.

More broadly, this work opens the doors to a number of intriguing directions. First, the measurement of multi-body correlation functions can capture the intricate correlations that characterise complex phases of matter such as time reversal symmetry breaking and topological order [225]. Second, the ability to measure along arbitrary bases enables the implementation of novel certification protocols [226]. Finally, by interspersing unitary rotations with analog quantum simulation, one can study multi-time correlation functions as well as more varied dynamical protocols [126, 208, 227].

Overall, in this chapter, we have demonstrated the use of Rydberg based quantum simulators to study with the dipolar XY model frustration in spins system. In the following part of this thesis, we

will return to the Ising model and show how industrial applications can also be tackled, especially with the help of combinatorial optimisation and machine learning methods.

Part III

Quantum algorithms : exploring innovative applications with graphs

6 - Combinatorial optimisation with Rydberg atoms

Contents

Representing data in the form of graphs is ubiquitous in a multitude of scientific, engineering and industrial scenarios as they naturally represent complex connections between entities. They naturally describe how individuals are connected based on their relationships in social networks [228], characterise interactions of proteins and genes [229] and can represent the structure of sentences in linguistics [230]. The visual abstractions and universality that graphs provide make them invaluable tools for addressing and understanding complex problems, such as combinatorial optimisation ones. Harnessing the innate properties of neutral atoms offers an appealing avenue in achieving enhanced computational capabilities in graph-related optimisation tasks, such as Maximum Independent Set (MIS) problems. The ability to solve the MIS problem has broad implications, extending to various other graph-related problems, given their connection in graph theory, and aligning with numerous industrial challenges [231].

In this chapter, we delve into the process of mapping graphs onto arrays of trapped neutral atoms to faithfully reproduce their topologies through pairwise van der Waals interactions. We explain how variational quantum algorithms can be used to find MIS solutions and implement the corresponding hybrid classical-quantum loops on Fresnel for several approaches. Subsequently, we construct an adiabatic protocol applicable to a family of similarly structured graphs, demonstrating its applicability in an industrial scenario involving smart-charging tasks for electric vehicles. Finally, we address the

limits on the kind of combinatorial problems we can tackle with current hardware, while also surveying diverse strategies to augment these capacities.

Some elements of this chapter are taken from the following articles to which I contributed :

- [232] C. Dalyac et al., “Graph Algorithms with neutral atom quantum processors”, arXiv, 2024.
- S. Tibaldi et al., "Analog QAOA with Bayesian Optimization on a neutral atom quantum processor", *in preparation*.

6.1 - From graphs to spins to Rydberg atoms

6.1.1 - Mapping combinatorial problems to spins systems

Combinatorial problems involving graphs typically revolve around selecting an optimal configuration from a finite set of distinct elements, all while adhering to specified constraints. Let us start with a simple but stressful enough example for anyone who has ever faced it, entitled *making a guest list*.

Inviting people is hard, seriously

For any event, be it a wedding or a thesis defense celebration, selecting the attendees involves navigating a delicate balance. One must carefully choose from a pool of individuals, bearing in mind potential conflicts or strained relationships among them. This intricate network can be modelled by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where each vertex $i \in \mathcal{V}$ represents a prospective guest and an edge $(i, j) \in \mathcal{E}$ between two guest i, j represents their relationship as showcased in Fig. 6.1(a).

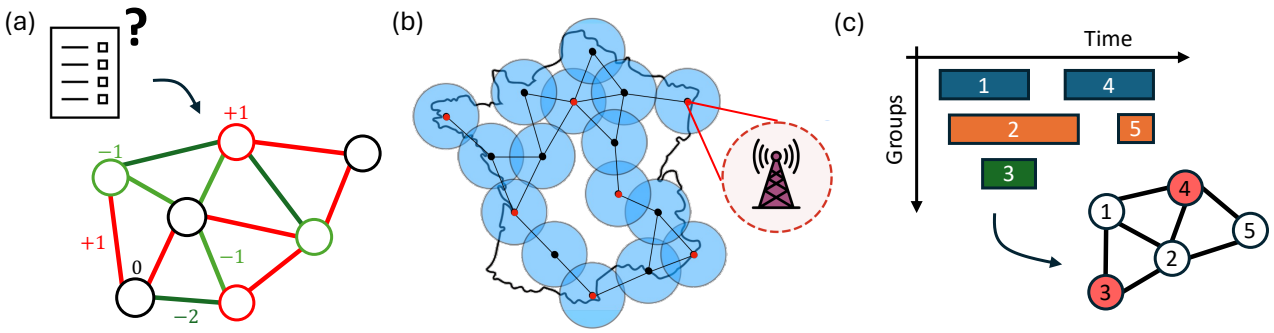


Figure 6.1: **Instances of combinatorial problems solved by finding MIS.** (a) A list of potential guests can be represented as a graph where persons/nodes are connected by weighted edges according to their relationships. Nodes can also be weighted according to the relationship between the host and their guest for instance. (b) Broadcasting problem where identical radio transmitters over French cities have the same radius of transmission. One needs to assign distinct frequencies to nearby towers to avoid interference but the finite bandwidth requires using as few frequencies as possible. The MIS indicated by red dots shows which towers could share the same frequency. (c) Example of a Group Interval Scheduling problem described in Sec. 6.4.

Depending on the situation, one will assign weights w_i to each vertex and w_{ij} to each edge with $w \in \mathbb{R}$, arbitrary scores. Planning successfully the event would then require minimising the following

cost function

$$C_{\mathcal{G}}(\boldsymbol{z}) = \sum_{i\in\mathcal{V}} w_i z_i + \sum_{(i,j)\in\mathcal{E}} w_{ij} z_i z_j, \tag{6.1}$$

where z_i is a binary variable, equal to 1 when node i is selected (guest i is invited) and 0 otherwise. A configuration $\boldsymbol{z}$, or guest list, is thus successful (optimal), according to the scenario given by the weights, when the cost $C_{\mathcal{G}}(\boldsymbol{z})$ is low (the lowest). For instance, consider a scenario where each individual, regardless of their specific identity, is assigned a negative weight $w_i = -1$. This directly translates into the more people the merrier as adding a new person decreases $C_{\mathcal{G}}(\boldsymbol{z})$. By then assigning $w_{ij} = 0$ for pairs of people getting along and $w_{ij} = c > 0$ for pairs who don't, adding someone on bad terms with an already invited guest results in an increase of $C_{\mathcal{G}}(\boldsymbol{z})$. Here, the ratio c accounts for the personal balancing on inviting as many guests as possible ($c \ll 1$) or making sure that the event would be trouble-free ($c \gg 1$). Finding a solution to this problem implies finding a maximum subset of nodes not connected to each other, i.e. with $w_{ij} = 0$, and is known in the optimisation field as the Maximum Independent Set problem. As more and more details and constraints are taken into account, this problem can even be made more complex. For instance, the personal benefit of choosing someone could depend on its identify as inviting your best friend i to your birthday holds more significance than a total stranger j, i.e. $w_i < w_j < 0$. Moreover, the links between people can be further quantified by the scores $w_{ij} > 0$ with a thousand-year-old rivalry weighing more than a mere feeling of discomfort. Bringing along a group of friends who all appreciate each other will benefit the event, and can be modelled by a cluster of connected nodes with negative weights on the edges, i.e. $w_{ij} < 0$. The MIS thus become a weighted MIS (MWIS).

As many people have noticed in the run-up to their wedding, planning this guest list/solving this MIS problem is in fact hard. With the number of elements N increasing, the exponential number of possible configurations $|\mathbb{B}^N| = 2^N$ explodes rendering an exhaustive search quickly intractable. With $N = 100$ guests for instance, checking the $2^{100} \approx 10^{30}$ arrangements would take thousand of years to even the most advanced supercomputer. Interestingly, no other method than this straightforward but painfully time-consuming one has been found to solve this problem exactly. For this reason, MIS problems have been cast to the *NP-hard* complexity class, stating that no known polynomial-time algorithm can solve them exactly on all instances [233]. Approximate methods will be discussed in Sec. 6.4.4.

Parallel with statistical physics

Using graphs to embed instances, several other problems, such as network flows, scheduling, max-cut, max-clique or vertex cover, can be tackled [234]. They can be formalised as Quadratic Unconstrained Binary Optimisation (QUBO) problems, encompassing the problem and the specific graph instance $\mathcal{G}$ to solve such that the cost function to optimise reads

$$C_Q(\boldsymbol{z}) = \boldsymbol{z}^T Q \boldsymbol{z} = \sum_{i=1}^{N} Q_i z_i + \sum_{i,j=1}^{N} Q_{ij} z_i z_j \tag{6.2}$$

with $N = |\mathcal{V}|$, $\boldsymbol{z} \in \mathbb{B}^N$, a binary vector/bitstring and $Q \in \mathcal{M}_N(\mathbb{R})$, a symmetric matrix containing the weights of the problem. Thus solving a QUBO Q amounts to find the bitstring minimising C_Q.

Interestingly, this formulation is exactly the same as classical Ising models of spins σ_i as the Hamiltonian of such system reads

$$H = \sum_{i=1}^{N} h_i \sigma_i + \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j = C_Q(\boldsymbol{\sigma}) \tag{6.3}$$

with Q encompassing the terms related to J the interaction exchange and h the magnetic field. Finding the ground state of this Ising Hamiltonian is equivalent to finding the optimal configuration $\boldsymbol{z}$ minimising C_Q. This connection between two seemingly distant fields helped leveraging methodologies from statistical physics, notably *simulated annealing* described in Box. 21, to tackle challenging QUBO instances [235, 236]. For instance, we utilise this approach in chapter 8 to benchmark a quantum-enhanced QUBO-solving algorithm. Conversely, computational complexity theory also yielded significant insights into the hardness of polymer protein folding [237]. Notably, in 2014, all of Karp's 21 NP-complete problems were explicitly translated into equivalent Ising spin formulations [238].

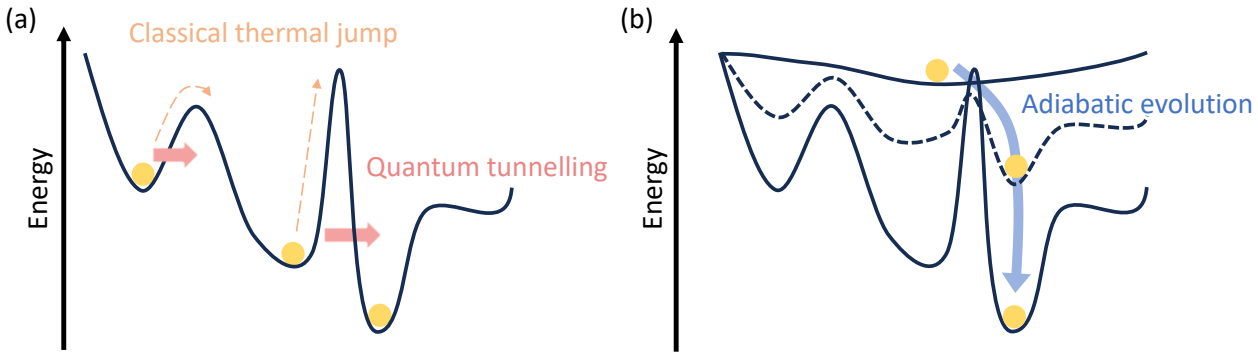


Figure 6.2: **Minimising a system energy through annealing.** The energies of the possible states (yellow ball) of a system forms a landscape (black line) with local minima and a global minimum. (a) Classical thermal jumps (orange), used in simulated annealing, enables to jump over low and wide potential barriers while tunnelling (red), used in quantum annealing, can help transfer through thin and high barriers. (b) Initialising a state in the global minimum of a landscape (highest line) and evolving the system slowly enough to a final landscape (lowest line) should keep the system in the instantaneous global minimum through adiabatic evolution (blue), as seen in quantum adiabatic algorithms.

6.1.2 - Native embedding of Unit Disk graphs with Rydberg atoms

The topology of the interaction Hamiltonians obtained with Rydberg atoms natively echoes the ones of the classical spins models, making the neutral atom technology a potential resource to tackle specific graph problems. By utilising the spatial dependency $U(r_{ij})$ of the two body terms in Eq. 1.19/1.21, one can for instance reproduce the adjacency matrix $\mathcal{A}$ of a graph $\mathcal{G}$ as shown in Fig. 6.3. We define the corresponding graph Hamiltonian as

$$\hat{H}_{\mathcal{G}} = \sum_{(i,j)} \mathcal{A}_{ij} \hat{n}_i \hat{n}_j \tag{6.6}$$

Box 21: Simulated annealing

Simulated annealing is a probabilistic black-box optimisation algorithm inspired by thermodynamics, wherein solutions to complex problems are iteratively refined by probabilistically accepting uphill moves based on a decreasing temperature parameter as shown in Fig. 6.2(a). For a cost/energy function C with N binary variables, one can introduce a Boltzmann probability measure p_β on the space of configurations $\mathbb{B}^N$ such that for each configuration $\boldsymbol{z}$,

$$p_\beta(\boldsymbol{z}) = \frac{1}{\mathcal{Z}(\beta)} e^{-\beta C(\boldsymbol{z})}, \quad \text{where } \mathcal{Z}(\beta) = \sum_{\boldsymbol{z}\in\mathbb{B}^N} e^{-\beta C(\boldsymbol{z})}. \tag{6.4}$$

$\beta > 0$ acts as an inverse temperature and the partition function $\mathcal{Z}(\beta)$ ensures the normalisation of the probability law. This temperature can be related to a kinetic energy term which can help the system escape local minima in a *thermal hopping* fashion. At high temperatures ($\beta \approx 0$), the probabilities are spread out across the configuration space. As the temperature decreases ($\beta \to \infty$), the probabilities concentrate around configurations with lower energy, much like cooling a material causes it to solidify into its more stable shape.
The simulated annealing algorithm mimicks this cooling process using a Markov chain Monte Carlo method. Starting from a random configuration, the temperature is gradually lowered over a series of steps called the *annealing schedule*. At each temperature $1/\beta_i$, the algorithm explores the configuration space through a random walk, with transitions between configurations occurring depending on p_{β_i}. For instance, a proposed move from $\boldsymbol{z}$ to $\boldsymbol{z}'$ can be accepted with the following law :

$$P(\boldsymbol{z} \to \boldsymbol{z}') = \begin{cases} 1 & \text{if } C(\boldsymbol{z}') < C(\boldsymbol{z}) \\ p_{\beta_i}(\boldsymbol{z}')/p_{\beta_i}(\boldsymbol{z}) & \text{else} \end{cases} \tag{6.5}$$

The random walk focuses more and more on the low-cost regions as the temperature is reduced. An important remark is that this method is guaranteed to converge to the optimum if for every temperature β_i the random walk, seen as a Markov chain, converges to its stationary distribution. By the end of the annealing schedule, at very low temperatures, the algorithm converges to the configuration with the minimum energy, which corresponds to the optimal solution to the optimisation problem. A main drawback of simulated annealing lies in the arbitrary selection of numerous hyperparameters, such as the initial temperature, the temperature decay schedule, the stopping criteria, or the length of each temperature stages. Moreover, for some optimisation problems/graphs, the required number of iterations of the random walks exhibit exponential scaling with N [239].

More precisely, assigning qubits to vertices enables a direct correspondence between their pairwise interactions and the weights of the edges. Mapping a N-node graph to an interaction Hamiltonian thus requires placing N qubits such as to satisfy the $N(N-1)/2$ constraints $U(r_{ij}) = \mathcal{A}_{ij}\ \forall (i<j)$ up to a global scaling coefficient. The task of solving this *embedding* problem poses challenges in itself, and the subsequent sections will delve into various options for resolution.

As a reminder, the Rydberg blockade effect can be approximated as enforcing that pairs of atoms

can not be both excited at the same time, if located closer than a certain threshold r_b. Incidentally, distant pairs remain free from this constraint. This binary perspective, solely depending on the threshold r_b, natively mimicks the topology of Unit-Disk (UD) graphs where two vertices i, j are sharing an edge if and only if they lie within a threshold distance from each other in the Euclidean plane. These graphs are part of the larger family of intersection graphs, that appear in concrete problems such as wireless networking [240] as shown in Fig. 6.1(b), computational biology [241], map labelling [242] or to maximise yield when cutting chips from a large chip wafer [243].

Thus, embedding UD graphs into a neutral atom system amounts to a free-space positioning problem with $2N$ variables $\boldsymbol{r} = (x_i, y_i)_{i=1\cdots N}$ and the constraints $U(r_{ij})/U(r_b) \gg 1$ for $(i,j) \in \mathcal{E}$ and $\ll 1$ otherwise.

The resulting Ising interaction Hamiltonian reads

$$\begin{aligned}\hat{H}_{\text{int}}(\mathbf{r}) &= \sum_{(i,j)\in\mathcal{E}} U(r_{ij})\hat{n}_i\hat{n}_j + \sum_{(i,j)\notin\mathcal{E}} U(r_{ij})\hat{n}_i\hat{n}_j \\ &= U(r_b)\hat{H}_{\mathcal{G}} + U(r_b)\sum_{(i,j)\in\mathcal{E}}\left(\frac{r_b^6}{r_{ij}^6} - 1\right)\hat{n}_i\hat{n}_j + \sum_{(i,j)\notin\mathcal{E}} U(r_{ij})\hat{n}_i\hat{n}_j\end{aligned} \tag{6.7}$$

where we artificially separate into two the sum over pairs sharing an edge in order to highlight the presence of $\hat{H}_{\mathcal{G}}$.

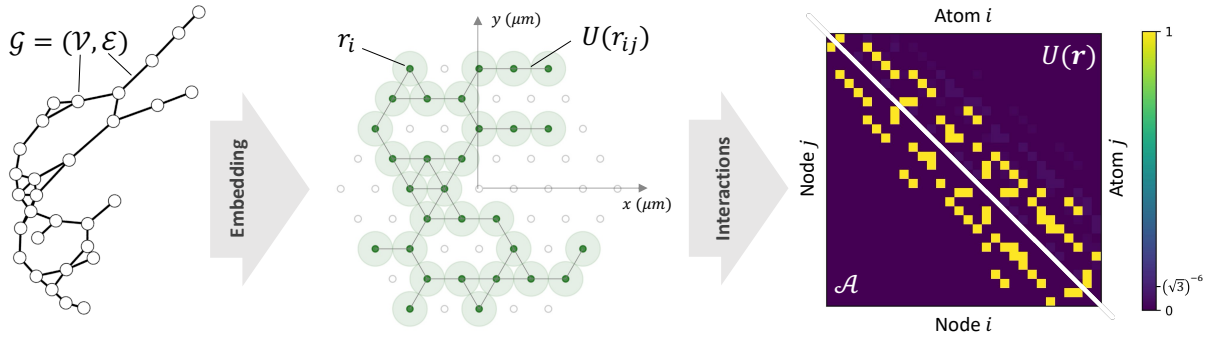


Figure 6.3: **Representing graph connectivity with interactions.** With an adequate spatial positioning $\boldsymbol{r}$ of the atoms in a register (green), their interactions (upper corner of the matrix) can reproduce the adjacency matrix $\mathcal{A}$ (lower corner of the matrix) of an input graph $\mathcal{G}$, up to a global scaling factor. The sharp decay of the interaction allows neglecting terms between second (and further) nearest neighbours.

The first sum accounts for the variation of distances between pairs representing an edge. For a perfect embedding, all linked pairs are spaced by r_b, cancelling the sum. Non-zero terms in that sum can lead to lifts of degeneracy of blockaded states in the spectrum of $\hat{H}_{\text{int}}(\mathbf{r})$. For instance, a 2×2 square with atoms spaced by a represents in a fully blockaded regime, i.e. $r_b > \sqrt{2}a$, the simplest King's graph where all four nodes are connected. While all combinations of highlighting two nodes out of four hold the same cost in the graph, exciting two atoms on the same border or two atoms

across a diagonal does not yield the same energy in $\hat{H}_{\text{int}}(\mathbf{r})$. It can be worth to note when wanting to reproduce perfectly the spectrum but harmless when only considering ground state mapping.

The second sum gathers the spurious terms generated by the tail of the interactions which are usually neglected in the blockade approximation. One has to ensure that the largest distance between a pair sharing an edge in the graph is always far less than the shortest distance between a pair not sharing an edge. For instance, when embedding onto a regular register, one needs to be aware of the ratio between nearest-neighbour (NN) distance r_{nn} and next nearest-neighbour (NNN) r_{nnn} distance. For a theoretical clean embedding of a binary adjacency matrix, where all edges have the same weight, the distances should be chosen as $r_{nn} \ll r_b \ll r_{nnn}$, in order for the NN pairs to represent the edges and the interactions between NNN pairs to be neglected. More practically, a good choice for the blockade approximation to work is $r_b \approx \sqrt{r_{nn} r_{nnn}}$. However, for regular lattice or arbitrary configurations of atoms, having $r_{nnn}/r_{nn} \gtrapprox 1$ will translate into erroneous blockade approximation and thus smaller energy separation between blockaded and non blockaded states. Going back to the previous square geometry, this ratio is $\sqrt{2}$. Therefore, embedding a graph corresponding to a square without diagonal connections is especially tricky as the diagonal terms, supposedly not representing edges, will still contribute with strength 1/8 regardless of the chosen blockade radius.

More involved strategies can be used to entirely get rid of those unwanted terms. For instance, one can combine dynamical decoupling/Floquet engineering techniques [244] together with Trotterisation, in order to implement stroboscopically all the various Hamiltonian terms. Another approach valid in the Ising encoding consists of taking advantage of the multi-level nature of atomic qubits. One can thus use an additional non-interacting ground state to temporarily store the quantum state of a group of qubits, building on the idea of Ref. [29] where this approach was used for digital-analog computing. Along the same line of thought, switching on and off desired interaction terms for selected durations can be achieved through Rydberg dressing [245, 246].

A perfect embedding should result in each term of the two remaining sums of Eq. 6.7 to be zero such that $\hat{H}_{\text{int}}(\mathbf{r}) \propto \hat{H}_{\mathcal{G}}$. This establishes a perfect correspondence between the two spectra, ensuring that the order of the energy levels is preserved. For specific graphs such as UD ones generated on bi-dimensional regular lattices, the mapping process is relatively straightforward. Consider $\mathcal{G}$ as a UD graph derived from selecting a subset of sites on a triangular lattice, with r_b matching the lattice spacing. In this scenario, the first sum cancels out, and the dominant terms in the second sum exhibit a strength of $U(r_b)/27$, allowing for their neglect. While the mapping is effective for this particular case and situations where UD graphs can be mapped to a triangular layout, for most graphs, embedding them into an atomic register necessitates an optimisation procedure, keeping in mind the hardware constraints.

In the following, we will focus on graphs mappable to layouts obtained from regular lattices, such as triangular, and build variational quantum algorithms (VQA) tasked to solve their associated MIS problems.

6.2 - Solving the UD-MIS problem with variational algorithms

Exploiting the representation of a graph's adjacency matrix within the interaction Hamiltonian of a set of Rydberg atoms has spurred the development of various proposals for solving MIS problems [247]. For a graph $\mathcal{G}$, the MIS cost function outlined in Eq. 6.1 can be encoded as an operator of the form:

$$C_{\mathcal{G}}(\hat{n}) = \hat{C}_{\mathcal{G}} = -\sum_{i\in\mathcal{V}} \hat{n}_i + c\hat{H}_{\mathcal{G}}. \tag{6.8}$$

Here, nodes are uniformly weighted with -1, and edges between adjacent nodes carry a uniform weight of $c > 0$. The ground state of this operator manifests as a coherent superposition of MIS classical product states. Such an operator finds correspondence with the Rydberg version of the Ising Hamiltonian (with $\Omega = 0$ and $\delta > 0$) of $N = |\mathcal{V}|$ atoms well-positioned such that:

$$\hat{H}/\hbar = -\delta \sum_{i=1}^{N} \hat{n}_i + \hat{H}_{\text{int}}/\hbar(\boldsymbol{r}) \approx \delta \left[-\sum_{i=1}^{N} \hat{n}_i + \underbrace{\frac{U}{\delta}}_{=c} \hat{H}_{\mathcal{G}} \right] = \delta \hat{C}_{\mathcal{G}} \tag{6.9}$$

where $U = U(r_b)/\hbar$ and the ratio between the next-nearest interactions and the detuning encodes the edge weight c. By preparing the ground state of $\hat{H}$ and subsequently sampling it n_{shots} times, it becomes possible to measure not just one, but all the MISs of $\mathcal{G}$ as they are degenerate. The paradigm of VQA involves leveraging a variational optimisation scheme to discover a control protocol capable of preparing such ground states with good efficiency. This approach tackles in essence the challenges of ground state preparation occurring in the quantum simulation field and discussed in chapters 4 and 5.

6.2.1 - Main ingredients of a VQA

As showcased by Fig. 6.4, a VQA is essentially a hybrid loop between an optimisation procedure fully handled by classical resources (CPU) and a QPU able to perform the quantum dynamics associated with a given control. Communication between the two entities boils down to new protocols being sent from the CPU to the QPU, and to measurements being sent back by the QPU and used by the CPU to compute a better protocol. The loop is iterated until a convergence criterion is met, the simplest being until a fixed number of iterations has been performed. Each iteration k of the loop involves a protocol parameterised by $\boldsymbol{\theta}^{(k)}$, obtained following the three steps described below.

Cost estimation: MIS probability and approximation ratio

The cost estimation procedure attributes a cost $C(\boldsymbol{\theta})$ to each produced quantum state $|\psi(\boldsymbol{\theta})\rangle = \sum_{\boldsymbol{z}} a_{\boldsymbol{z}}(\boldsymbol{\theta})|\boldsymbol{z}\rangle$ or to each set of measured bitstrings $P^{\psi}(\boldsymbol{\theta})$. Presence of bitstrings close to MIS configurations should lower the cost so as to incite the optimiser to work in this direction at the next iteration. A straightforward procedure consists in only looking at the proportion of MISs in the measured distribution (and flip it to lower the cost), i.e. $1 - P(MIS)$. This can only be done if the size of the MIS has been classically obtained beforehand, either with an exhaustive approach doable for $N \lesssim 50$ or with a reliable classical optimisation procedure for larger sizes. However, as the Hilbert space increases exponentially with N, the probability of obtaining a non-zero projection to the MIS subspace approaches 0, especially for small n_{shots}. It is thus preferable to use a more continuous metric

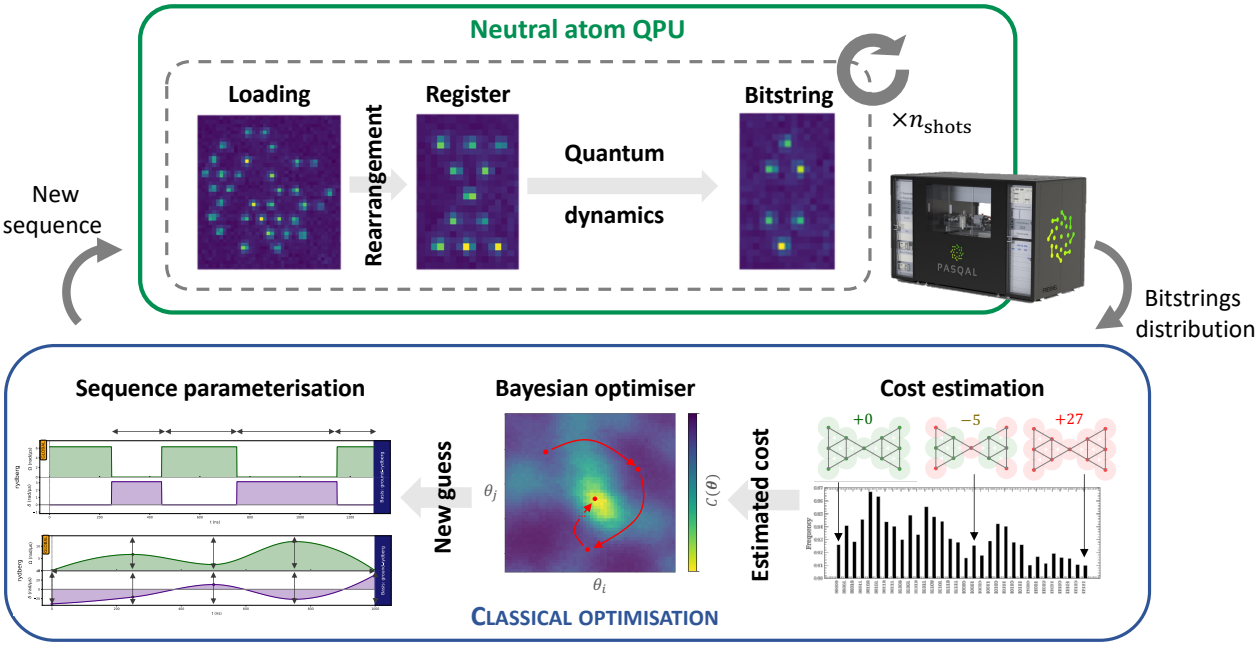


Figure 6.4: **Variational quantum algorithm loop**. A classical optimisation part is composed of a cost estimation method, favouring MIS configurations, a decision making Bayesian optimiser, navigating the parameter space with an exploitation-exploration strategy, and a sequence parameterisation step, building from a parameter instance a driving protocol sent to a neutral atom QPU. In the latter, atoms are loaded, arranged into positions to reproduce the graph considered and the quantum system evolves according to the built driving protocol. Measuring the system provides one bitstring and a distribution is acquired by repeating the quantum procedure several times before being sent to the classical part. This hybrid loop runs until the VQA converges.

in the sense that large IS should also lower the overall cost function. A fairly often used metric is the *approximation ratio* defined as :

$$R_{\mathcal{G}}(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\hat{C}_{\mathcal{G}}|\psi(\boldsymbol{\theta})\rangle = \sum_{\boldsymbol{z}} |a_{\boldsymbol{z}}(\boldsymbol{\theta})|^2 C_{\mathcal{G}}(\boldsymbol{z}) \tag{6.10}$$

with associated normalised cost $C(\boldsymbol{\theta}) = 1 + R_{\mathcal{G}}(\boldsymbol{\theta})/S_{\mathcal{G}}$ when the MIS size of the graph $S_{\mathcal{G}}$ is known. This cost helps discriminate between two distributions without MISs and its asymmetry favours IS configurations over non IS ones, as removing a node from a MIS (making it an IS) increases the cost by 1 but adding one (making it a non IS) increases it by at least $c-1$. Except for its normalised version, it does not require to know the MIS size beforehand. The cost function can be further multiplied by an additional function $g(\boldsymbol{\theta}) \geq 1$ encompassing other objectives to fulfil such as time optimality. The computed cost $C(\boldsymbol{\theta}^{(k)})$ is then passed to a classical optimisation routine for further processing.

Classical optimiser: the Bayesian algorithm

C can be seen as a noisy and expensive-to-evaluate black-box function. The Bayesian optimisation method outlined in Sec. 3.3 proves particularly advantageous in this specific scenario, thanks to its resilience against noisy evaluations, its efficiency even within a limited budget of iterations, and its gradient-free approach. Having previously provided the parameter instance $\boldsymbol{\theta}^{(k)}$, it can update its model with the computed value of $C(\boldsymbol{\theta}^{(k)})$. Balancing exploration of unknown regions of the landscape and exploitation of promising known regions, it then provides a new guess of optimal parameter $\boldsymbol{\theta}^{(k+1)}$ to try on the QPU. The values of this guess define a control protocol depending on the chosen parameterisation.

Controls parameterisation: layered and adiabatic protocols

The state evolution obtained when driving a N-qubit system under a parameterised control protocol can be written as :

$$|\psi(\boldsymbol{\theta})\rangle = \mathcal{T}\left[\exp\left(-i\int_{s=0}^{t_{\boldsymbol{\theta}}}\left(\frac{\Omega_{\boldsymbol{\theta}}(s)}{2}\sum_{j=1}^{N}\hat{\sigma}_j^x - \delta_{\boldsymbol{\theta}}(s)\sum_{i=j}^{N}\hat{n}_j + U\hat{H}_{\mathcal{G}}\right)ds\right)\right]|0\rangle^{\otimes N} \tag{6.11}$$

with the evolution operator obtained from Eq. 1.28 by parameterising the control fields Ω and δ (not using the phase ϕ) of Eq. 1.27. Several parameterisations of the control fields have already been mentioned in chapter 3 and we focus here on a layered approach, as well as a smoothly interpolated one which constitutes the building block of the variational quantum adiabatic algorithm used in Sec. 6.2.3.

Quantum Approximate Optimisation Algorithm (QAOA) approach

In this formalism, specific to the Rydberg setup as explained in Box. 22, $\boldsymbol{\theta}$ effectively encompasses the duration of each of the p control layers. The layer l includes a period of free evolution with detuning δ and duration t^l_{cost} followed by a resonant pulse with amplitude Ω and duration t^l_{mix}. The sequence is initialised with a $\pi/2$ resonant pulse which, without interactions, would prepare the system in $|+\rangle^{\otimes N}$ but here, only spreads the state as much as possible. Implementing such a control sequence with Pulser can be done similarly to Sec. 2.1.2 using EOMs as shaping devices. While this mode enables to produce sharp rising pulses, it also fixes a degree of freedom as the value of δ in between pulses fully depends on the value of Ω due to the three-to-two levels approximation (see Sec. 1.2.2). Indeed, compensating for the change in light shifts due to turning off one of the two excitation lasers is challenging at this timescale. Nevertheless, this "feature" enables to have a fast varying detuning between 0 and a fixed value. The theoretical performance of a perfect QAOA with Rydberg atoms has been extensively studied, notably by tackling industrial use cases [248] similar to the ones studied in Sec. 6.4. A few experimental implementations of QAOA were performed [249] with the notable use of more than 280 atoms in [250]. For long enough evolution times, the authors experimentally identified a nearly quadratic quantum speedup with respect to classical simulated annealing on hard graphs. An identified limitation of the QAOA approach is the need of increased p to further improve a score, leading to longer sequences, as the layers keep piling up, and to longer convergence, as the parameter space expands. One interesting idea to circumvent this issue consists in making an educated guess from previous layers to the new one [251], hence limiting the size of the parameter space to optimise on.

Variational Quantum Adiabatic Algorithm (VQAA)

This formalism builds on the adiabatic theorem introduced in Sec. 4.2.1 stating that with a slow enough time evolution, one can transfer a state from the ground state of an initial Hamiltonian to the one of a final Hamiltonian as pictured in Fig. 6.2(b). Quantum adiabatic algorithms (QAA) [252] were developed based on this principle in order to solve optimisation problems with a quantum computer. In this specific version, we parameterise a QAA so as to produce a VQAA tailored to solve MIS problems. $\boldsymbol{\theta}$ encompasses the values of the two control fields at fixed points in time, i.e. $\{\Omega_{\boldsymbol{\theta}}(t_i), \delta_{\boldsymbol{\theta}}(t_i)\}_{i=1\cdots m}$. These points can be evenly spaced, slicing the time interval regularly, they can be focused in regions where one needs more flexibility in the evolution, or they can even be set as optimisation parameters. In addition, the total pulse duration T can also be used as a tunable parameter and included in $\boldsymbol{\theta}$. The resulting parameter space of dimension $3(m+2)$ can be shrunk to a $(2m+3)$-dimensional one by fixing $t_i = T(i-1)/(m-1)$ and $\Omega_{\boldsymbol{\theta}}(0) = \Omega_{\boldsymbol{\theta}}(T) = 0$. The control fields are finally derived using interpolation by monotonic cubic splines which ensures that intermediate values will not cross the control bounds. Since these sequences are being produced by the AOM shaping devices, they are subjected to its finite bandwidth and deformation effect.

Examples of sequence obtained with the two parameterisations are displayed in Fig. 6.4. In both cases, the parameter space is bounded by hardware-related constraints such as the maximum coherence time of the machine or the available laser power. In the following, we study these two approaches on a single graph instance composed of 6 atoms arranged on a triangular layout, displayed at the top of Fig. 6.5(a).

6.2.2 - Standard QAOA-like approach

We apply the QAOA-like approach as a way to benchmark both the ability of the QPU to perform quantum dynamics under such protocols and the ability for the Bayesian optimiser to find optimal driving parameters even in the presence of noise.

Dynamics benchmark and noise mitigation

We emulate an optimisation procedure for increasing values of $p = 1, 2, 4$ over a parameter space $\Theta = [0.1\mu\text{s}, 1\mu\text{s}]^{2p}$, using 100 iterations (10 of them being reserved for initialisation) and $n_{\text{shots}} = 64$ shots per iteration. The normalised approximation ratio $1 - R = 1 - (R_{\mathcal{G}}/(-4))$ is used as cost function. We obtain three optimised protocols displayed in Fig. 6.5(a), yielding $1 - R = 0.45,\ 0.16$ and 0.27 respectively. Increasing p from 1 to 2 improves the score but the limited budget of iterations implies that at larger depth, like with $p = 4$, the optimiser will likely not converge, resulting in poorer performance as shown in Fig. 6.5(b,c).

We test these three protocols on Fresnel and probe the dynamics of $1-R$ at intermediate points in time, comparing it with a noiseless simulation using Pulser. As displayed in Fig. 6.5(b), in both cases, $1 - R$ is modified after each layer, ultimately reaching a minimum value at $T = \pi/2/\Omega + \sum_l (t^l_{\text{mix}} + t^l_{\text{cost}})$. However, a large discrepancy between simulated and experimental data remains. Adding detection errors to the simulation to match the experimental data reveals that the approximation ratio is especially sensitive to ε due to the asymmetry already mentioned. As more population is transferred to MIS states, the number of measurements violating the Rydberg blockade constraint is increasing as non zero ε produces non IS bitstrings. With detection errors included, the final values of $1 - R$ increase up to ≈ 1, close to its initial value, making the idea of directly optimising on the

Box 22: Quantum Approximate Optimisation Algorithm with Rydberg atoms

The Quantum Approximate Optimisation Algorithm (QAOA) [253] is a prominent method within the quantum optimisation paradigm [251, 254], inspired by the Trotterised version of adiabatic evolution. In QAOA, the quantum unitary evolution operator is parameterised by two sets of p angles, p being the *depth* of the algorithm, and alternates between two types of operations: mixing and problem-specific evolution.

Hardware-agnostic version: The algorithm applies p successive layers of two distinct Hamiltonians: a cost operator $\hat{C}$ and a mixing operator $\hat{M}$. Mathematically, this can be represented as:

$$|\psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \prod_{j=1}^{p} \hat{U}_M(\beta_j)\hat{U}_C(\gamma_j)\,|\psi_0\rangle \tag{6.12}$$

$\hat{U}_X(\theta) = e^{-i\theta\hat{X}}$ for $X \in \{M, C\}$ denotes the unitary operators corresponding to mixing and cost evolution, respectively. The algorithm initialises by preparing the system in $|\psi_0\rangle = |+\rangle^{\otimes N} = 2^{-N/2}\sum_{\boldsymbol{z}\in\mathbb{B}^N}|\boldsymbol{z}\rangle$, representing the ground state of the mixing Hamiltonian, often chosen as $\hat{M} = \sum_i \hat{\sigma}_i^x$. By evolving from this state, and with the aid of a classical optimiser, it becomes possible, as $p \to \infty$, to determine optimal angles $\boldsymbol{\gamma}^{opt}$ and $\boldsymbol{\beta}^{opt}$ such that:

$$|\psi(\boldsymbol{\gamma}^{opt},\boldsymbol{\beta}^{opt})\rangle = \mathrm{argmin}_{\psi}\,\langle\psi|\,\hat{C}\,|\psi\rangle \tag{6.13}$$

This approach can be viewed as a quantum version of annealing using quantum tunnelling to locate the minimum of a landscape as described in Fig. 6.2(a).

Rydberg version: In the Rydberg version, the cost operator $\hat{C}$ can be replicated for specific graphs and the mixing part can be tackled with resonant pulses of amplitude Ω. A major limitation preventing a straightforward implementation of QAOA on a Rydberg setup lies in the inability to turn off the interaction component of the Rydberg Hamiltonian during the mixing phase of each layer. This also implies that a clean preparation of $|\psi_0\rangle$ might require a more complex protocol or be even not possible. Applying pulses with $\Omega \gg U$ enables to neglect the interaction effects but the maximum amplitude reachable on current devices is only around $\Omega_{\max}/2\pi \sim$ few MHz. Consequently, this requirement mandates low values of U and hence of δ, which in turn prolongs the sequence and renders the evolution vulnerable to decoherence. One can nonetheless program an evolution with a QAOA-like protocol, applying series of resonant pulses with fixed Ω interleaved by periods of free evolution under fixed detuning δ. Each layer duration constitutes a parameter to optimise on.

QPU challenging.

A first solution consists in correcting for the detection errors using the scheme described in Box. 11, at the expense of possible additional computing costs. With $\varepsilon = 3\%$ and $\varepsilon = 8\%$, the corrected data are more in line with noiseless simulation and the remaining gap between the two at longer times can be attributed to decoherence effects. It is worth noting that imperfect correction of the measured distributions can nonetheless happen due to finite sampling effects and wrong estimation of ε and ε'.

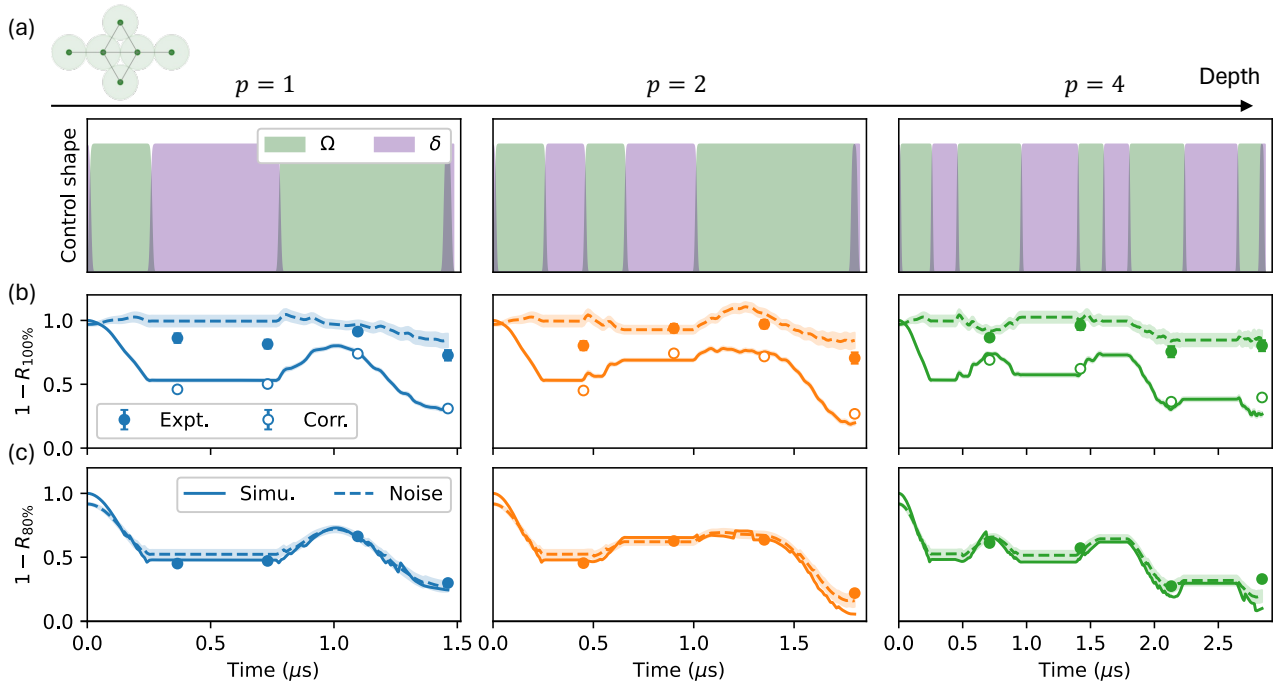


Figure 6.5: **Benchmark runs of the quantum dynamics with QAOA-like approach**. (a) Sequences parameterised with QAOA-like approach are applied on the atomic register given in the inset. For increasing depth p, the control shapes of Ω and δ are displayed as sent to the atoms, i.e. distorted by the shaping device. The evolution of (b) the normalised approximation ratio $1 - R_{100\%}$ and (c) the normalised truncated approximation ratio $1 - R_{80\%}$ during the dynamics is obtained using noiseless emulation (solid line) or emulation with shot noise ($n_{\text{shots}} = 1000$) and detection errors ($\varepsilon = 3\%$ and $\varepsilon' = 8\%$) (dashed). This enables to benchmark raw experimental measurements (filled dots) and SPAM corrected ones (white dots). The standard deviation (filled area-error bars) over the finite sampling with detection errors is obtained using the Jackknife resampling method.

A second possibility to make the approximation ratio more resilient to detection errors is to discard all the bitstrings violating the Rydberg blockade condition, assuming that their presence can only be due to noise. In the hard blockade limit where $U \gg \delta, \Omega$, the resulting discarding slightly modify the initial distribution and is an efficient way to mitigate ε. However in the regime of control parameters of this implementation, we can not rule out that the dynamics itself, through facilitation mechanisms for instance, could not produce bitstrings with unwanted excitations. Discarding bitstrings with the previous rule may alter significantly the optimisation landscape as some measured distributions could be almost emptied. Another straightforward approach consists in erasing a fixed percentage of each measured distribution, discarding bitstrings $\boldsymbol{z}$ among the ones with the highest energies $C_{\mathcal{G}}(\boldsymbol{z})$. This percentage can either be estimated knowing the value of ε or set arbitrarily high, at the expense of maybe discarding bitstrings not produced by the imperfect detection events. As N increases, the effect of detection errors intensifies and a large discarding percentage might be needed, at the expense of acquiring more statistics to balance the loss. The effect of this truncation is shown in Figure 6.5(c) with $R_{80\%}$ being the approximation ratio computed when keeping only the 80% best bitstrings of

the distributions. Both noisy emulated and experimental refined data present good agreement with noiseless emulated one when discarding 20% at $N = 6$. Overall, at this time scale, the main effect impacting the MIS preparation remains the detection errors and possible miscalibration of the controls, as emphasised in the next part. However, by optimising directly on the QPU, one can still locate an optimum as long as the landscape is not too flattened.

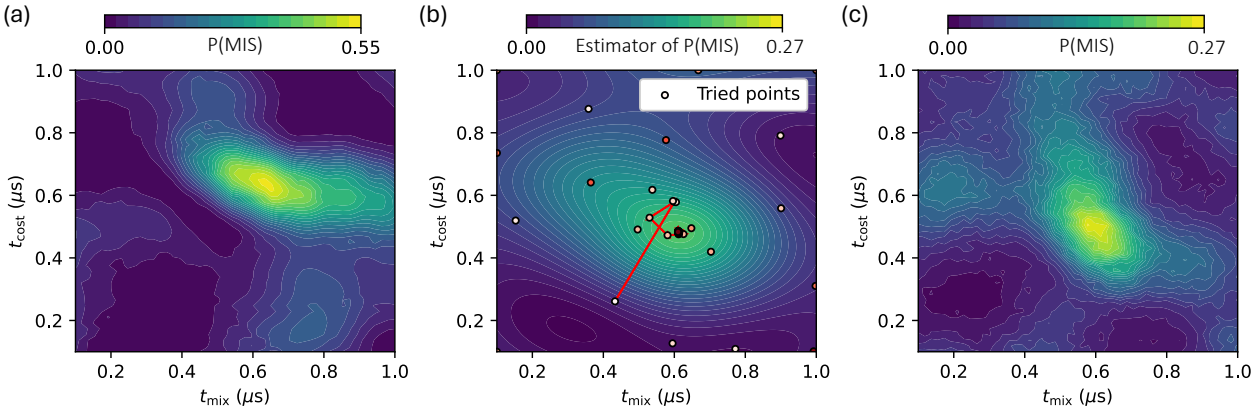


Figure 6.6: **Landscape of MIS probability for QAOA with depth 1.** (a) Varying the two parameters t_{mix} and t_{cost} of a QAOA-like sequence with $p = 1$, the noiseless landscape browsed when optimising P(MIS) can be retrieved. (b) The Bayesian optimiser constructs an estimator of P(MIS) by strategically sampling the parameter space (dots) with the help of the QPU. The redder the dot, the later it has been sampled. The minimum is updated (red line) each time a better one is found. (c) Adding detection and calibration errors in the emulation moves the global minimum location in the parameter space and explains the model built by the optimiser when using the noisy QPU.

Navigating a noisy landscape

We first perform a closed loop with $p = 1$ on the $N = 6$ graph using the MIS probability as a figure of merit in this case. The landscape of optimisation in the noiseless case is displayed in Fig. 6.6(a) with a unique global optimum located at $t^*_{\text{mix}} = t^*_{\text{cost}} = 0.633$ μs. We explore this landscape with the Fresnel QPU using a Bayesian optimiser provided with $10 + 50$ iterations with 256 shots each. The optimiser starts by using its initialisation budget to randomly sample the parameter space in an efficient covering as shown by clearer points in Fig. 6.6(b). Building a model to estimate the MIS probability at unexplored points, it then exploits the central region (green), sometimes exploring a bit outside of it (red points) to reduce its uncertainty and ultimately converging towards a believed optimum $t^*_{\text{mix}} = 0.611$ μs and $t^*_{\text{cost}} = 0.483$ μs (darker red points).

Surprisingly, the found optimum and constructed estimator do not coincide with the noiseless case. We checked that detection errors alone do not move the optimum, revealing that a miscalibration of parameters is involved. Indeed, when measuring Ω and δ with independent Rabi and Ramsey protocols, we obtain $\Omega/2\pi = 1.08$ MHz and $\delta/2\pi = -0.775$ MHz, instead of 1 Mhz and -0.5 Mhz, respectively. Fig. 6.6(c) exhibits the noisy landscape computed with detection errors and those miscalibrations, highlighting a displacement of the optimum location and a reduced optimal value of MIS probability

in accordance with the results found on the QPU. In the presence of limited noise, the Bayesian optimiser can still navigate the modified landscape and helps construct an optimised protocol to sample MIS configurations. This is promising as noise levels on QPU are, although not negligible, often maintained below known thresholds. However, a downside of such an optimised protocol is that it becomes useless once the QPU is recalibrated. Another optimisation procedure needs therefore to be computed, increasing the total cost. Consequently, this stresses the significance of building a VQA able to generate protocols with greater noise resilience and *generalisability*.

6.2.3 - Designing a VQAA with pulse shaping optimisation

We now switch to the second formalism, utilising a smoother parameterisation of the pulses to work with adiabatic protocols.

Avoiding diabatic errors with optimised drive

Similarly to the approach used in Sec. 4.2.4, we seek to find an optimised path adiabatically driving a system towards the ground state of $\hat{C}_{\mathcal{G}}$. For the same graph with $N = 6$ atoms, we diagonalise the Rydberg Hamiltonian over $\Omega/U \in [0, 1.5]$ and $\delta/U \in [-1, 3]$ and plot in Fig. 6.7(a) the MIS probability obtained when sampling the ground state. A lobe-shape portion of the phase diagram (yellow) constrained between $\delta/U = 0$ and $\delta/U = 3$ exhibits MIS probabilities higher than 95%. We emulate an optimisation procedure with smooth pulse shaping using a periodic temporal slicing with $m = 4$ over the parameter space $\Theta = [0.5\mu\text{s}, T] \times [0, 1.5U]^m \times [-2U, 3U]^{m+2}$ and use $10 + 90$ iterations with perfect access to the states produced. We set a maximum duration of $T = 3$ μs for the protocols explored by the optimiser, compelling it to manipulate the shapes of Ω and δ to ensure adiabaticity within these constraints.

With the optimised path (red) displayed in Fig. 6.7(a), we obtain a MIS probability of 99.98% after 3μs. While the optimiser has hit the temporal bound, hinting at the fact that longer protocols could be even more successful, both amplitude and detuning were used with parsimony, without any explosion of the parameters. Although the path may initially appear convoluted, Fig. 6.7(b) illustrates that it effectively circumvents regions characterised by vanishing energy gaps, i.e. $\Omega/U \approx 0$ and $\delta/U \approx 0/3$. The gap Δ is defined between the MIS manifold of states (here only including the ground state since the graph has only one MIS) and the first above excited state. The parameter T dictates the minimum acceptable gap value (here around $0.7U$) to maintain adiabaticity, as excessively small values might induce undesired population transfers to excited states. The protocol is not perfectly adiabatic as shown by the instantaneous energy spectrum of Fig. 6.7(c). While most of the population remains in the ground state, transitioning from the *IS phase* to the *MIS phase*, a bit of population leaks to the first excited state around 1 μs and between 2.5 μs and T, when the gap is minimum. Overall, the Bayesian optimiser constitutes an effective pathfinder, avoiding the areas most likely to cause diabatic errors.

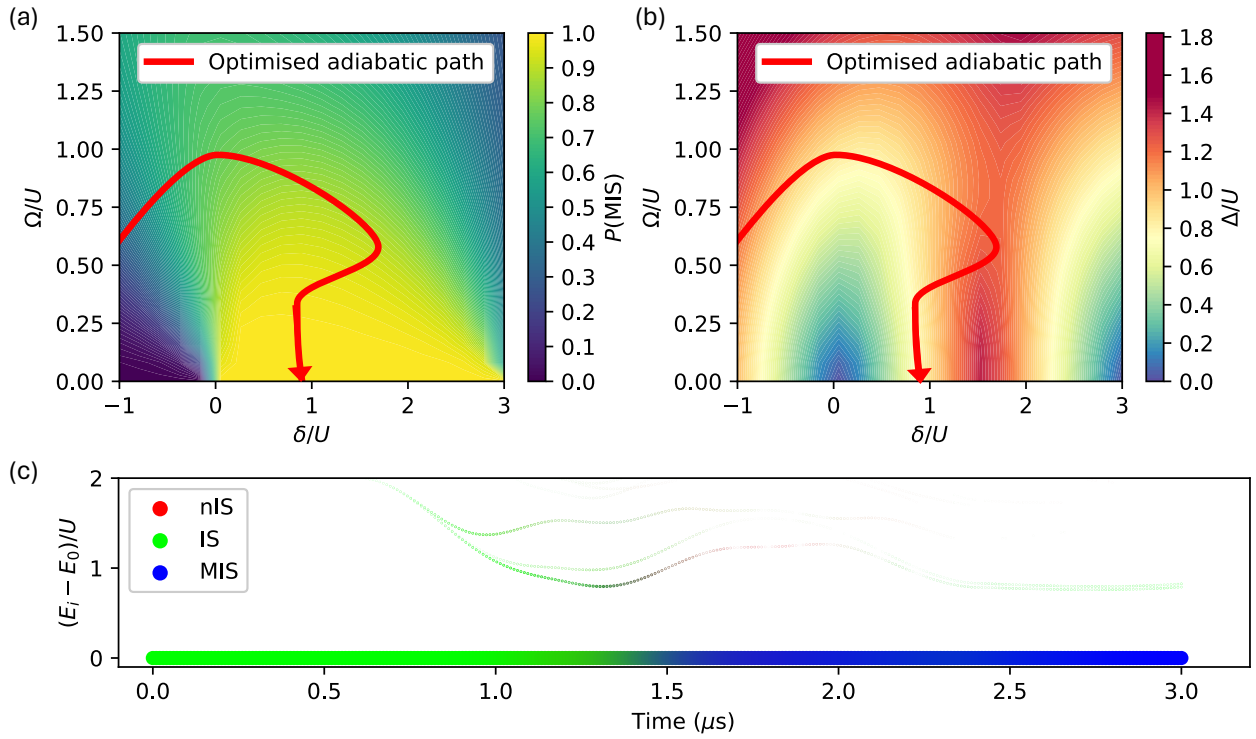


Figure 6.7: **Optimised adiabatic driving to the MIS phase.** (a) Phase diagram of the MIS probability obtained by computing the ground state of the Rydberg Hamiltonian with constant parameters Ω, δ and U. The path taken by the optimised protocol (red) is constrained by $\Omega/U \leq 1.5$ and $|\delta|/U \leq 3$. (b) Similar as (a) for the gap Δ between ground state and first excited state. The optimised protocol avoids regions of vanishing gaps (dark blue). (c) Instantaneous spectrum of the system during the optimised dynamic with levels coloured with their proportion of non IS (red), IS (green) and MIS (blue) configurations and thickened by population (levels are only appearing when populated).

Noise resilience of VQAA

In order to assess the resilience of such almost adiabatic protocols, we conduct a noise analysis detailed in Fig. 6.8. The protocol appears resilient to parameter miscalibrations, such as rescaling of Ω and shifts of δ, as depicted in Fig. 6.8(a). In fact, even with similar miscalibrations observed in the QAOA case ($\tilde{\Omega}/\Omega = 1.08$ and $(\tilde{\delta} - \delta)/U = 0.03$), the obtained MIS probability would only decrease from 99.98% to 99.95%, hinting at minimal deviation from the optimum found. The influence of T is also indicated in Fig. 6.8(b) where noiseless simulation exhibits a significant diabatic drop below 1 μs but rather flat behaviour otherwise. Adding the detection errors only rescales the curve by a factor $\approx 71\%$. However, when considering decoherence with an effective model of relaxation T_1 and dephasing T_2, lengthening the duration becomes detrimental and an optimal duration starts to appear for this rather small instance. At $T = 3$ μs, the MIS probability is already reduced by a factor 2. For larger graphs, this balanced optimum between adiabaticity and decoherence should move towards longer times, thereby reducing the optimum value of P(MIS) that can be achieved. Despite all this,

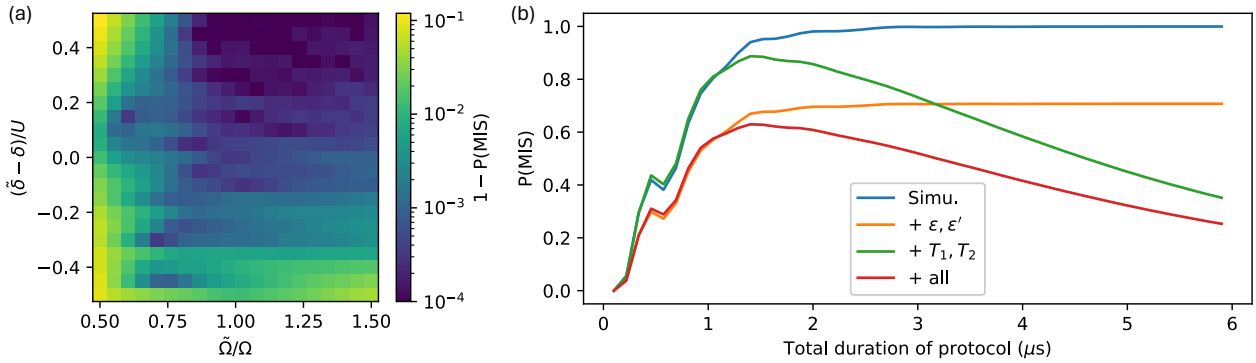


Figure 6.8: **Noise analysis of an adiabatic protocol.** (a) Colour map of 1−P(MIS) when rescaling the global amplitude drive Ω to $\tilde{\Omega}$ and shifting the detuning δ to $\tilde{\delta}$. (b) Emulated evolution of the MIS probability when stretching in time an adiabatic protocol. In the noiseless case (blue), only diabatic errors impact the result at short times. Detection errors $\varepsilon = 1\%$ and $\varepsilon' = 8\%$ (orange) rescales the behaviour while decoherence $T_1 = 100$ μs and $T_2 = 4.5$ μs (green) impacts the probability for longer sequences. With this noise model (red), an optimal duration can be derived, here around 1.4 μs to balance adiabaticity and decoherence.

the VQAA approach remains more resilient than the QAOA-like and, as explained in the next section, more generalisable.

6.3 - Generalisation of optimised protocols: QAOA vs. VQAA

The previously introduced VQAs provide a means to derive an optimised protocol for preparing a quantum state tailored to a specific graph, which when sampled, yields MIS configurations of the said graph. Nevertheless, they are quite resource consuming, requiring dozens if not hundreds of iterations to converge. An appealing idea explored in this section involves constructing a general protocol capable of preparing MIS states for a family of structurally similar graphs, with specific training on only a portion of the family, for instance the smallest graphs.

6.3.1 - Cost landscapes of various geometries

A family of graphs with similar structure can be generated by randomly sampling a lattice with regular geometry, using random walks. Given a periodic trap layout, we can select subsets of adjacent traps with or without cycles and therefore obtain similar looking graphs. Examples of such families are given in the first column of Fig. 6.9 using triangular geometry, square geometry or a combination of the two, using a periodic Shastry-Sutherland lattice, well-known in condensed-matter physics [255]. For each layout, we generate a family of 30 graphs of size ranging from $N = 5$ to 10 and apply to each graph either a QAOA-like or a QAA-like sequence. Each protocol, showcased on the first line of Fig. 6.9, is parameterised by two parameters (red) and we obtain the corresponding landscapes of $C(\boldsymbol{\theta})$ by sweeping over their possible values within the bounds $\Theta = [0.1, 0.5]^2$ for QAOA and $\Theta = [0, U]^2$ for the QAA. Fig. 6.9 depicts the cost landscapes associated with the normalised approximation ratio, averaged over the graphs of each family. Additionally, the location of the global minimum for each

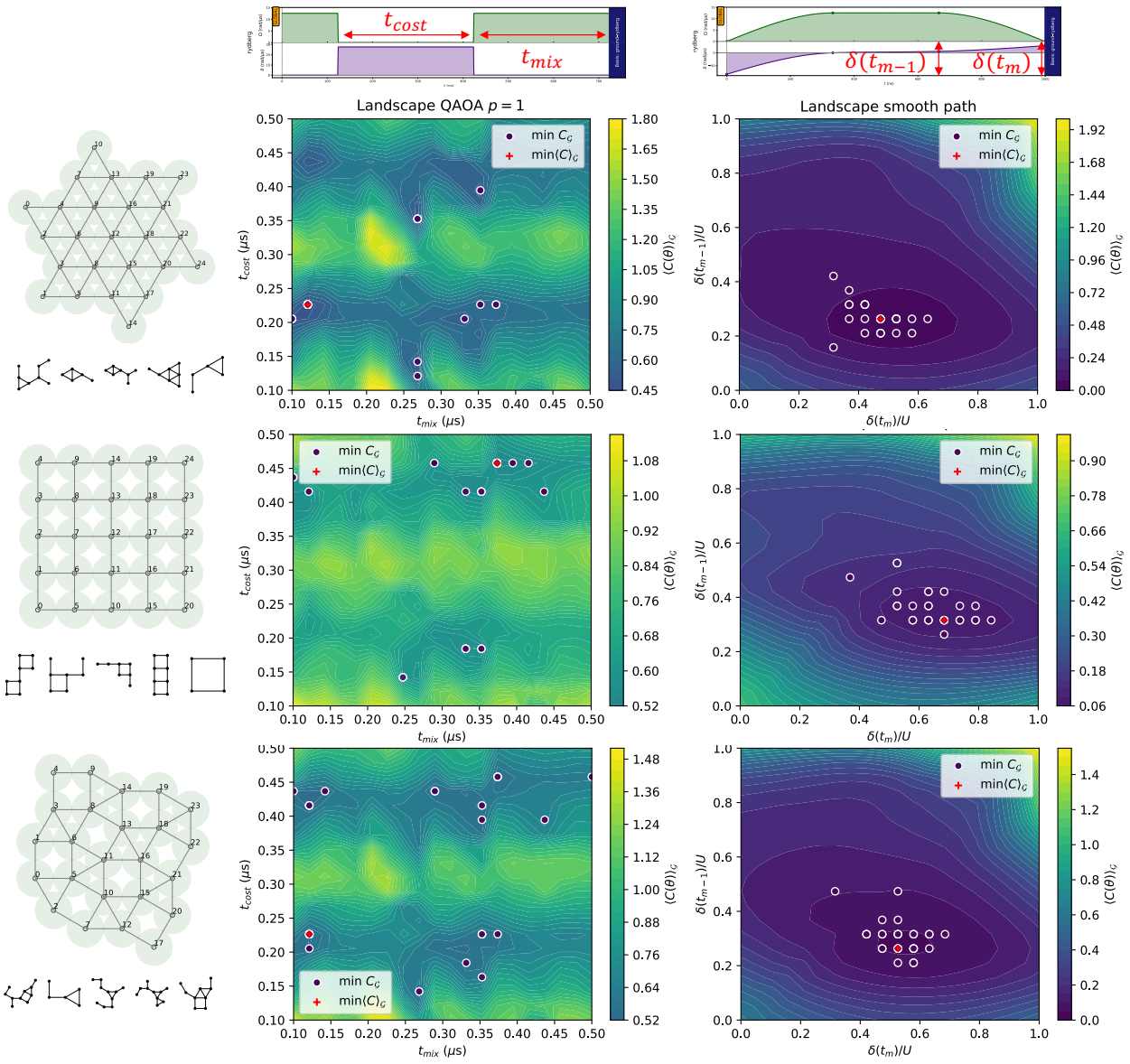


Figure 6.9: **Study of generalisability of VQAs over graphs generated from regular lattices.** Datasets of graphs generated from either triangular, square or Shastry-Sutherland lattices (layouts plotted on the left) are subjected to QAOA-like or QAA-like protocols parameterised by two parameters (Pulser sequences plotted above): either the two durations of a standard depth 1 QAOA protocol or the last two detuning values of a standard QAA protocol. The corresponding colour maps are obtained by averaging over the dataset the landscapes obtained when computing the normalised approximation ratio $1 + R_\mathcal{G}(\boldsymbol{\theta})/S_\mathcal{G}$ for each graph over the parameter space. The minimum found for the averaged landscape is plotted (red cross) as well as the minimum found for each graph over this space (dots). Regions with lower cost are displayed in darker colours.

graph is also indicated to highlight possible grouping effects. In the QAOA-like case, despite the

value of the global minimum for each graph all being close to 0, the lack of similarities among the distinct landscapes creates numerous local minima on the averaged landscapes. For none of the three families does the averaged landscape exhibit a global minimum (red cross) close to 0 in value. For the three geometries considered, the individual minimum only partially group together, invalidating the hypothesis of a single QAOA-like protocol enabling preparation of MIS over a graph family.

On the contrary, for the QAA-like approach, the individual landscapes overlap to create a global minimum for the averaged case, closely surrounded by the clustered individual global minimum locations. For the triangular and Shastry-Sutherland layouts, the global minimum of the averaged landscape reaches 0 whereas for the square case, it only approaches 0.06. Two elements can contribute to explain this discrepancy : firstly, some graphs in this family may not be as similar as ones in the triangular family, which is supported by the existence of some individual minima located far from each other. Secondly, in the square case, the NNN interactions are often only reduced by a factor 1/8 compared to the NN ones, making such graphs more prone to have imperfect embeddings. If $\hat{H}_{\mathcal{G}}$ is not well reproduced, a QAA protocol will only yield a finite efficiency. This effect seems reduced when combining square and triangular geometries. In addition, we can also see that variations of the parameters will only poorly affect the averaged preparation as the landscapes of QAA remains quite flat around the global minima. Finally, it is interesting to notice that the optimal locations varies between the geometries with $\delta(t_m = T)$ increasing from triangular to square, passing by Shastry-Sutherland. This gives an interesting insight on the values around which the MIS phases of the various geometries are centred. Those observations motivate the choice of the VQAA approach to find a general optimal protocol for graphs with triangular structure.

6.3.2 - General protocol for triangular graphs

The procedure pursued here involves utilising the VQAA on a family of triangular graphs, split into training and testing sets. The goal is to efficiently derive a generalised pulse capable of preparing MIS configurations for all graphs within the family, including those for which the protocol has not been specifically optimised on. This strategy is driven by the considerable expense associated with running a VQA for multiple graphs on a neutral atom QPU, given the current sampling rate. By simulating the protocol optimisation on the training set, composed of small-sized graphs, we keep the classical computation cost low. Then, we can directly employ on the QPU the optimised protocol for the test set, composed of unseen larger graphs, mitigating both the classical and quantum computation cost.

The optimisation is performed over 50 graphs, 10 of each size ranging from 5 to 9, including instances with and without cycles. The cost $C(\boldsymbol{\theta})$ is obtained summing the average $\langle\ \rangle_{\mathcal{G}}$ of the MIS probability obtained for each graph and the associated standard deviation $\sigma_{\mathcal{G}}$. Adding the standard deviation in the cost function helps the optimiser to favour protocols with uniform scores across the family over ones working very well for specific instances but less generalisable. Fig. 6.10(a) showcases the convergence over $25 + 175$ iterations for both a $4-$depth QAOA and a VQAA with $m = 3$ and $T_{\max} = 4\ \mu$s. As expected, even a deeper QAOA fails to find a working general protocol and remains stuck at $\langle \mathrm{P(MIS)} \rangle_{\mathcal{G}} \approx 0$. The VQAA locates an interesting region with cost below 0.5% after 100 iterations and ultimately converges towards $\langle \mathrm{P(MIS)} \rangle_{\mathcal{G}} \approx 99.73\%$, where all the graphs have a MIS probability above 98.5%. We can check that running separate optimisation routines for each graph gives very similar average convergence both in terms of speed and quality of preparation but with final scores more spread. The optimiser history can be boiled down to three main steps : before the

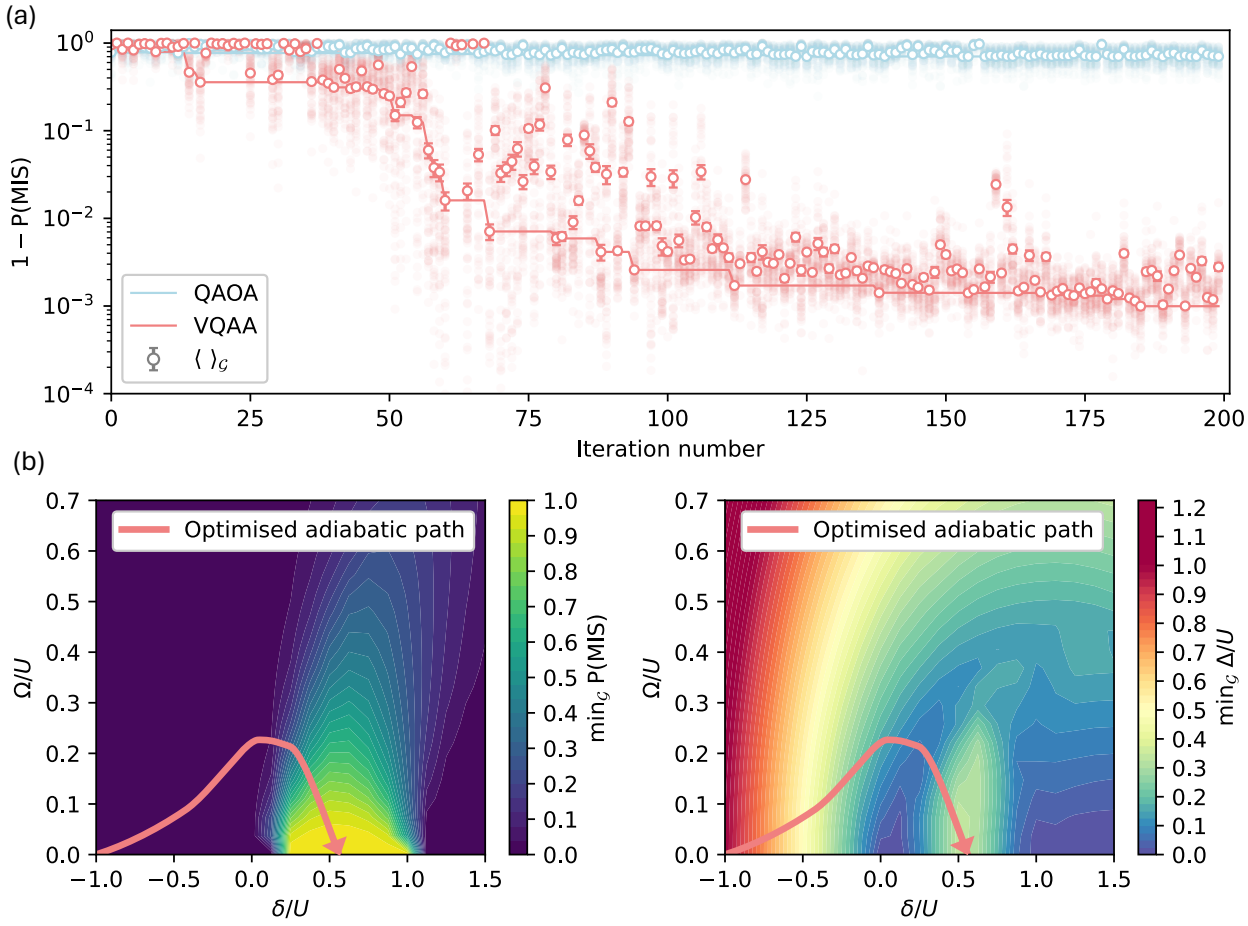


Figure 6.10: **Optimisation scheme to devise a generalisable adiabatic protocol for triangular graphs.** (a) Convergence of the optimiser when searching for regions of parameters minimising $1 - \langle \mathrm{P(MIS)} \rangle_\mathcal{G}$ (dots) using a QAOA-like parameterisation (light blue) or a QAA parameterisation (red). The cost obtained for each graph at each iteration is also plotted in transparent. (b) Similar maps as in Fig. 6.7(a,b) where the minimal P(MIS) and gap calculated across the dataset of graphs are plotted. The optimised path obtained for the QAA parameterisation (red) finds the MIS phase (yellow) and avoids regions with vanishing gaps (dark blue).

50^{th} iteration, it tries to work on all the graphs at the same time, while between 50 and 125, it allows itself to try pulses working very well, and even better than the final one, only a few graphs while the score of the others can be worse than before. Finally, after iteration 125, it only refines the averaged cost by a few tens of %.

The optimised path given in Fig. 6.10(b) starts deep in the IS phase and ends in the middle of the MIS phase where the MIS probability of the ground state goes to 1 for all graphs. We have constrained the optimiser to use hardware realistic bounds similar to the used for Fig. 4.3, explaining the shorter path taken compared to Fig. 6.7(a). When taking the minimal value over the family of graphs, the MIS phase is narrower, i.e. $0.2 \leq \delta/U \leq 1$. Additionally, the optimiser identifies the optimal path

while adhering to both hardware constraints and the adiabatic criteria, ensuring avoidance of regions with low energy gaps Δ, for instance around $(\Omega, \delta)/U \approx (0, 0)$. Δ is defined here as the maximum between the gaps within the MIS manifold and the gap between MIS manifold and the first excited state above it. The protocol seems to work well on the training set and can now be tested on larger graphs exhibiting similar triangular structure.

6.3.3 - Scaling of MIS probability with the graph size using TN emulation

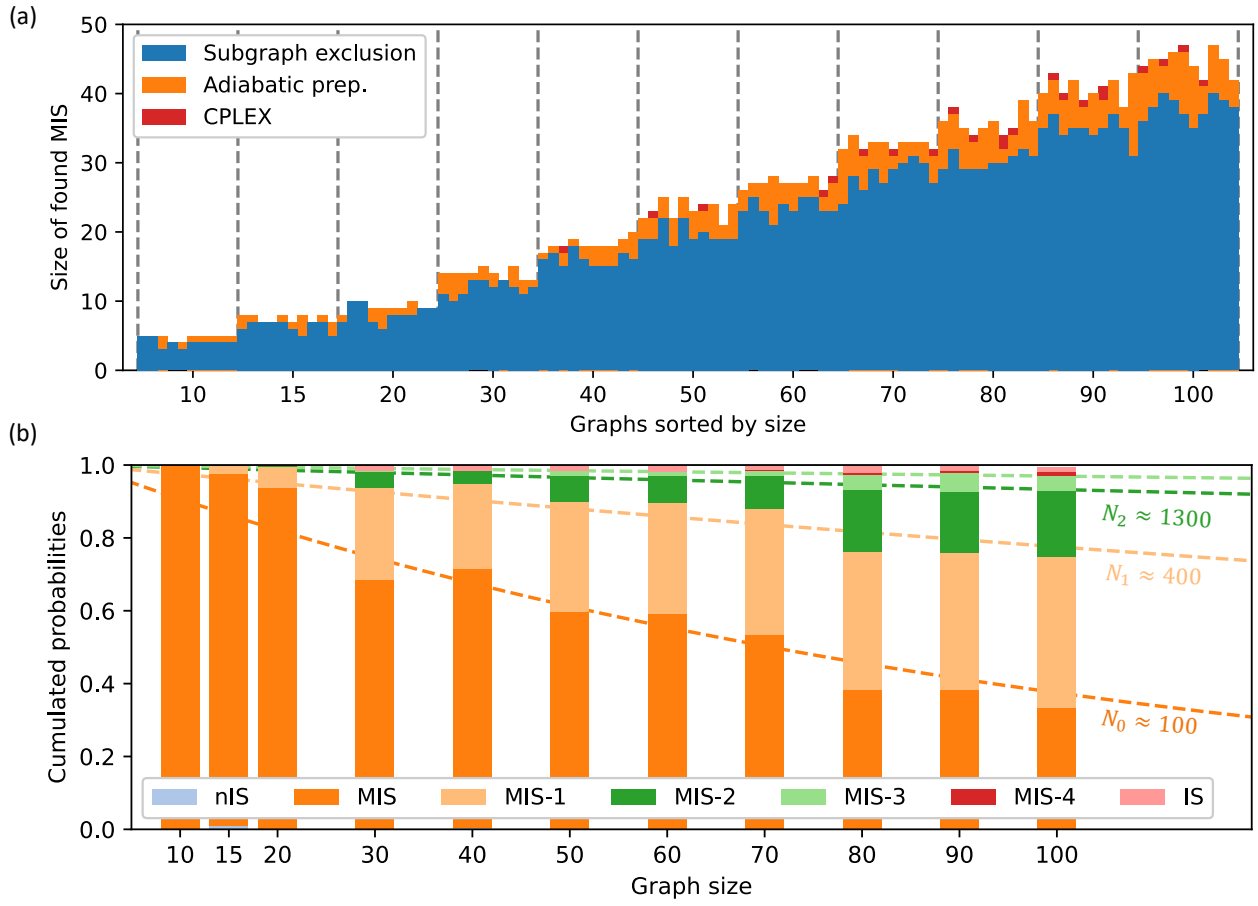


Figure 6.11: **Scaling of the adiabatic preparation efficiency with the system size.** (a) For graphs with sizes ranging from 10 to 100, the size of the MIS is computed using classical methods, like subgraph exclusion (blue) and CPLEX (red) or quantum methods (orange). (b) Cumulative probabilities of finding non IS, IS, MIS-4, ..., MIS configurations after the adiabatic preparation, averaged over graphs of the same size. The decline of $\sum_{i<k}$P(MIS$-i$) with the graph size N is fitted by an exponential decay with exponent $1/N_k$ (dashed line) for $k = 0, 1$ and 2.

The test set is comprised of 110 graphs, 10 of each size ranging from 10 up to 100. Emulating the dynamics over such large systems requires using tensor network methods described in Sec. 2.2. Another scaling problem to deal with is the classical computation of the MIS size for each graph $S_\mathcal{G}$.

For sizes smaller than 50, a classical computer can still perform an exhaustive search over the space of configurations, but for larger instances, such a method is not an option. We therefore rely on two classical algorithms : an approximate algorithm using the subgraph exclusion (SE) method [256] implemented in the NetworkX python package and the state-of-the-art CPLEX algorithm from IBM. CPLEX employs a variety of advanced optimisation techniques, including linear programming, branch-and-bound, cutting planes, and heuristics, to efficiently solve large-scale optimisation problems in various domains such as supply chain management, resource allocation in logistics and decision-making processes in finance. We apply the optimised protocol obtained in the previous section to all the graphs of the test set and we check the size of the maximal IS found for each of them. Fig. 6.11(a) displays the found MIS size with graphs sorted by size for the two classical approaches and the quantum one. For all graphs, we have $S_{\mathcal{G}}^{\mathrm{SE}} \leq S_{\mathcal{G}}^{\mathrm{adia.}} \leq S_{\mathcal{G}}^{\mathrm{CPLEX}}$. The SE method (blue) proves to be unreliable as it can not even reach the MIS sizes found by an exhaustive search at graph size ≤ 50 and is always matched or exceeded by the general QAA approach (orange). The latter, when compared to the sizes found by CPLEX (red) in less than 10 ms, seems to give the right MIS size in 82% of the cases, with more errors as the graph size increases. Keep in mind that no optimisation was performed on those graphs as only the generalised adiabatic drive of Fig. 6.10 was applied.

In order to further quantify the quality of the states prepared at large sizes, we study the distributions of probabilities of the configurations for each graph. Fig. 6.11(b) depicts the average cumulative probabilities for each type of configuration, i.e. non IS, IS (with the largest ones being labelled as MIS$-k$, meaning they are missing k elements to be MIS) and MIS. It is worth noting that even in the case where MISs were not found, a large portion of the distributions is still composed of MIS-1 and MIS-2 which is a promising feature for including such a QAA approach in a hybrid algorithm. Indeed, a good enough solution can then be classically enhanced as checking if configurations of MIS$-1+1$ (MIS$-2+2$) are in fact MISs only scales with N (N^2). As the graph size increases, the probabilities summed over the MIS$-k$ decline and can be naively fitted with an exponential decay $\exp[-N/N_k]$ with $N_0 \approx 100$, $N_1 \approx 400$, $N_2 \approx 1300$, etc. Following this fit, we can extrapolate that measuring a MIS with $F = 99\%$ probability at size $N = 1000$ would require around $n_{\mathrm{shots}} > \log(1-F)/\log(1-\exp(-N/N_0)) \approx 10^5$ shots but measuring a MIS or a MIS-1 would only need around 50 shots with this method. This whole study was carried out without taking noise into account and many effects, not necessarily related to the adiabatic driving itself, could impact the scaling when implemented on a noisy QPU.

6.4 - Using adiabatic quantum computing for an industrial use case

We will now utilise the adiabatic protocol optimised above to solve with a QPU an industrial use case provided by the French electric utility company EDF and similar to the one tackled in [248].

6.4.1 - Modelling a smart charging task as optimal scheduling of load time intervals

Smart charging is an umbrella term encompassing all optimisation problems related to electric vehicle (EV) charging. The recent increase in the number of EVs presents both new challenges and opportunities for electricity management. Issues like charging task allocation, scheduling, and cost optimisation emerge due to the significant charging times of EVs and their unpredictable load on the electrical system [257]. Vehicle batteries can serve as energy storage and power supply devices, a technique known as vehicle-to-grid (V2G), which significantly enhances the flexibility of the electric system and reduces the reliance on fossil fuels during high-peak demand [258].

Charging an EV takes a finite duration and its starting time can be fixed by the users of the vehicles, imposed by technical constraints or decided by the smart-charging manager. Given a set of charging/load tasks, one can represent them as intervals on a timeline as shown in Fig. 6.1(c), such that each of them belongs to a specific group, for example distinct vehicle fleets of a company. One needs to select a subset of these tasks to maximise the number of non-overlapping ones, such that at most one task in each group is completed. Therefore, the completion time of the selected loads will be minimised and no group will be over-represented in the schedule. This problem belongs to the class of *Group Interval Planning* (GIP) problems [259]. Adding fake empty loads enables to consider instances where all groups contain the same number k of tasks. Such a problem is NP-hard for $k \geq 3$ [260]. We can easily map such a problem to a MIS one by building the graph where a node i corresponds to an interval and i and j are connected if their corresponding intervals overlap in time or if they belong to the same group. Therefore, solving a GIP problem modelling a smart charging task can be tackled by a MIS-solving QAA algorithm if one can map the instances to implementable graphs.

The graphs used originate from a data set of 2250 loads performed during May 2017 on identical charging points of the Belib's network of load stations located in Paris [261]. GIP instances can be derived by randomly sampling the loads and we limit this work to specific instances that can be formulated as MIS on two-dimensional UD graphs.

6.4.2 - Embedding of GIP graphs into structurally similar atomic registers

Embedding UD graphs into a plane can be achieved in a relative straightforward way using *force-directed graph drawing* methods such as the Fruchterman-Reingold algorithm [262]. The latter aims at positioning nodes such that connected pairs of nodes are placed close together, while unconnected pairs of nodes are positioned farther apart. The algorithm simulates a physical system where nodes are represented as particles that repel each other due to an electrical charge and are connected by springs. The optimal positions, iteratively obtained, correspond to the equilibrium state reached by the system.

However, the additional terms of Eq. 6.7 appearing due to the $1/r^6$ scaling of the vdW interactions can cause unwanted dynamics if the mapping obtained with this method is not cleaned. Formally, for a UD graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the interaction energy $U^{\mathcal{E}}_{\min} = \min\{U_{ij}, (i,j) \in \mathcal{E}\}$ should be an order of magnitude stronger than $U^{\bar{\mathcal{E}}}_{\max} = \max\{U_{ij}, (i,j) \notin \mathcal{E}\}$ so as not to have unconnected pairs of atoms exerting comparable blockade as connected pairs.

Improved free-space mapping of the UD graphs

We develop a heuristic that locally modifies the positions of the vertices of a UD graph, starting from an initial guess. The latter can be obtained with a Fruchterman-Reingold algorithm [262] or with a manifold learning algorithm as described in Box. 27. Because UD graphs are local, we can first devise a procedure that finds optimal positions for sub-groups of vertices, before concatenating the positions together. In this procedure, we keep track of the vertices for which we have already found optimal positions. Then, we select a vertex i and its neighbourhood $\mathcal{N}(i)$ and build the subgraph $\tilde{\mathcal{G}}$ induced by $\mathcal{N}(i)$. We then minimise the ratio $U^{\bar{\mathcal{E}}}_{\max}/U^{\mathcal{E}}_{\min}$ on this subgraph only using Sequential Least Squares Programming [263]. Hardware constraints such as the limited field of view can also be enforced by slightly penalising atoms outside of the system bounds or pairs of atoms too far from each other. We add a contribution of the variance of the distance across all connected pairs to ensure that the significant terms of the interaction matrix are all of the same order and the atoms dynamics do not happen at too different timescales. The result of such optimisation is shown in Fig. 6.12(a). The positions can be globally rescaled so as to respect the hardware constraints on minimum inter atomic distance.

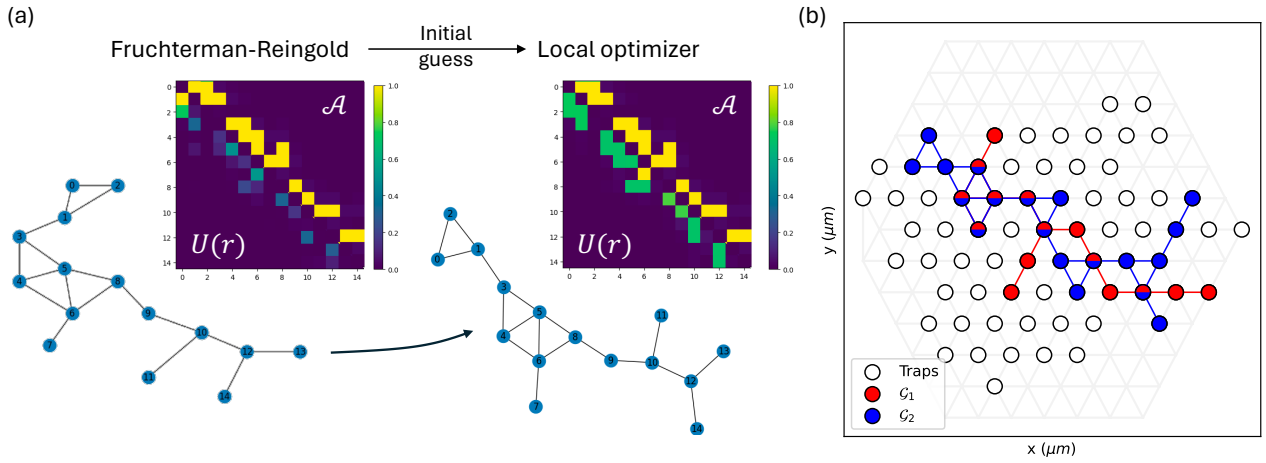


Figure 6.12: **Mapping and batching techniques used to embed datasets of UD graphs.** (a) An initial mapping using the Fruchterman-Reingold algorithm can try to reproduce the adjacency matrix $\mathcal{A}$ of a graph $\mathcal{G}$ (upper corner of the matrix) with the atomic register interactions (lower corner of the matrix). However, the scale differences between representative terms can lead to poor embeddings. A local optimiser improves the initial guess and reduces the scale difference between the terms. (b) Illustration of the batching method where two graphs $\mathcal{G}_1$ (red) and $\mathcal{G}_2$ (blue) are mapped to an underlying triangular lattice (grey). By sharing common traps, a single SLM pattern (white circles) can be used for several atomic registers.

Batching SLM patterns for time saving

The optimised free-space positions calculated above usually define the locations of the filled traps in a trap layout. The additional traps required to get enough trapped atoms from the initial random filling of the layout can be put randomly around them, or in a designated area. For each graph of the dataset, an SLM pattern is then derived and the atomic register can be constructed in the device. In practice however, the SLM calibration step can be quite time-consuming, i.e. of the order of a few minutes, getting close in duration to a few hundreds of shots. In the interests of speeding up the experimental implementation, we therefore seek to *batch* several graphs onto the same SLM pattern, reducing the number of calibrations needed for the whole dataset.

The batching method first implies clustering the dataset of graphs according to similarities in their structures. By identifying identical structures across graphs, such as hexagons or crosses, we can separate graphs into several families and build a dedicated trap pattern for each. The similarity metric used for the clustering is a linear combination over the types of recurring cycles in the dataset of the following quantity

$$s(\mathcal{G}_1, \mathcal{G}_2) = 1 - \exp(-\alpha|c_1 - c_2|), \tag{6.14}$$

where $c_{1/2}$ represents the number of a given cycle in $\mathcal{G}_{1/2}$ and α is a hyper-parameter. For instance, the graph displayed in Fig. 6.12(a) has $c = 4$ 3-node cycles. The similarity matrix between all graphs of the dataset is fed to a *k-means clustering* algorithm [264]. Furthermore, since the laser power is distributed over all the traps, we want to reduce the total number of traps, in order to maximise the intensity provided to each trap. This ensures that the traps are deep enough to obtain a satisfying filling efficiency ($p_{\text{fill}} \approx 55\%$) over the whole pattern. Once we have clustered the graphs into different batches, we embed them onto an underling grid structure. For each batch, we thus select a periodic lattice including the types of cycle appearing in its graphs and sequentially select sites to include in the pattern, conserving the pairwise distance between free-space mapped nodes of a graph and favouring the use of already selected sites. Once all graphs have been embedded, the obtained layouts are shifted and rotated to further try to maximise the occupation of the same traps. Finally, if only M sites have been selected and the largest graph has more than $p_{\text{fill}}M$ nodes, random adjacent sites are added to the pattern to ensure proper initial filling.

In the case of the smart-charging dataset, we select a batch consisting of 33 graphs, ranging in size from 9 to 23 nodes. The mapping-and-batching procedure outputs a unique trap layout using 61 traps. An example of overlapping graphs is given in Fig. 6.12(b).

6.4.3 - Experimental implementation and results

We embed the graphs into atomic registers on Fresnel as shown by the fluorescence pictures of Fig. 6.13(a). Applying the adiabatic protocol obtained above and imaging again, we can directly retrieve MIS configurations composed of the missing atoms between the two pictures. We acquire around $n_{\text{shots}} \approx 1000 \times (p_{\text{move}})^{23} = 1000 \times 80\% = 800$ shots for each graph. On average, in noiseless emulation, we obtain $\langle \text{P(MIS)} \rangle_{\mathcal{G}} = 98.3(4)\%$ while in experiment, we get $29(2)\%$. The quality of the distributions, either emulated or experimentally obtained, can be summarised by looking at the evolution of the truncated approximation ratio introduced in Sec. 6.2.2.

Fig. 6.13(b) displays how the normalised approximation ratio approaches 1 as bitstrings with high cost are discarded more and more. For instance, the averaged cost obtained in noiseless simulation

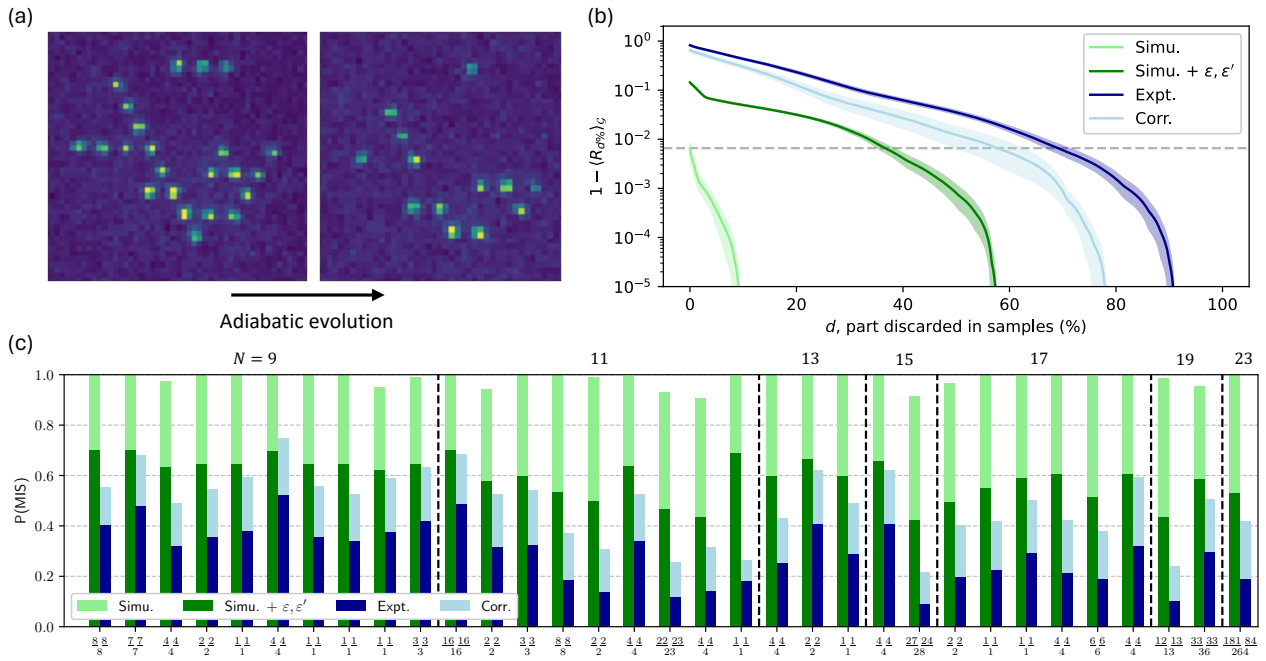


Figure 6.13: **Experimental results when applying the generalisable QAA protocol to smart charging use case.** (a) Fluorescence pictures of the atomic register taken before and after the adiabatic dynamics. Missing atoms forms a MIS. (b) Evolution of the approximation ratio, averaged over the dataset, when discarding the d% worst bitstrings of each distribution. Mean (solid line) and standard deviation (transparent filled areas) are plotted for (light green) noiseless emulation, (green) noisy emulation, experimentally obtained (blue) raw and (light blue) corrected data. (c) Histograms of the MIS probabilities obtained for each graph. The dataset is sorted and separated (dashed black vertical lines) according to the graphs size. On the x axis, the number of different MIS configurations found is given in the format $\frac{\text{Simu. Expt.}}{\text{Exact}}$.

(light green) is around 1% but discarding the 10 worst percent of the distributions makes it plummet to 0, highlighting that the minimum MIS probability reached across the dataset is around 90%. We can also quantify how much the detection errors $\varepsilon = 1\%$ and $\varepsilon' = 8\%$ (green) modify the scores by checking the proportion that needs to be discarded to retrieve the original score (following the grey dashed line), here about 38%. The detection errors are not the only contributors to the discrepancies between emulated and experimental data as correcting for those does not reproduce the noiseless behaviour. Indeed, from Fig. 6.8(b), we expect from a 4 μs adiabatic pulse a drop in MIS probability of at least 60% due to both imperfect detection and decoherence. The experimental curve (dark blue) drops at large d emphasises that the worst preparation over the dataset still includes around 10% of MISs in its distributions. Fig. 6.13(c) gives a more detailed picture of the experimental results with the smaller graphs ($N = 9$) exhibiting between 35% and 50% of MIS configurations while few graphs at sizes 11, 15 and 19 only have around 10% of MISs. Correlating these significant drops with the number of MISs of the graph and the efficiency of preparation in the noiseless case for instance remains challenging as various experimental parameters may also have influenced the results. Spatial

inhomogeneities of the control fields could result in distinct protocols being applied across positions, even though the devised protocol seemed resilient to control fluctuations. Overall, for almost all graphs, except the ones with the largest MIS subspaces, such as the 23−nodes instance with 264 possible MIS configurations, all the MISs were found at least once. On average, the first MIS was found after a few shots/seconds, two orders of magnitude longer than with the 10 ms of CPLEX. A gain in speed/performance from the quantum method is not to be expected at those sizes. The threshold at which the tables could turn, inducing a so called quantum advantage, is hard to predict both because of the cost of simulating large quantum systems, even with tensor networks, and because constant improvements of classical algorithms makes it a moving target to reach.

6.4.4 - Conclusion : UD-MIS seems not so hard

Notably, the triangular UD graphs of this dataset are all *chordal*, meaning that every induced cycle in them has exactly three vertices. For this specific family, there exist classical solvers able to find a MIS in polynomial time [265]. Keep in mind that we could devise the same method of generalised adiabatic driving for other families of graphs, exhibiting longer or more complex cycles, and apply it to a batch of graphs with such structure obtained from the smart-charging dataset. However, the locality of UD-MIS makes it relatively easy to approximate classically, and there even exists a so-called polynomial-time approximation scheme that guarantees a $1-\epsilon$ approximation ratio found in polynomial time with the instance size [266], as already noticed in early publications [247, 267]. Extrapolating from realistic coherence levels and current repetition rate of neutral atom devices, the authors of [267] estimated that a quantum algorithm exploiting this locality aspect could potentially surpass their classical benchmark with 8000 atoms for a time budget of 2 seconds, or with $1000-1200$ atoms for a time budget of 0.2 seconds. Authors of [268] introduced an additional local Hamiltonian to the annealing scheme to achieve a quadratic speed-up over effective classical techniques. Recent works assess the use of quantum dynamics in excited regimes to solve combinatorial problems, using for example many-body localisation [269] or quantum critical dynamics [270]. Specifically, quantum quenches with neutral atoms have been explored to circumvent super-exponential closing gaps in annealing schemes [271].

6.5 - Beyond the UD-MIS problem

The existence of efficient classical approximation schemes for UD-MIS problems motivated efforts towards quantum procedures capable of tackling more complex groups of problems such as UD-MWIS, where nodes are marked with distinct weights. Those weights represent an additional degree of complexity that can be tuned to make the instances harder to solve. Employing site-dependent detuning as explained in Box. 12, facilitates the encoding of solutions to the UD-MWIS problem in the Rydberg Hamiltonian. However, only relying on a global driving Ω proves inadequate in this context. Overcoming this limitation demands a more extensive set of operations, prompting the implementation of advanced sequences with local controls.

An alternative route involves the exploration of atomic embeddings of graphs that go beyond the UD property. It has indeed been observed that classical solvers exhibit a time-to-solution order of magnitude higher when applied to instances with larger connectivity or less local structure [272]. In particular, the incorporation of *ancillary* vertices, coupled with the ability to construct atomic arrays

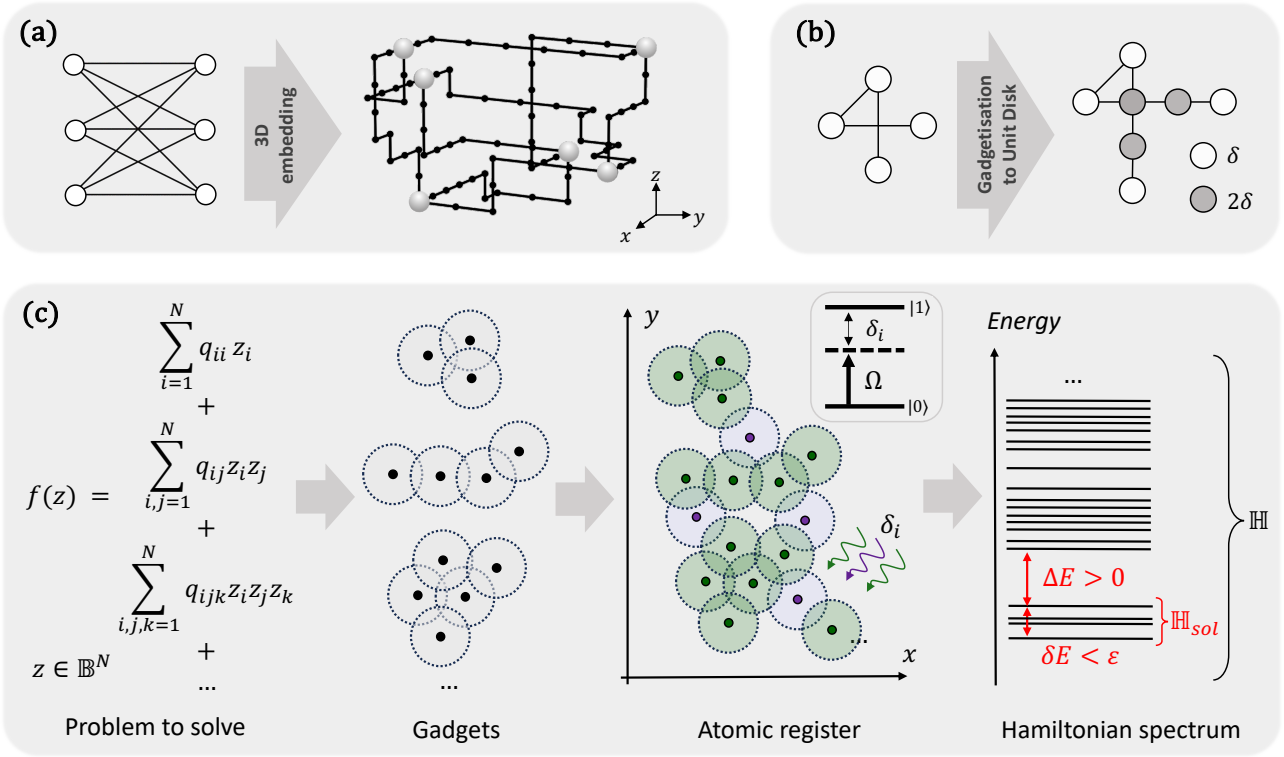


Figure 6.14: **Going beyond UD using local addressing and ancillary atoms.** (a) Embedding a non-planar graph in 3D using ancillary chains of atoms. (b) A 2D-gadget implementing a crossing edge using additional local detunings (grey nodes). (c) Framework of geometric programming: a general unconstrained binary optimisation function f is translated into a sum of terms that can each be encoded in an arrangement of M atoms and local detunings $\{\delta_i, i \in 1, \dots, M\}$. Then, one can assemble these *gadgets* together using predefined rules: this usually requires additional atoms (green). By construction, the Hamiltonian spans a Hilbert space where the low-energy bulk $\mathbb{H}_{\text{sol}}$ encodes the solutions to the QUBO in an $\varepsilon-$energy band. The excited states are gaped from the solutions by an energy difference $\Delta E > 0$.

in 3D, facilitated the examination of the MIS problem on various platonic [273] and non-planar, high-degree graphs [249]. The key idea driving these developments is the utilisation of a chain of auxiliary qubits to mediate strong interactions between distant qubits, allowing for the effective tackling of complex non-UD graphs. A subsequent demonstration in Ref. [274] showcased a deterministic and polynomial approach capable of embedding any bounded degree graph. As an illustration, this method successfully addressed the MIS problem for the complete bipartite graph $K_{3,3}$ using a 3D arrangement of qubits shown in Fig. 6.14(a), in conjunction with local detunings. The ground state of this 3D graph, found using DMRG, can be viewed in Fig. 2.4(b). By discarding the ancillary vertices, one can check that this ground state is actually a MIS configuration of $K_{3,3}$.

Advancements in local controls and ancillary chains of atoms have spurred the utilisation of the native UD-MIS solving capabilities of neutral atoms to address optimisation problems in higher com-

plexity classes. Solutions to the original problem are mapped to the lower energy states of the Rydberg Hamiltonian, a process referred to as *geometric programming* on Rydberg platforms [275]. This involves finding a geometric arrangement of atoms and a set of laser controls to achieve the desired hierarchy of eigenstates. Various methods of geometric programming have been recently proposed, relying on *gadgets*—small atomic ensembles encoding parts of an optimisation problem. Assembling these gadgets using defined gluing rules constructs a faithful UD-MIS problem. For instance, Ref. [276] introduces copy and crossing gadgets that facilitate all-to-all connectivity on a crossing lattice, as illustrated in Fig. 6.14(b). This enables the encoding of any node-unweighted QUBO problem into a UD-MWIS problem, using at most $4N^2$ atoms and global detuning masks. Gadgets described in Ref. [277] specifically address the implementation of 3- and 4-body parity constraints in the parity encoding of connected higher-order binary optimisation (HOBO) problems, facilitating the encoding of any HOBO as a UD-MWIS.

Additional gadgetisation techniques, such as the QUBO-centric approach proposed in Ref. [278], offer robust solutions without requiring local fields. Triangular gadgets and anti-copy wires, introduced in Ref. [279], encode 3-Satisfiability (3-SAT) problems as UD-MIS problems, demonstrating experimental success on small systems. This toolbox has been experimentally applied to factorise small semi-prime numbers in [280]. Geometric programming techniques have also been developed for quantum simulation, expanding the range of algorithmic problems solvable by neutral atoms, as seen in [275] or [281]. The geometric programming formalism is underpinned by the functional completeness of PXP models, as demonstrated in [275]. The hard-blockade approximation, enabled by an embedding respecting $r_{nn} \ll r_b \ll r_{nnn}$, allows encoding solutions of various algorithmic problems into the bottom end of a Rydberg Hamiltonian spectrum with a finite gap and perfect degeneracy, as depicted on the right of Fig. 6.14(c). These techniques significantly broaden the spectrum of algorithmic problems solvable by neutral atoms, albeit at the cost of introducing auxiliary atoms to propagate Rydberg blockade and employing local controls to fix different maximum independent configurations at the required energy levels on the Hamiltonian spectrum.

7 - Quantum kernel for classification in Graph Machine Learning

Contents

Graphs are not only employed as representations of combinatorial problems but also as pivotal tools in machine learning (ML). Impactful applications abound, such as predicting potential edges in recommendation systems [282], identifying fraudulent activities in communication networks [283], and forecasting protein function [284]. While graphs offer a rich structure for manipulating complex data, the level of freedom they afford can prohibit straightforward analysis. For instance, graphs with identical adjacency matrices yet disparate node and edge labels constitute challenges in directly applying linear algebra techniques. It is therefore essential to create efficient ML models that correctly and effectively learn and extract information from graph structures. Using the exponentially large Hilbert space accessible to a quantum computer in order to generate new types of graph embeddings is an appealing idea, with many proposals and theoretical studies over the past few years [285–288].

In this chapter, after introducing the procedure followed by classical ML algorithms to perform classification tasks on graphs, we will show how a neutral-atom QPU can be seen as a learning model for graph-based data and incorporate the latter, as a quantum kernel, into a classification algorithm implemented experimentally on a real-world dataset of toxic molecules [289]. Finally, we will evaluate the potential advantage of our method by means of a novel metric that is sensitive to the similarity between the geometry of the feature spaces of two kernels [290].

The following chapter draws from the following article to which I contributed:

- [291] B. Albrecht et al., "Quantum feature maps for graph machine learning on a neutral atom quantum processor," Physical Review A, vol. 107, no. 4., 2023.

7.1 - Classification task in Graph ML

7.1.1 - Supervised learning

In supervised learning, a ML algorithm is provided a dataset of associated pairs of input-output, $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), \cdots, (\boldsymbol{x}_M, y_M)\} \subset (\mathcal{X} \times \mathcal{Y})^M$, and is tasked with predicting the corresponding output $y \in \mathcal{Y}$ of an unknown input $\boldsymbol{x} \in \mathcal{X}$. Each input $\boldsymbol{x}$ is characterised by $N = |\mathcal{X}|$ features, $x^{(1)}, \cdots, x^{(N)}$. Depending on whether $\mathcal{Y}$ is continuous or discrete, the prediction is referred to as either a *regression* or a *classification* task. Whereas in regression, y can take a range of possible values, in classification, the goal amounts to sort inputs into a finite set of categories. The algorithm is also given a parameterised family of models $\{f(\boldsymbol{x}; \boldsymbol{\theta}) : \mathcal{X} \times \Theta \to \mathcal{Y}\}$ from which it must choose the best mapping based on the patterns observed in the training data. This is achieved through *training*, i.e. the the optimisation of a model-dependent score function penalising discrepancies between the model guesses and observed data. The trained model is then trialled on unseen data points to assess its *generalisability*. A lack of training data or over-simplifying assumptions can cause a model to underfit, making it perform poorly on both training and test sets. On the contrary, optimising a highly complex model for too long can lead to overfitting of the data, taking, for example, random fluctuations in the training set for significant patterns. Such a model will score well on its training dataset, but will be of no help on new data, even if the latter displays easy-to-recognise patterns. To address those issues, *regularisation* terms, often labelled as λ, can be added to the loss function to penalises overly complex models and redundant features can be removed during pre-processing of the data. In addition, *cross-validation* procedure, in which $\mathcal{D}$ is split into multiple subsets for training and pre-testing, can help assess the generalisation performance of the model and identify overfitting.

7.1.2 - Classification task and performance metrics

We focus in this chapter on binary classification for which $\mathcal{Y} = \mathbb{B} = \{0, 1\}$, i.e. the inputs are labelled as positive or negative. As an example, let us take a recipe book where each instance is a cake recipe with features being, among others, k the amount of sugar and l, the baking time. Through trials-and-experiments, making each recipe once, one can assign a binary label, good or bad, to each recipe and thus collect a training dataset as displayed in Fig. 7.1(a). When experimenting with new ingredients or variations in a recipe, it becomes valuable to anticipate the success of these alterations with a reasonable level of accuracy.

A classification model is typically evaluated on its ability to correctly predict the class labels using the following metrics :

$$P = \frac{t_p}{t_p + f_p}, \quad R = \frac{t_p}{t_p + f_n}, \quad F_\beta = (1 + \beta^2)\frac{PR}{\beta^2 R + P} \tag{7.1}$$

Here, $t/f_{p/n}$ represents the proportion of true/false positives/negatives of the predicted distribution. For instance, a good recipe wrongly labelled as bad will increase the f_n proportion. The precision P is the ability of a classifier to not mistake a negative sample for a positive one; it thus represents the quality of a positive prediction made by the model. Similarly, the recall R can be understood as the

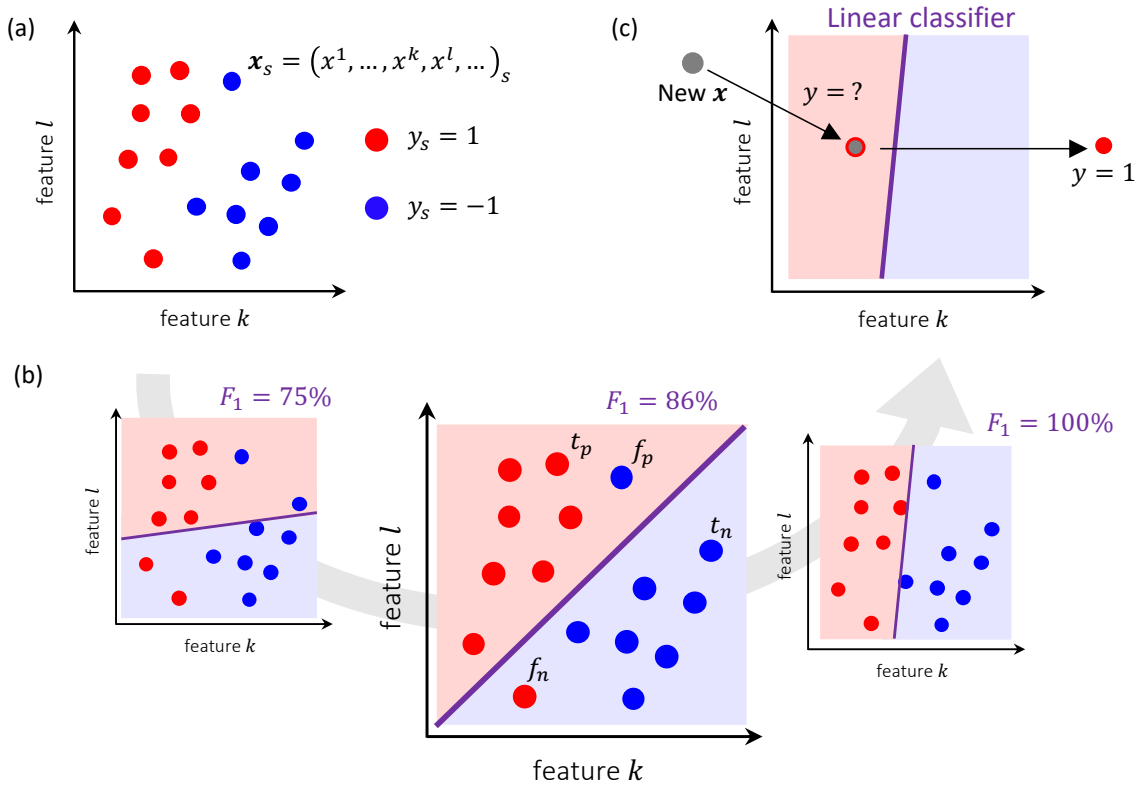


Figure 7.1: **Classification task in ML.** (a) A dataset of input-output pairs $(\boldsymbol{x}_i, y_i)$ is gathered : for each input $\boldsymbol{x}_i$, the binary output y_i is known. (b) A linear classifier is trained over the known dataset, trying to find the best way of separating the two classes. (c) The trained classifier can be used to label with known performance a new input $\boldsymbol{x}$. Its generalisability can be assessed by applying it to a test dataset, distinct from the training dataset used for its training.

ability of the model to find all the positive samples. F_β is a combination of the two and will be used in the following with $\beta = 1$.

The *linear discriminant* model described in Box. 23 is one of the simplest classifier to try. Training this linear classifier on the collected data, i.e. maximising F_1 in the case displayed in Fig. 7.1(b), enables to get a clear cut over the dataset. Fig. 7.1(c) reveals that while the baking time l is not a defining feature of good cakes, the sugar amount k, on the contrary, will only produce bad cakes if excessive. Keep in mind, that this conclusion is only valid for the dataset presented in this example. While this simplest model proves to be sufficiently appropriate for the dataset at hand, it may perform poorly on more complex collections, such as the non linearly separable distribution of Fig. 7.2 or ones containing graphs for instance. More advanced models such as Random Forest of decision trees or Gradient Boosted Machines aggregating weak learners can prove useful and will be used in the following chapter 8. Rather than increasing the complexity of the classification model itself, an alternative approach described below involves transforming the data by mapping it into a *feature space* where the distributions can exhibit different shapes.

Box 23: Linear regression models

Linear discriminant analysis consists in finding a linear combination $\boldsymbol{\theta}$ of features that will help to directly characterise classes or lead to problem dimensionality reduction by feature selection. Such a parameterised model can be represented by

$$f(\boldsymbol{x};\boldsymbol{\theta}) = \boldsymbol{\theta}^T\boldsymbol{x} + b \tag{7.2}$$

with the bias b taken as 0 for centred data. Finding the best set of parameters $\boldsymbol{\theta}$ is typically achieved by minimising the *empirical risk*

$$R(\boldsymbol{\theta}) = \frac{1}{M}\sum_{(\boldsymbol{x},y)\in\mathcal{D}} \mathcal{L}(f(\boldsymbol{x};\boldsymbol{\theta}), y) \tag{7.3}$$

with $\mathcal{L}(y, y')$ being a loss function characterising discrepancies between true labels and inferred ones such as $\mathbb{1}_{y\neq y'}$ or $||y-y'||^2/2$. This optimisation can be tackled by gradient-descent methods [292, 293] where $\boldsymbol{\theta}$ is updated until convergence with the following rule $\boldsymbol{\theta}' := \boldsymbol{\theta} - \gamma\partial R(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ with γ, a hyper parameter. A possible choice of R reads

$$R(\boldsymbol{\theta}) = \frac{1}{2M}||\boldsymbol{\theta}^T\boldsymbol{X} - \boldsymbol{y}||^2 + \frac{\lambda}{2}||\boldsymbol{\theta}||^2, \tag{7.4}$$

where $\boldsymbol{X}$ encompasses the inputs $\boldsymbol{x}_i$, $\boldsymbol{y}$ encompasses the labels y_i and λ is the hyper parameter controlling the ridge regularisation which prevents excessively high values of the parameters $\boldsymbol{\theta}$. This quadratic optimisation problem is known as the *primal* problem. In this particular least-squares minimisation, we can solve analytically the optimisation with

$$\frac{\partial R(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = 0 \Rightarrow \boldsymbol{\theta}^* = (\boldsymbol{X}^T\boldsymbol{X} + \lambda I)^{-1}\boldsymbol{X}^T\boldsymbol{y}. \tag{7.5}$$

but this requires the inversion of a matrix of size $M \times M$ which can quickly become resource consuming if the inputs are characterised by many features.

7.1.3 - From feature maps to kernels

In many classical machine learning methods, one seeks to map input data into a different space called the feature space, making it easier to work with, using a transform called the *feature map*. An example is shown in Fig. 7.2, where the distribution of the binary classes into concentric circles makes the dataset not linearly separable in $\mathbb{R}^2$ and thus not manageable with the simple linear discriminant approach described in Fig. 7.1. The feature map ϕ solves this issue by mapping data points from a two-dimensional plane into three-dimensional space, making the data linearly separable in $\mathbb{R}^3$. One of the drawbacks of this formalism, alongside finding the right expression for the feature map, is that the higher the dimension of the feature space used to integrate the data, the more resource-intensive the calculation of $\phi(\boldsymbol{x})$ is likely to become. Fortunately, rather than needing the absolute values of $\phi(\boldsymbol{x}_i)$, most ML algorithms merely require understanding the relative distance between mapped data points in this transformed space: this is referred to as the *kernel trick*.

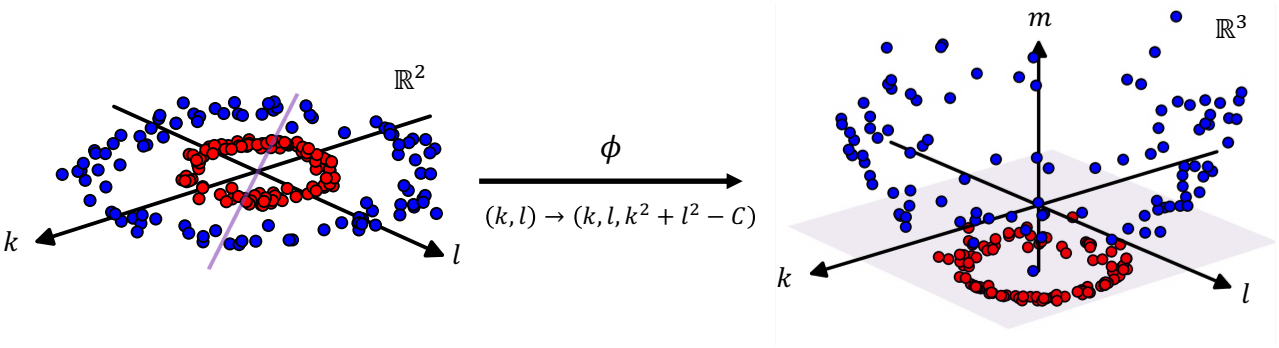


Figure 7.2: **Example of classical feature map.** We seek to find a binary classifier enabling the separation of the two data classes by a hyperplane (purple). After a transformation $\phi((k,l)) = (k,l,k^2+l^2-C)$ mapping the data point from $\mathbb{R}^2$ to $\mathbb{R}^3$, the transformed data points become linearly separable.

From the feature map ϕ we can define an associated kernel function $K : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}$ with $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle$. The kernel function K is a positive semi-definite bivariate function and serves as a similarity measure between data points. The interesting property of kernel functions is that they can be used directly in the linear regression model instead of the explicit feature vectors. Using a feature map ϕ, we can define a linear model on the mapped feature space by $f(\boldsymbol{x}; \boldsymbol{\theta}) = \boldsymbol{\theta}^T \phi(\boldsymbol{x})$. By replacing $\boldsymbol{X}$ with $\boldsymbol{\Phi}$, the matrix encompassing the mapped features $\phi(\boldsymbol{x}_i)$, in the equations of Box. 23, we can derive an expression of the trained model :

$$y(\boldsymbol{x}) = f(\boldsymbol{x}; \boldsymbol{\theta}^*) = K(\boldsymbol{x})^T (\boldsymbol{\Phi}\boldsymbol{\Phi}^T + \lambda I)^{-1} \boldsymbol{y}. \tag{7.6}$$

where $K(\boldsymbol{x}) = (K(\boldsymbol{x}, \boldsymbol{x}_1), \ldots, K(\boldsymbol{x}, \boldsymbol{x}_M))$. The symmetric matrix $\boldsymbol{\Phi}\boldsymbol{\Phi}^T$ can be written as a kernel matrix $\boldsymbol{K}$ since $(\boldsymbol{\Phi}\boldsymbol{\Phi}^T)_{ij} = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle = K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{K}_{ij}$. As previously emphasised, the solution to this linear regression problem can be solely expressed in terms of the kernel function K and no longer requires the expression of ϕ. This trick is especially utilised by Support Vector Machine (SVM) algorithms, which aim to delineate data points by identifying a separating hyperplane within the feature space as conveyed in Box. 24. In this space, the coordinates of each data point can be determined with respect to the kernel function K. Therefore, it is often interesting to directly come up with a kernel function rather than an explicit feature map. For example, the Radial basis function kernel, with hyper parameter $\gamma > 0$ can be written as

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = e^{-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2} = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle \text{ with } \phi^{(k)}(\boldsymbol{x}) = e^{-\gamma \boldsymbol{x}^2} \sqrt{\frac{(2\gamma)^k}{k!}} \boldsymbol{x}^k \tag{7.7}$$

where $\phi(\boldsymbol{x})$ is a feature vector of infinite size which would require truncation to be used in a ML numerical model. Feeding such as kernel to an SVM enables it to effectively handle non-linear relationships between data points without explicitly mapping them to a higher-dimensional space.

7.1.4 - Classical kernels for graphs

Kernels also offer the versatility to handle symbolic objects, extending beyond vectors to encompass more complex structures such as graphs. Graph embedding techniques [294], which refer to finding a representation of a graph or of its individual nodes in a vector space, serve as an imperative component in numerous GML algorithms. By finding node representatives which preserve different types of relational information from the graph, node embedding can be used for prediction tasks such as node classification [295] or link prediction [296]. Embedding can also be done at the graph-level to distinguish graphs of different nature. Notions of distances and similarities between the representative vectors can then be used to find the best boundary between datapoints with different labels in the context of supervised machine learning. This is the main idea behind the notion of *graph kernel*, which represents a measure of similarity between input graphs in the form of a scalar product between their representative vectors obtained either using a feature map or directly a kernel. A simple example of a graph kernel is the *size* kernel which for two graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, is defined as:

$$K_{\text{size}}(\mathcal{G}_1, \mathcal{G}_2) := e^{-\gamma(|\mathcal{V}_1| - |\mathcal{V}_2|)^2}, \tag{7.8}$$

with a choice of hyper parameter $\gamma > 0$. It considers two graphs to be similar solely based on their sizes, lacking in-depth analysis of their topological structures. Different graph kernels have been introduced to capture different aspects of graph structures and we describe several of them, including the SVM-θ, the Graphlet Sampling (GS), the Random Walk (RW) and the Shortest Path (SP) kernels, in Appendix A. The latter classical kernels will be used to benchmark the quantum kernel built in the next section.

7.2 - Quantum Evolution Kernel for graphs

ML algorithms are both versatile tools able to handle complex analysis of various types of data and expensive models requiring substantial computational resources. The Quantum Machine Learning (QML) field gathers diverse methodologies for infusing quantum principles into classical ML, with the aim of enriching classical techniques through the potential of the quantum information paradigm.

7.2.1 - QML: a matter of mixing

Classifying QML algorithms involves understanding where the quantum part is added. A noteworthy classification [297] focuses on the nature of both the data and the processing device, dividing QML into four sub-fields :

- **(CC)** Classical data analysed with classical devices, utilising quantum-inspired methods such as tensor networks (see Sec. 2) for neural network training [298] and surprisingly, dequantised quantum algorithms [299].
- **(QQ)** Quantum data analysed with quantum devices, envisioning tasks like quantum simulations preparing ground-states and analysing them with specific quantum methods [300]. However, this approach is seen as a long-term perspective due to challenges in storing and extracting quantum states.

Box 24: Support Vector Machine (SVM) algorithm

The SVM algorithm aims at splitting a dataset into two classes by finding the best hyperplane that separates the data points in the feature space, in which the coordinates of each data point (here each graph) is determined according to the kernel K.
Suppose we have access to a training graph dataset $\{\mathcal{G}_i\}_{i=1\dots M}$ with a set of labels $\boldsymbol{y} = \{y_i\}_{i=1\dots M}$ where $y_i = \pm 1$ depends on which class the graph $\mathcal{G}_i$ belongs to. Mathematically, the *dual* formulation of the SVM problem consists in finding $\boldsymbol{\alpha}^*$, the parameters in $\mathcal{A}_C(\mathbf{y}) = \left\{\boldsymbol{\alpha} \in [0, C]^M \middle| \boldsymbol{\alpha}^T \mathbf{y} = 0\right\}$, minimising

$$\frac{1}{2}\boldsymbol{\alpha}^T Q \boldsymbol{\alpha} - \mathbf{e}^T \boldsymbol{\alpha}, \tag{7.9}$$

where $\mathbf{e}$ is the vector of all ones, Q is a $M \times M$ matrix such that $Q_{ij} = y_i y_j K(\mathcal{G}_i, \mathcal{G}_j)$, and $C > 0$ is an adjustable penalty hyper parameter. Setting C to a large value increases the range of possible values of $\boldsymbol{\alpha}$ and therefore the flexibility of the model. On the other hand, it also increases the training time and overfitting risk. The data points for which $\alpha_i^* > 0$ are called support vectors (SV). Once the α_i are trained, the class of a new graph $\mathcal{G}$ is predicted by the decision function, given by:

$$y(\mathcal{G}) = \operatorname{sgn}\left\{\sum_{i \in SV} y_i \alpha_i^* K(\mathcal{G}, \mathcal{G}_i)\right\}. \tag{7.10}$$

Once the kernel is trained, the prediction relies only on estimating with the kernel the distances between the unseen graph and the support-vector graphs. The choice of the kernel function determines the nature of the decision boundary that the SVM can learn.

- **(QC)** Quantum states analysed using classical ML tools, where ML techniques predict properties of quantum systems from limited measurements [301]. In the context of quantum simulation, the behaviour of the phase space over certain regimes of parameters can be predicted, aiding experimentalists in locating phase transitions [302].
- **(CQ)** Classical data processed on a quantum device, with the initial motivation on exploring whether quantum computing could accelerate ML sub-routines, akin to Shor algorithm's impact on factorisation [303]. Driven by complexity theory, researchers sought speed-ups in linear algebra routines with the notable, yet controversial, example of exponential speed up of the Harrow-Hassidim-Lloyd algorithm [304, 305].

While the following chapter 8 utilises the **CQ** approach as a way to speed up a combinatorial optimisation problem, this chapter focuses on the capabilities of quantum devices to act as ML models themselves rather than accelerating sub-routines. We thus treat classical data with a quantum processor in a **CQ** approach.

7.2.2 - Quantum dynamics as a learning model

With the recent advances in geometric QML, works have shown how classical graph-structured data could be encoded into quantum states and manipulated for classification, clustering or regression tasks. These efforts started with quantum convolutional neural networks [300, 306], and attempts were made to translate classical Graph Neural Network (GNN) architectures to Quantum Neural Networks (QNN) [307]. These procedures involve realising a *quantum feature map* [285, 286].

Using both the ability to encode graphs in the interaction Hamiltonian of neutral atom registers (see Sec. 6.1.2) and to dynamically evolve a quantum state with the control Hamiltonian of Eq. 1.27, we can generate from qubits initialised in $|0\rangle$ a graph-dependent wavefunction

$$|\psi_{\mathcal{G}}\rangle = U(\mathcal{G};t)\,|0\rangle^{\otimes|\mathcal{G}|}, \tag{7.11}$$

where the time evolution operator $U(\mathcal{G};t)$ reads

$$U(\mathcal{G};t) := \mathcal{T}\left[\exp\left(-i/\hbar\int_{s=0}^{t}(\hat{H}_{\mathcal{G}}+\hat{H}_{ctrl}(s))ds\right)\right]. \tag{7.12}$$

The state $|\psi_{\mathcal{G}}\rangle$ depends on the graph topology and on the parameterised pulse sequence applied to the qubits over time. This state can then be probed using a quantum observable $\hat{O}$ with eigenstates $\{|o_k\rangle\}_{k=1\cdots K}$, with $K = 2^{|\mathcal{G}|}$, resulting in the measurement of a probability distribution:

$$\mathcal{P}_{\mathcal{G}} = (p_1,\cdots,p_k,\cdots,p_K),\ \text{with}\ p_k = |\langle o_k|\psi_{\mathcal{G}}\rangle|^2. \tag{7.13}$$

Note that if some eigenvalues λ_k are degenerate, one would get instead

$$p_k = \sum_{i,\lambda_i=\lambda_k}|\langle o_i|\psi_{\mathcal{G}}\rangle|^2, \tag{7.14}$$

where p_k is restricted to the distinct eigenvalues of $\hat{O}$. In practice, if K is large, one would resort to binning the values of λ_i by defining a set of $K' < K$ intervals $\{I_k = [\tilde{\lambda}_k, \tilde{\lambda}_{k+1}]\}_{k=1\ldots,K'}$, with $\tilde{\lambda}_1 \leq \min_k \lambda_k$ and $\tilde{\lambda}_{K+1'} \geq \max_k \lambda_k$, such that $\mathcal{P}_{\mathcal{G}} = (\tilde{p}_1,\ldots\tilde{p}_{K'})$, where

$$\tilde{p}_k = \frac{|\{m_i|m_i\in I_k\}|}{m} \underset{m\to\infty}{\equiv} \sum_{i|\lambda_i\in I_k}|\langle o_i|\psi_{\mathcal{G}}\rangle|^2. \tag{7.15}$$

with m the number of measurements made to approximate the probability distribution.

We will call quantum feature map the function $\mathcal{G}\to\mathcal{P}_{\mathcal{G}}$ displayed in Fig. 7.3 with for ease of writing no mention on the observable $\hat{O}$ or the pulse parameterisation, such as the total time t, used to compute it.

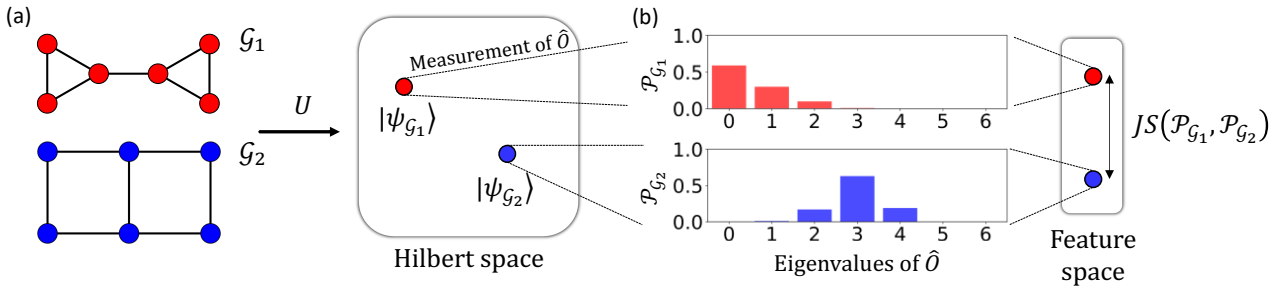


Figure 7.3: **Building a quantum feature map** (a) For each graph $\mathcal{G}$, a quantum system is evolved according to the propagator U. (b) Measuring $|\psi_{\mathcal{G}}\rangle$ with an observable $\hat{O}$ returns a probability distribution $\mathcal{P}_{\mathcal{G}}$ of the eigenvalues of $\hat{O}$. The distance between two distributions in the feature space can be obtained using the Jensen-Shannon divergence

7.2.3 - Expressive power of the quantum feature map

The graph quantum feature map already shows interesting properties when associated with single-body observables $\langle \hat{O}_{j=1,\dots,|\mathcal{G}|} \rangle$. The measured values are not only affected by local graph properties such as node degrees, but also by more global ones such as the presence of cycles. This enrichment provided by the quantum dynamics contrasts with the locality of node representations in many classical GML algorithms. This key feature comes from the fact that the quantum dynamics of a given spin model (*e.g.* an Ising model) will be significantly influenced, beyond short times (given by the Lieb-Robinson bound [308, 309]), by the complete structure of the graph.

We illustrate experimentally this behaviour for two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ displayed in Fig. 7.3(a) that are non-isomorphic but locally identical. In these graphs, nodes can be separated into two equivalence classes according to their neighbourhood: border nodes B have one degree-3 neighbour and one degree-2 neighbour, while centre nodes C have two degree-2 neighbours and one degree-3 neighbour. These graphs hold significance as they represent one of the smallest examples of non-isomorphic graphs where the Weisfeiler-Lehman (WL) test [310], presented in Box. 25, fails. We will see that the presence of interactions will enable us to discriminate between $\mathcal{G}_1$ and $\mathcal{G}_2$ by comparing the dynamics of local observables on border and centre nodes.

We first embed the graphs in a tweezer array with a nearest-neighbour (NN) distance of $r_{NN} = 5.3\ \mu$m and apply a constant pulse with $\Omega/2\pi = 1.0$ MHz and $\delta/2\pi = 0.7$ MHz. We then measure the local mean Rydberg excitation $\langle \hat{n}_j \rangle_{j \in B/C}$ for varying pulse duration $t \in [0, 2.5]\,\mu$s. As illustrated in Fig. 7.4, a qualitative difference in the dynamics of both graph appears after $t \sim 0.25\ \mu$s. Precisely, the excitation of the border nodes in Fig. 7.4(a) is initially increasing with indistinguishable behaviour between the two graphs. Then, a distinction appears between the two graph instances: the mean density for the border qubits of $\mathcal{G}_1$ exhibits damped oscillations around $\langle n_B \rangle \sim 0.15$ with period of the order of $0.5\ \mu$s while for $\mathcal{G}_2$ it exhibits flatter oscillations centered around 0.25 with period around $1\ \mu$s. We can observe a comparable distinction between the two graphs for the centre qubits in Fig. 7.4(b). The experimental measurements are consistent with the theoretical predictions supplemented by an effective noise model with detection errors and dephasing included. First, due to the nature of the quantum state and the limited budget of shots, measurements are subject to sampling noise. For

Box 25: Weisfeiler-Lehman test and Message passing NN

The WL graph isomorphism test serves as a heuristic method for determining the presence of an isomorphism between two graphs. Operating as a polynomial-time algorithm, the WL test is renowned for its local approach, primarily reliant on the vertex neighbourhoods within the graph. Initially, each vertex is assigned a unique label based on its immediate neighbours, followed by iterative refinements incorporating information from neighbouring vertices and their respective neighbourhoods. This process continues until a distinguishing characteristic emerges or the labels stabilize [311].
By focusing on local neighbourhood information, the WL test captures the inherent structure of the graphs under comparison. It exploits the notion that if two graphs are isomorphic, their vertices and respective neighbourhoods should exhibit similar label patterns throughout iterations. Conversely, non-isomorphic graphs should reveal distinguishing characteristics in their labels, signifying structural differences. The WL test also bears relevance to standard message-passing neural networks (MPNN) architectures, where information propagation occurs solely along the graph's edges. Notably, these successful MPNN models [312] have been proven to be, at most, as powerful in discerning graph structures [313] as the WL test.
Despite its potency, the WL test fails when applied to the two graphs depicted in Fig. 7.3(a). Hence, it becomes intriguing to explore the behaviour of these graphs when embedded with atoms and allowed to evolve under a quantum Ising Hamiltonian.

instance, on average, each of the 25 experimental points on Fig. 7.4 is obtained using 600 shots and the uncertainty related to this effect (vertical error bars) is estimated using the Jackknife resampling method [314]. Second, the detection errors, measured at $\varepsilon = 3\%$ and $\varepsilon' = 8\%$ for this experiment can modify the measured excitation distributions, with a noticeable effect shown on at $t = 0$ where the simulated $\langle n_j \rangle$ does not start at 0 despite $|\psi(t = 0)\rangle = |0 \dots 0\rangle$. Finally, we take into account positional disorder, Doppler effect and laser phase noise in a effective dephasing term $2\pi \times 0.06$ MHz obtained by fitting damped Rabi oscillations.

When restricted to the mean-field approximation (or similarly in the classical limit), the qubits' dynamics on either graphs are far more similar, as illustrated in the insets of Fig. 7.4. We still observe distinct dynamics between the two graphs, which is due to next nearest neighbours (NNN) interactions, more pronounced for the centre nodes. If we neglected those NNN interactions, the mean-field equations governing the dynamics of each qubit would only depend on its direct neighbourhood, *i.e.* the local structure of the graph. In that case, the qubits dynamics for $\mathcal{G}_1$ and $\mathcal{G}_2$ obey the exact same equations (see black dashed line in the insets). We therefore conclude that the presence of interactions in the system enables us to discriminate between the two non-isomorphic graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ by evaluating node-level local observables $\langle n_B \rangle$ or $\langle n_C \rangle$.

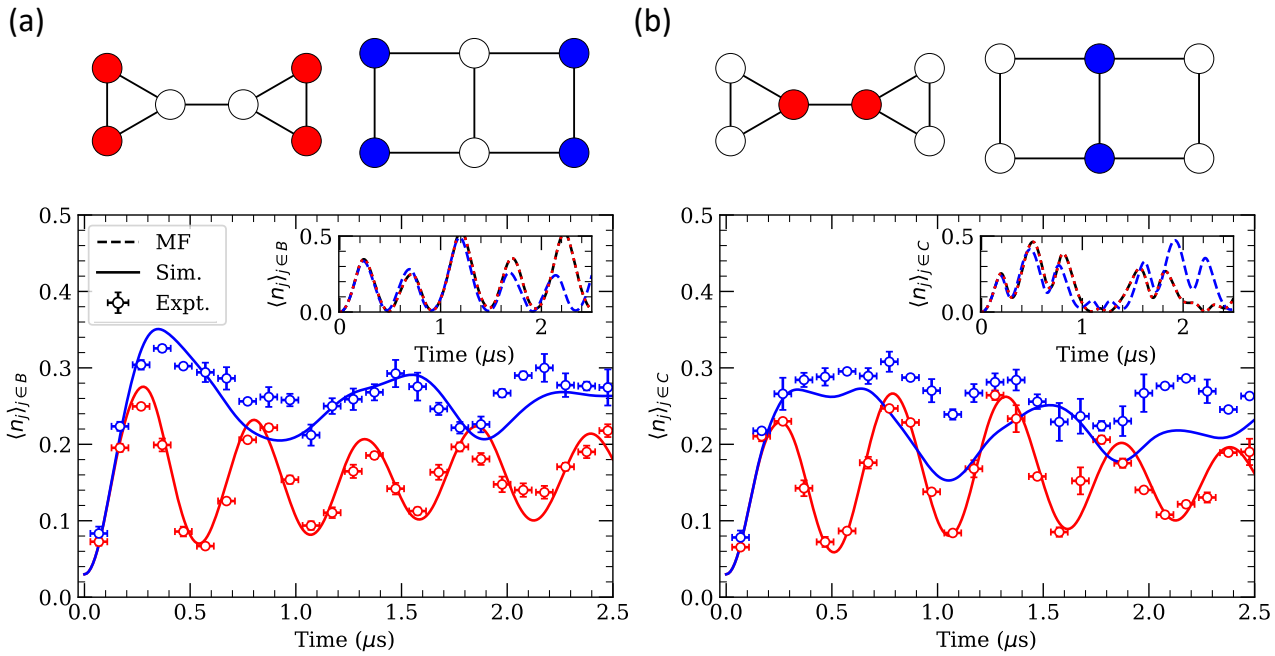


Figure 7.4: **Different Rydberg dynamics for locally-equivalent nodes.** Based on their neighbourhood, the nodes either belong to (a) the *border* region, B, or (b) to the *centre* region, C. Precisely, we plot the evolution of the mean occupation $\langle n_i \rangle$ of the two regions B and C for both graphs $\mathcal{G}_1$ (red) and $\mathcal{G}_2$ (blue). The dots represent the experimental results while the full curves show noisy simulation results. Horizontal error bars account for the sequence-trigger uncertainty ($\approx$ 40ns) while the vertical ones account for the sampling noise. The insets show the corresponding mean field dynamics (dashed) with only NN (black) or full (coloured) interactions.

7.2.4 - Building a quantum kernel

By looking at $\hat{O} = \sum_{i=1}^{6} \hat{n}_i$, we can quantify even more the difference in the dynamics between two graphs. To this end, we first compute the histogram $\mathcal{P}_{\mathcal{G}_i}(n)$ of the number of excitations observed in each shot on graph $\mathcal{G}_i$. The difference between those graphs is then estimated via the Jensen-Shannon divergence JS of their respective histograms [315], a commonly used distance measure between probability distributions, which is defined as

$$JS(\mathcal{P}, \mathcal{P}') = H\left(\frac{\mathcal{P} + \mathcal{P}'}{2}\right) - \frac{H(\mathcal{P}) + H(\mathcal{P}')}{2}. \tag{7.16}$$

Here $H(\mathcal{P}) = -\sum_k p_k \log p_k$ is the Shannon entropy of $\mathcal{P}$. The Jensen-Shannon divergence reaches its minimum at $JS(\mathcal{P}, \mathcal{P}) = 0$ and reaches its maximum at $\log 2$ if $\mathcal{P}$ and $\mathcal{P}'$ have disjoint supports. This is illustrated in Fig. 7.5, where the largest difference $JS_{max} \approx 0.28$ (roughly 40% of the theoretical maximal value) is achieved at a time $t \sim 0.57$ μs where the two distributions are the most distinct. At this duration, the distribution for $\mathcal{G}_1$ is sharply peaked at $n = 0$ while that of $\mathcal{G}_2$ is wider and peaks around $n = 2$, as illustrated in the inset. We note that the local observables $\langle n_j \rangle_{j\in B/C}$ also exhibit maximal deviation at this same duration t, indicating direct correspondence between measurements at the node and graph levels.

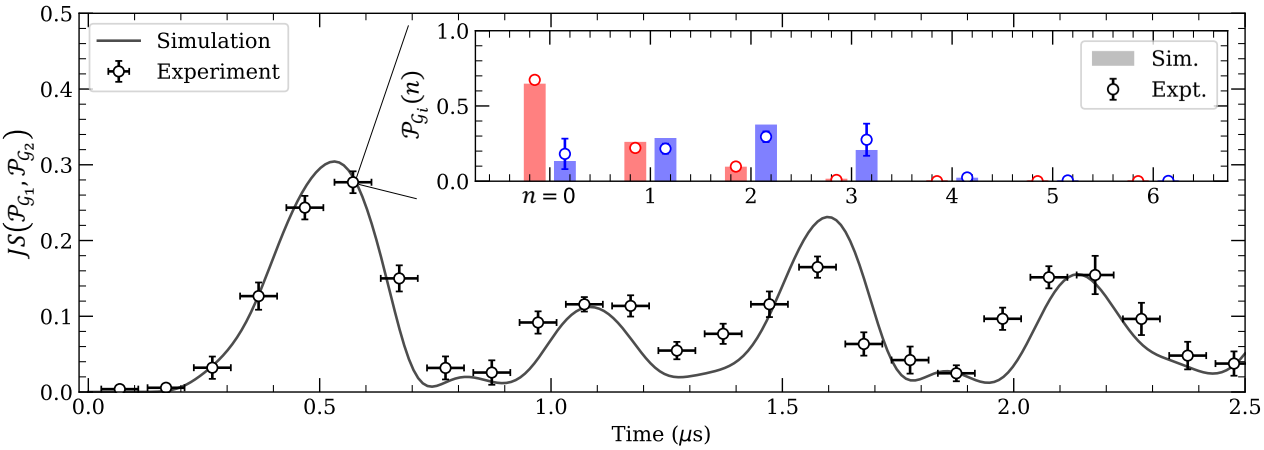


Figure 7.5: **Evolution of the Jensen-Shannon divergence of a global observable**. The experimental values (dot) are compared to the noisy simulation (plain). At each point in time, $JS(\mathcal{P}_{\mathcal{G}_1}, \mathcal{P}_{\mathcal{G}_2})$ is computed using the excitation distributions $\mathcal{P}_{\mathcal{G}_{1/2}}(n = 0, \ldots, 6)$ obtained either numerically (bar) or experimentally (dot). The inset depicts $\mathcal{P}_{\mathcal{G}_{1/2}}$ obtained at $t \approx 0.57$ μs which yields the maximum value $JS_{max} \approx 0.28$.

The dependency of local observables evaluated after the application of the quantum feature map on global graph structures have interesting consequences regarding quantum-enhanced versions of GNN [307, 316, 317]. In standard GNN architectures, information is only propagated along the graphs edges. Incorporating a propagation rule built from the quantum feature map above enables to go beyond this well-known limitation of GNN architectures.

We now turn the quantum feature map into a Quantum Evolution Kernel (QEK). In essence, our proposed approach involves associating each graph $\mathcal{G}$ with a probability distribution $\mathcal{P}_{\mathcal{G}}$ derived from the measurement of an observable on a quantum system governed by the graph topology. Subsequently, the kernel is computed utilising the Jensen-Shannon divergence introduced in Eq. 7.16. For two graphs $\mathcal{G}$ and $\mathcal{G}'$, and their respective probability distributions $\mathcal{P}_{\mathcal{G}}$ and $\mathcal{P}_{\mathcal{G}'}$ we define the graph kernel as

$$K_\mu(\mathcal{G}, \mathcal{G}') = \exp\left[-\mu\, JS(\mathcal{P}_{\mathcal{G}}, \mathcal{P}_{\mathcal{G}'})\right] \in \left[2^{-\mu}, 1\right]. \tag{7.17}$$

This kernel is well-defined, *i.e.* the kernel matrix is always positive definite. Throughout the rest of this chapter we set $\mu = 1$. Depending on the type of Hamiltonian evolution performed or observable measured, QEK can be related to other kernels, a subject detailed in Ref [318]. In the following section, we tackle binary classification tasks on graphs with a ML algorithm based on utilising our quantum kernel inside a classical SVM.

7.3 - Classification using QEK

7.3.1 - Binary classification task on PTC-FM

We now use the graph kernel introduced in Eq. (7.16) to tackle a binary classification task on a dataset of chemical compounds, called PTC-FM (Predictive Toxicity Challenge on Female Mice) [289, 319]. The objective is to accurately predict the reactivity of chemical compounds (toxic/positive or non-toxic/negative) based on their structural properties. Indeed in many cases poisonous proteins act as enzyme inhibitors, where the geometry of the protein fits to the binding site of an enzyme and perturbs its usual functioning [320, 321]. In the next sections, we will detail the hardware implementation protocol for obtaining a QEK-based SVM trained on this dataset. The whole process is illustrated in Figure 7.6. We will estimate the quality of the classification with the F_1 score defined in Eq. 7.1. On the one hand, correctly predicting the toxicity of a compound (increasing t_p) leads to better F_1 scores. On the other hand, classifying a toxic compound as harmless (increasing f_n) or a harmless compound as toxic (increasing f_p) results in a lower score.

Dataset and mapping on hardware

In the original PTC-FM dataset, the 349 molecules are represented under the form of graphs where each node is labelled by atomic type and each edge is labelled according to its bond type. We first truncate the dataset to small graph sizes in order to be able to numerically train the kernel in reasonable time, and discard larger molecules. For the $M = 286$ remaining graphs of this dataset, we take into account the adjacency matrix of the graphs representing the compounds and discard the nodes and edges labels. Note that the results of our implementation are therefore not directly comparable to kernel results in the literature which take into account edge and node labels (see Ref. [322] for example).

Each node of a graph will be represented by a qubit in the QPU. In the same fashion as for the UD graphs in combinatorial optimisation schemes described in chapter 6, we first need to determine the positions of these atoms in order to implement an interaction term that effectively reflects the graph topology. To this end we use the same mapping and batching techniques. As a reminder, starting from a Fruchterman-Reingold layout, the position optimiser needs to even out the distances between two connected nodes while locating unconnected nodes further away than connected ones between unconnected nodes. Using as a score to optimise the ratio between the minimum interaction term across connected pairs and the maximum interaction term between unconnected pairs, one can tailor the free-space optimiser to produce representative mapping of UD graphs. We achieve a significant increase of the mean ratio across the dataset used, starting from 5.9 with the classical Fruchterman-Reingold method and going up to 16.8. More than half the dataset exhibits a ratio higher than 10 and less than 5% of the dataset is embedded with some defects, *i.e.* a ratio smaller than 1.

We assess the benefit of this two-step approach by comparing the distributions of distance of pairs sharing an edge and pairs not sharing an edge before and after the optimisation as shown in Fig. 7.7. While some defects such as fake or missing bonds frequently appear in the pre-optimisation embedding, the optimised positions are constrained in such a way that a clear cut is visible between the two distributions.

Once the mapping is deemed good enough, we apply the batching method described in Sec. 6.4.2 to the dataset and successfully map the entire dataset of 286 graphs into only 6 SLM patterns. For

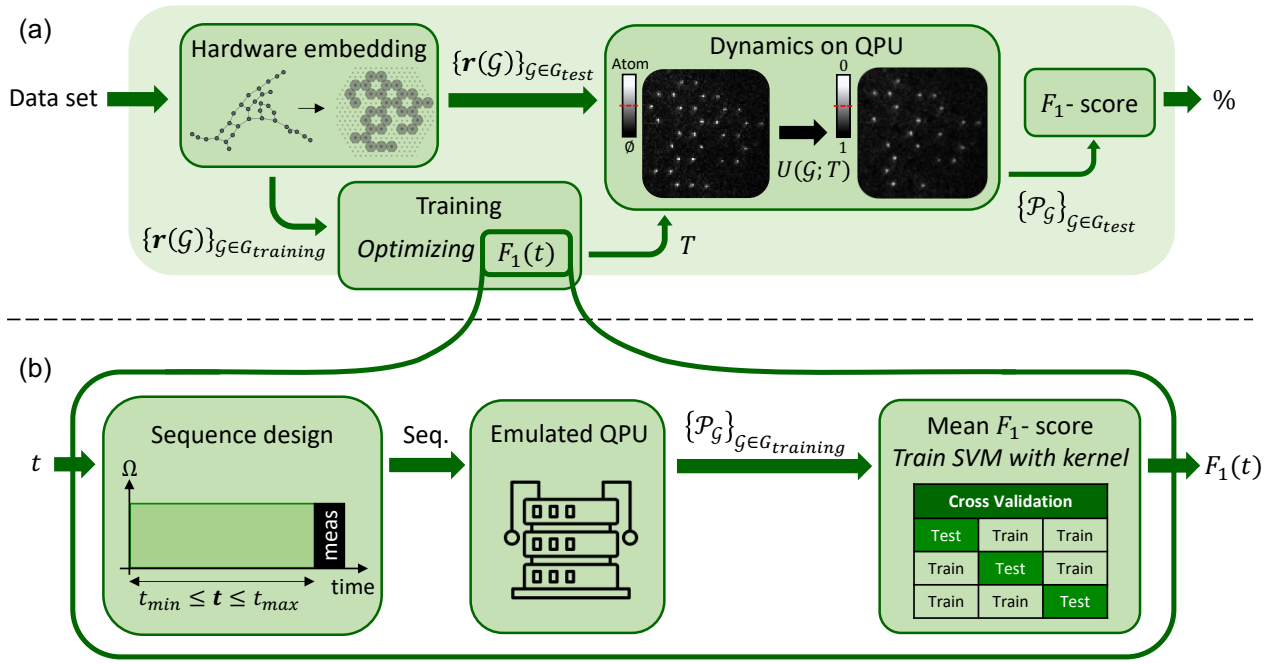


Figure 7.6: **Overview of QEK implementation.** (a) A dataset of graphs $\mathcal{G}$ is first mapped onto atomic registers $\mathbf{r}(\mathcal{G})$ implementable on the QPU, and separated between a training set $\mathcal{G}_{\text{training}}$ and a test set $\mathcal{G}_{\text{test}}$. We use the training set to determine numerically the optimal pulse sequence to be applied on the hardware using a grid search algorithm for optimising $F_1(t)$ (see (b)). This training phase outputs the optimal parameter T used to design the laser-pulse sequence applied experimentally on each register of the test set. The resulting dynamics performed on the QPU generates $U(\mathcal{G};T)$, driving the system from $|0\rangle^{\otimes\mathcal{G}}$ to $|\psi_\mathcal{G}\rangle$. A final F_1 score is then derived from assessing if the measured probability distributions $\{\mathcal{P}_\mathcal{G}\}_{\mathcal{G}\in G_{test}}$ are well labelled by the trained SVM . (b) The optimisation of the score function F_1 during the training includes several steps. The parameter t, taken from the constrained parameter space $[t_{\min}, t_{\max}]$ defines a sequence with the laser parameters Ω and δ being fixed. The dynamics of the system is emulated and enables to compute the probability distributions associated to this particular value of t for the whole training part of the dataset. Finally, $F_1(t)$ is obtained by fitting the SVM with the kernel constructed from those probability distributions.

example, we batch 66 graphs together onto a 71-trap pattern and on average, the 6 SLM patterns use 70 traps each to encode 48 graphs each. Then, tailoring the spatial disposition of the tweezers generated by the SLM to fit the optimised layouts, we can replicate the graphs in the hardware.

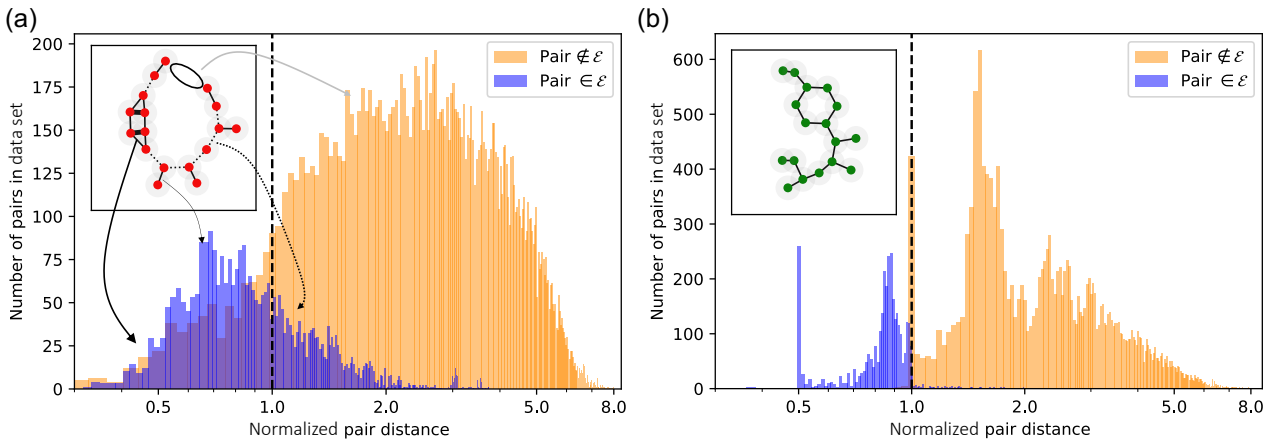


Figure 7.7: **Optimising positions for embedding graph topology in atomic registers.** Histograms of normalised pairwise distances between atoms in the 286 graphs of the truncated dataset when performing the embedding with (a) only a Fruchterman-Reingold layout or (b) when adding a local optimisation step afterwards. For a given graph (insets), two atoms forming a pair $\in \mathcal{E}$ (blue) can be close enough to form a bond via interaction (plain) or too far, creating a missing bond (dotted). Likewise, two atoms forming a pair $\notin \mathcal{E}$ can be placed too close and form a fake edge (thick line).

Model training

The quantum feature map used is thus based on embedding graphs into regular atomic patterns with a NN distance of $r_{NN} = 5.6$ μm and applying a resonant constant pulse with fixed $\Omega/2\pi = 1$ MHz and varying duration t between $t_{\min} = 0.1$ μs and $t_{\max} = 2.5$ μs. We first need to optimise the parameters of the quantum feature map, i.e. the hyper parameters of the ML model, to select those leading to the best classification score when combined with an SVM. Running the quantum dynamics on all the graphs of the training set with a chosen value of the hyper parameter $t \in [t_{\min}, t_{\max}]$ and fixed driving parameters will generate a set of probability distributions $\mathcal{D}_t = \{\mathcal{P}_\mathcal{G}\}_{\mathcal{G}\in G_{\text{training}}}$. From these distributions, a kernel matrix can be constructed and fed to an SVM. Through a cheap classical optimisation, the latter is trained and its decisions can be benchmarked against a test set. Given the modest size of the dataset, we can evaluate the performance of the SVM using repeated cross-validation. One realisation of the procedure entails randomly dividing $\mathcal{D}_t$ into d equal parts, termed "splits," where one split is reserved for testing while the remaining $d-1$ splits are used to construct a kernel matrix passed to the SVM for training. Subsequently, the SVM is tasked with classifying the elements of the reserved split and a F_1 score can be derived from its performance. This process is repeated for each of the d splits, putting one aside and training on the remaining, finally giving out a mean F_1 score computed by averaging the d values obtained. In addition, the overall procedure can be iterated multiple times with random dataset splits. We can consider this cross validation over the training of the SVM as a block (see Fig. 7.6(b)) taking as input a set of probability distributions obtained for hyper parameter t, or equivalently the constructed kernel, and returning the corresponding F_1 score. Performing the hyper parameter optimisation on the QPU would involve applying several sequences (usually $n_{\text{iter}} \sim 100$) on the whole graph dataset, requiring a tremendous

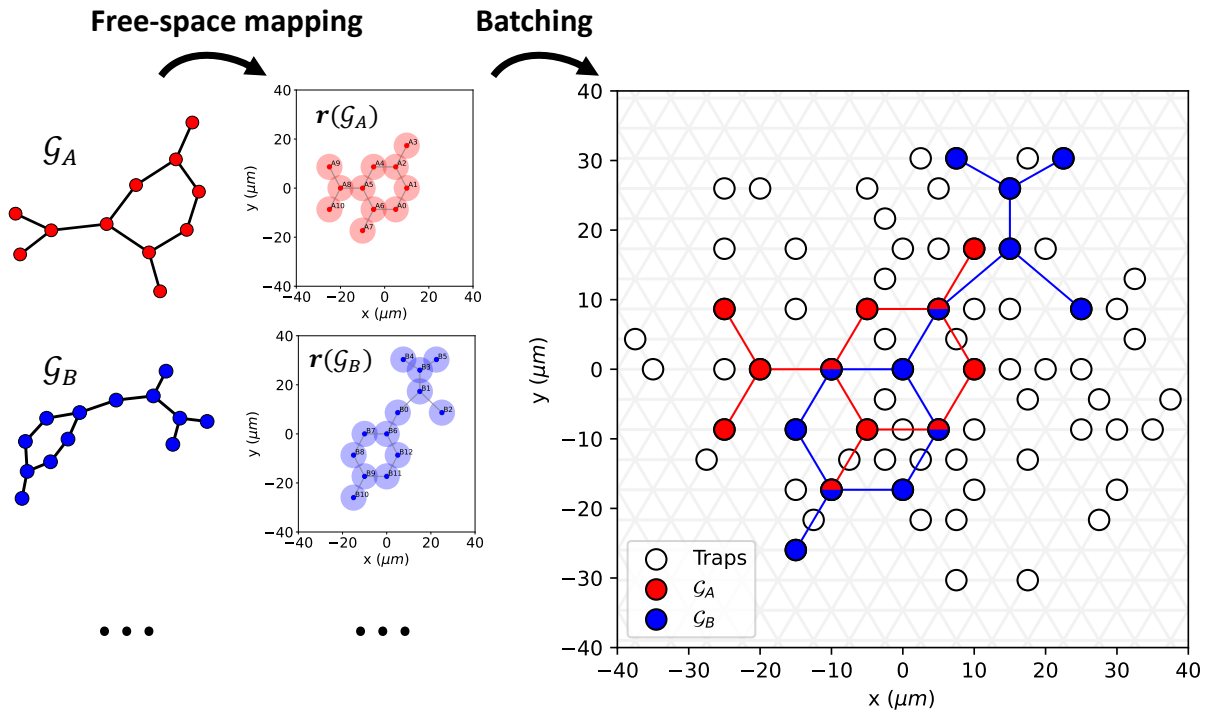


Figure 7.8: **Batching graphs to the same SLM pattern.** A family of 66 graphs, ranging in sizes from 4 to 19 nodes, is mapped in free-space and then batched to the same SLM pattern (white dots) over a triangular grid with spacing 5.6 μm. The traps used when implementing $\mathcal{G}_A$ ($\mathcal{G}_B$) are coloured in red (blue). The bi-coloured traps are those used for both graphs.

amount of computation ($n_{\text{iter}} \times M \times n_{\text{shots}} \sim 1500$ hours for this implementation). We therefore perform this optimisation numerically, only including in the training part of the dataset graphs with sizes $|\mathcal{G}| \leq 20$. The score to optimise $F_1(t)$ is defined as the average of the F_1-score over 10 realisations of the cross validation scheme with 5 splits. For each split, we perform a grid search on the penalty hyper parameter C of the SVM (see Box. 24) on the range $[10^{-3}, 10^3]$ such that the final score of a given pulse is the best cross-validation score among the tested values of C. This numerical procedure is described on Fig. 7.6(b).

We select the optimal duration $T = 0.66\ \mu s$ that exhibits the maximum $F_1(t)$-score and apply the constructed feature map to the whole dataset of $M = 286$ graphs. It should be noted that we are using the feature map optimised for graphs smaller than size 20 on the entire dataset where the graphs have up to 32 nodes. This may limit the benefit of hyper parameter optimisation, but remains better than a random choice of t. Without resource limitation on the QPU, we could resort to the procedure described in Fig. 7.6(a), with a training handled directly on the QPU and a fixed split between training and testing sets. However, due to limited size of the dataset, we apply a cross validation procedure in the last block of Fig. 7.6(a), randomly splitting the experimentally obtained probability distributions between training and testing sets and outputting an average F_1 score.

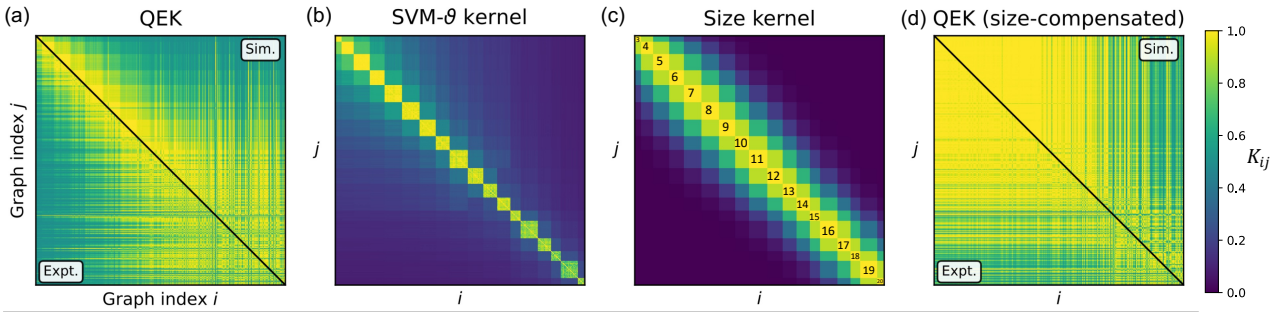


Figure 7.9: **Kernel matrices obtained numerically or experimentally.** Each kernel is represented by a $M \times M$ matrix where $K_{i,j} = K(\mathcal{G}_i, \mathcal{G}_j)$. The graph indices are sorted by increasing size. A separation (black line) is drawn between numerically simulated (top right) and experimentally measured (bottom left) QEK matrices. (a) QEK kernel obtained using directly the raw distributions $\mathcal{P}_{\mathcal{G}_i}$ and $\mathcal{P}_{\mathcal{G}_j}$. (b) Kernel obtained vian SVM-ϑ method. (c) Size kernel obtained with $\gamma = 0.1$. (d) K_{conv}, QEK kernel obtained using modified distributions $\tilde{\mathcal{P}}_{\mathcal{G}_i}$ and $\mathcal{P}_{\mathcal{G}_j}$, where graphs of smaller sizes are convoluted with binomial distributions when compared to larger graphs.

Classification results

After a training of our model, we experimentally obtain an F_1-score of $60.4 \pm 5.1\%$. For comparison purposes, we examine the performances of other kernels on this dataset: the GS, RW, SP and SVM-ϑ kernels, all these kernels being described in detail in Appendix A. The F_1-scores reached by the various kernels are collected in Table 7.1. Obtained scores range from $49.8 \pm 6.0\%$ up to $58.2 \pm 5.5\%$. Those results show that the Quantum Evolution Kernel is competitive with standard classical kernels on this dataset. The SVM-ϑ kernel is found to be, among the classical kernels tested, the one with the best performance. As described in Appendix A.1, it is defined up to a choice of base kernel between real numbers, which gives it a certain degree of flexibility.

We show in Fig. 7.9(a) the kernel matrix associated with QEK, with indices sorted by increasing size of the graphs. Using the same noise model as in the previous section, we find adequate agreement between the numerically $\mathcal{P}^{num}$ and experimentally $\mathcal{P}^{exp}$ obtained data. Quantitatively, we make use of the JS divergence to estimate this agreement for any $\mathcal{G}_i$ and observe that $\langle JS(\mathcal{P}^{num}_{\mathcal{G}_i}, \mathcal{P}^{exp}_{\mathcal{G}_i})\rangle_i \approx 0.03 \pm 0.01$ is one order of magnitude below $\langle JS(\mathcal{P}^{exp}_{\mathcal{G}_i}, \mathcal{P}^{exp}_{\mathcal{G}_j})\rangle_{i \neq j} \approx 0.33 \pm 0.01$. An interesting feature of both QEK and SVM-ϑ (Fig. 7.9(b)) kernel matrices is the emergence of size-related diagonal blocks, signalling that the models identify the size of the graphs as an important feature for classification. Examining more closely the dataset, we indeed remark that the subset of PTC-FM that we used is significantly size imbalanced, as illustrated in Fig. 7.10. Since the graph size seems to be a relevant feature for this particular dataset, we check that the size kernel defined in Eq. 7.8 reaches a F_1-score of $56.7 \pm 5.6\%$. The corresponding kernel matrix is displayed in Fig. 7.9(c) and exhibits a block-diagonal shape with a Gaussian tail. It is interesting to note that the quantum model was able to identify size as a relevant parameter for this dataset, leading to classification results which are on par with the best classical kernels.

Going forward, we modify the QEK procedure in order to make the kernel insensitive to size.

Kernel	$\boldsymbol{F_1}$**-score (%)**
QEK	60.4 ± 5.1
QEK (size-compensated)	45.1 ± 3.7
SVM-ϑ	58.2 ± 5.5
Size	56.7 ± 5.6
Graphlet Sampling	56.9 ± 5.0
Random Walk	55.1 ± 6.9
Shortest Path	49.8 ± 6.0
Random (baseline)	42

Table 7.1: F_1**-score reached experimentally on the PTC-FM dataset by QEK** ($\pm$ std. on the splits). In addition, the scores reached numerically by the classical kernels SVM$-\vartheta$, Size, Graphlet Sampling, Random Walk and Shortest-Path. The values reported are the average over a 5-fold cross-validation repeated 10 times. The baseline is set by looking at the score when randomly labelling instances which due to the class imbalance of the dataset is 42%.

To that end, we compare the measurement distributions obtained for different graph sizes using a convolution operation. Let us consider two graphs $\mathcal{G}_i$ and $\mathcal{G}_j$, of N_i and $N_j = N_i + \Delta N > N_i$ nodes respectively, and note their respective observable distributions $\mathcal{P}_{\mathcal{G}_i}$ and $\mathcal{P}_{\mathcal{G}_j}$. From $\mathcal{P}_{\mathcal{G}_i}$ we construct $\tilde{\mathcal{P}}_{\mathcal{G}_i} = \mathcal{P}_{\mathcal{G}_i} \star b_{\Delta N}^{(i/j)}$ the convolution of $\mathcal{P}_{\mathcal{G}_i}$ and a binomial distribution :

$$\binom{\Delta N}{n} p^n (1-p)^{\Delta N - n}. \tag{7.18}$$

$\tilde{\mathcal{P}}_i$ corresponds to the distribution one would get by adding to the graph ΔN non-interacting qubits, submitted to the same laser pulse as the other. Each of these isolated qubits undergoes Rabi os-

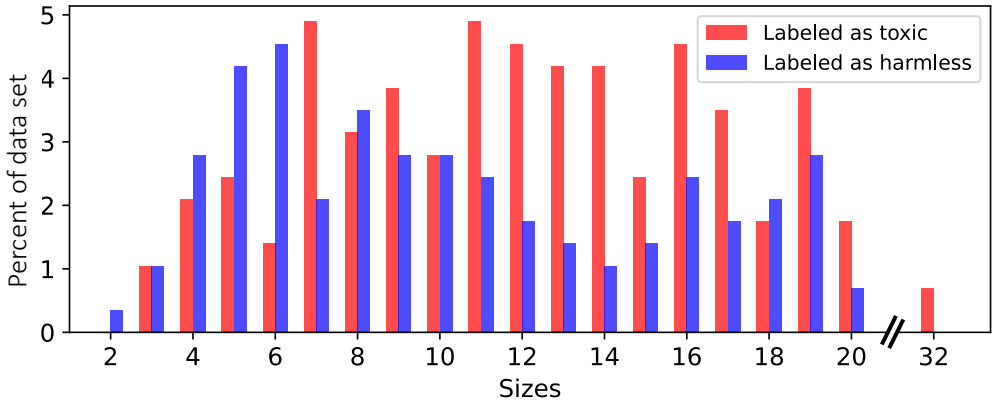


Figure 7.10: **Size imbalance in PTC-FM dataset.** For small number of nodes ($\lesssim 10$) more graphs are labelled as harmless (blue) while it is the opposite for larger graphs, more prone to be labelled as toxic (red).

cillations, induced by the applied pulse sequence. They are therefore measured either in $|0\rangle$ with probability p or in $|1\rangle$ with probability $1-p$, where $p = \sin^2(\pi\Omega T)$ (≈ 0.768 here). We finally define the modified graph kernel as

$$K_{conv}(\mathcal{G}_i, \mathcal{G}_j) = \exp\left[-JS(\tilde{\mathcal{P}}_{\mathcal{G}_i}, \mathcal{P}_{\mathcal{G}_j})\right]. \tag{7.19}$$

Using this procedure on the data obtained experimentally, we obtain the kernel matrix shown in Fig. 7.9(d), with a corresponding F_1-score of $45.1 \pm 3.7\%$. If this size-compensated version of QEK had been implemented without interaction between atoms, its score would be 42%, which is the lowest score reachable by any kernel. We therefore see that this compensated version of QEK cannot capture useful features beyond the graph size, meaning that the presence of interactions by itself is not sufficient to produce an interesting kernel for the task at hand. While the size-compensated QEK does not give results that are comparable with classical kernels, we study in the following part its expressive power, and show that the geometry induced by this method is hardly reproducible by a classical kernel.

7.3.2 - Geometric test with respect to classical kernels

Geometric difference

In order to obtain an advantage over classical approaches it is not sufficient to implement a quantum feature map based on quantum dynamics that are hard to simulate classically. As shown in [290], classical ML algorithms can in certain instances learn efficiently from intractable quantum evolution if they are allowed to be trained on data. The authors consequently propose another metric between kernels in the form of an asymmetric metric function called the *geometric difference* g_{12}, further detailed in Box. 26. Intuitively, g_{12} characterises the disparity regarding how each kernel K_1 and K_2 maps data points to their respective feature spaces. In our case, we take K_1 to be the size-compensated QEK K_{conv}, and K_2 is selected from a set of classical kernels. If the geometric difference is small, it means that there exists no underlying function mapping the data to the targets for which K_{conv} outperforms the classical kernel. On the other hand, a high geometric difference between a quantum and a classical kernel guarantees that there exists such a function for which the quantum model outperforms the classical one. Estimating the geometric difference is therefore a sanity check to stating that the encoding of data to the feature space through QEK could not be closely replicated by a classical model.

We compute the geometric difference between QEK and various classical kernels over the PTC-FM dataset and report the results in Table 7.2. The threshold for a high geometric difference is typically taken to be $\sqrt{M}$, where M is the size of the dataset. Here, the obtained g_{12} is always far beyond $\sqrt{M} \sim 10^1$, indicating that the embedding of data through our quantum-enhanced kernel is not trivial and cannot be replicated by a classical ML algorithm.

Box 26: Geometric difference and relabelling

Given two kernel functions K_1 and K_2, the geometric difference g_{12} described in [290] is an asymmetric distance function that quantifies whether or not the kernel K_2 has the potential to resolve data better than K_1 on some dataset. In its simplest form, the geometric difference is defined as:

$$g_{12} = \sqrt{||\sqrt{K_2}\,(K_1)^{-1}\,\sqrt{K_2}||_\infty} \tag{7.20}$$

where $||\cdot||_\infty$ denotes the spectral norm. The geometric difference becomes an especially useful metric when $K_1 = K_C$ is a classical kernel and $K_2 = K_Q$ is a quantum kernel. If M is the size of the dataset, a value of g_{CQ} of order $\sqrt{M}$ or greater indicates that the geometry of the feature space induced by the quantum kernel is rich enough to be hard to learn classically, and the quantum kernel can potentially perform better than classical kernels. In that case, it is possible to artificially relabel the dataset in order to maximally separate the kernels' performance. Such a relabelling process is a constructive proof of the existence of a certain dataset on which one kernel performs much better than the other. If v is the eigenvector of $\sqrt{K_2}\,(K_1)^{-1}\,\sqrt{K_2}$ corresponding to the eigenvalue g_{12}^2, the vector of new labels is given by $y_{\text{new}} = \sqrt{K_2}v$.
When dealing with a finite amount of training data, Eq. 7.20 should be regularised in order to stabilise the inversion of K_1. The regularised expression reads:

$$g_{12}(\lambda) = \sqrt{||\sqrt{K_2}\sqrt{K_1}\,(K_1 + \lambda I)^{-2}\,\sqrt{K_1}\sqrt{K_2}||_\infty} \tag{7.21}$$

where λ is the regularisation parameter. The geometric difference $g_{12}(\lambda)$ has a plateau for small λ, when the regularisation parameter becomes smaller than the smallest eigenvalue of K_1, and decreases for increasing λ. The effect of λ is to introduce a certain amount of training error. The training error can be upper bounded by a quantity proportional to:

$$g_{\text{tra}}(\lambda)^2 = \lambda^2||\sqrt{K_2}\,(K_1 + \lambda I)^{-2}\,\sqrt{K_2}||_\infty. \tag{7.22}$$

Practically, one should look at the regime where g_{12} has not plateaued but the training error is still small enough. A regularisation should be introduced also in the relabelling procedure. The new labels are taken to be $y_{\text{new}} = \sqrt{K_Q}v$, where v is the eigenvector of the regularised matrix

$$\sqrt{K_Q}\sqrt{K_C}\,(K_C + \lambda I)^{-2}\,\sqrt{K_C}\sqrt{K_Q}$$

corresponding to the eigenvalue $g_{12}(\lambda)^2$.

Relabelling the targets

To summarise, while the F_1-score on PTC-FM is rather similar using quantum or classical models, we see nonetheless that the geometry created by our quantum model is non-trivial. A possible interpretation of the non-superiority of quantum approaches on PTC-FM would be that the relationship between the data and the targets is not better captured by our quantum model, although its feature space is not reproducible by classical means. To further confirm this understanding, we find a function that increases and even maximises the utility of the potentially richer quantum feature space. We

build such a function by artificially relabelling the targets according to a procedure presented in [290] and outlined in Box. 26. We observe that QEK, without retraining, retains an F_1-score of around 99% on the relabelled dataset, while the closest classical kernel reaches a score of at most 82% even after retraining it on the new labels. The results are summarised in Table 7.2, where the difference in F_1-score between QEK and the various classical kernels after relabelling is shown.

Classical kernel	g_{12}	F_1**-gap** (%)
SVM-ϑ	10^3	17.2 ± 4.5
Size	10^5	17.8 ± 4.2
Graphlet Sampling	10^4	20.1 ± 4.5
Random Walk	10^5	17.3 ± 4.3
Shortest Path	10^5	18.2 ± 4.4

Table 7.2: **Comparison between QEK and classical kernels**. First column gives the order of magnitude of the geometric difference g_{12} between quantum and classical kernels. Second column gives the gap in F_1-score after relabelling the dataset.

In light of the geometric difference assessment and the observed gap of F_1-score between QEK and classical kernels on an artificial function, it remains an open question to generally characterise which types of dataset naturally offer a structure that better exploits the geometry offered by our quantum model, without requiring artificial tweaking of the labels. In the following section, we present a synthetic dataset on which QEK is able to outperform classical methods without any relabelling.

7.3.3 - Building and classifying a synthetic dataset

This binary classification dataset is created by sampling weighted random walks on a triangular lattice. In class A, sites belonging to a honeycomb-type sublattice are favoured. They are explored with a weight 1 while the rest of the triangular lattice sites are explored with a weight $p < 1$. Class B is constructed in a similar fashion, but taking a Kagome sublattice instead of a honeycomb one. The construction of this artificial dataset is illustrated in Fig. 7.11. For $p = 0$, the dissimilarities in local structure between graphs from the two classes facilitate easy differentiation. However, as p increases, their local structures become progressively more alike due to the incorporation of additional triangular lattice sites. When p reaches a sufficiently large value, numerous triangular local substructures are shared by graphs of both classes, potentially making them challenging to distinguish using classical methods. At $p = 1$, the underlying triangular lattice is uniformly explored when generating both datasets, rendering the two classes indistinguishable.

Exploiting our capability to discern between graphs with similar local but globally distinct structures, we apply QEK to this synthetic dataset. We anticipate our method to be minimally affected by the presence of sparse defects and thus to surpass classical approaches in performance.

We investigate numerically this assumption, for several values of p. In each case, we create 200 graphs of 20 nodes each, 100 in each class. The graphs are mapped to a triangular lattice with 5 μm spacing. Here, we consider two alternative schemes of pulse sequences. The first one remains almost the same as the experimentally implemented one, i.e. a unique resonant pulse of $\Omega/2\pi = 2$ MHz with parameterised duration up to 8 μs. The second one is an alternate layer scheme with 4 parameters as

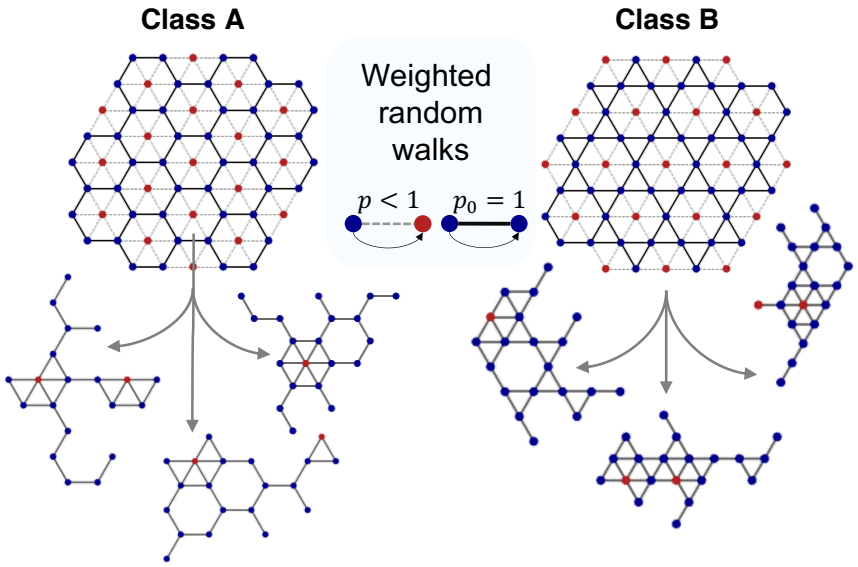


Figure 7.11: **Synthetic dataset.** Graphs in class A contains honeycomb sites (blue) with inclusions of non-honeycomb sites (red) with probability p. Graphs in class B contains Kagome sites (blue) with inclusions of non-Kagome sites (red) with probability p. We show examples of generated graph with the aforementioned process.

described in [318], where we evaluate 500 random values of the parameters and select the best one. The procedure is designed such that it would be directly implementable on the hardware, as we did for the PTC-FM dataset. We then compare the F_1-score reached by QEK to those reached by other classical kernels, namely: SVM-ϑ, GS, RW and SP. The results are summarised in Fig. 7.12. With decreasing proportion of defects, all methods perform increasingly better, as expected. Overall, regarding the mean F_1-score reached, the two QEK schemes outperform the four other classical kernels tested for all $p \leq 0.5$. Noticeably, at $p = 0.1$ (*resp* $p = 0.2$), the mean gap in F_1-score between the QEK scheme and the the best classical scheme is 4.5% (*resp* 7.1%) while the mean gap obtained with the alternate QEK scheme is even larger with 13.7% (*resp* 21%), thus showing that QEK can significantly surpass classical approaches on certain types of datasets. When adding too many defects, *i.e.* $p = 0.5$, our QEK exhibits similar performance to the SVM-ϑ.

7.3.4 - Conclusion

In this chapter, we reported the implementation of a quantum feature map for graph-structured data on a neutral atom quantum processor. We experimentally show that this embedding is not only sensitive to local graph properties but is also able to probe more global structures such as cycles. This property offers a promising way to expand the capabilities of standard GNN architectures, which have been shown to have the same expressiveness as the WL isomorphism test in terms of distinguishing non-isomorphic graphs [313, 323]. Some properties of quantum-enhanced versions of GNNs have been explored in [317].

We then use the quantum graph feature map for a toxicity screening procedure on a standard bio-chemistry dataset comprising 286 graphs of sizes ranging from 2 to 32 nodes. This procedure

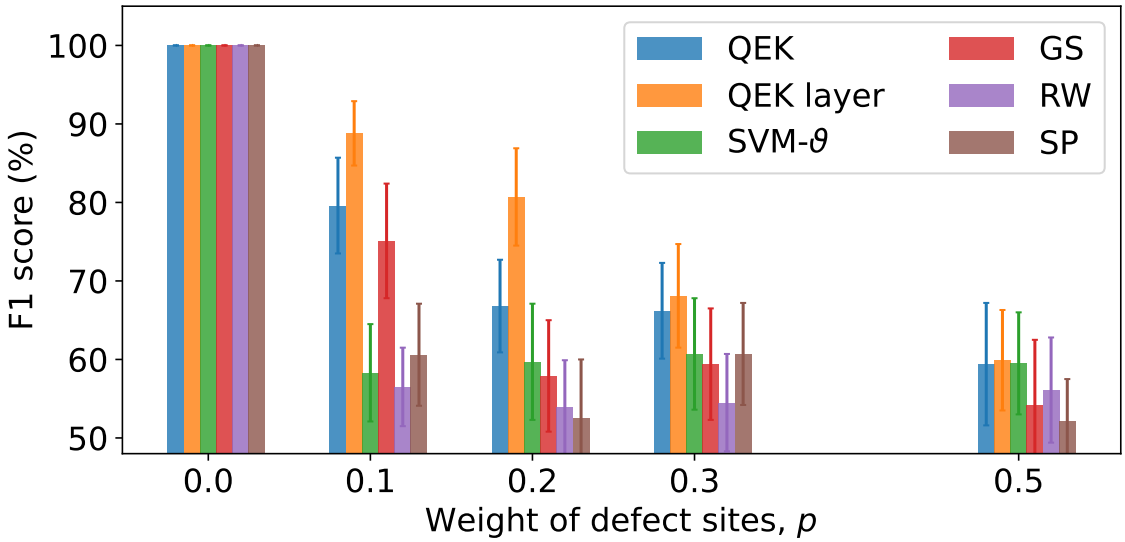


Figure 7.12: **F_1-score reached on the synthetic dataset** for different probabilities p of including defects using the Quantum Evolution Kernel (the alternate scheme is noted QEK layer) as well as by the best SVM-ϑ, GS, RW and SP kernels. The values reported are the average over a 5-fold cross-validation repeated 10 times. Each kernel reaches an F_1-score of 100% when $p = 0$.

achieves an F_1-score of $60.4 \pm 5.1\%$, on par with the best classical kernels. We intentionally do not include GNNs in the benchmark, as they belong to another distinct family of models. Beyond this pure performance assessment, we showcase the potential advantage of using a quantum feature map through the computation of geometric differences with respect to said classical kernels, which are metrics evaluating the degree of similarity between the kernels' feature spaces. We show that QEK captures features that are invisible to the classical kernels we consider. An artificial relabelling of the data enables us to create a synthetic dataset for which the performance of QEK could not be matched. We also identify another dataset made of bi-partite 2D lattices, for which the quantum procedure exhibits superior performances.

This proof-of-concept illustrates the potential of quantum-enhanced methods for graph machine learning tasks. Further work on more diverse datasets will be required to assess the viability of the approach compared to powerful state-of-the-art GNN architectures [324–327]. These experimental developments stand to benefit from recent advances in geometric quantum machine learning [328, 329], with the latter exploring applications to weighted graphs. Furthermore, in-depth theoretical studies on equivariant and geometric aspects of quantum machine learning have also been presented recently [330–332].

8 - Hybrid algorithm: quantum-enhanced classification in finance

Contents

Finance is one of the most promising fields where quantum computing could be useful in practical industrial applications [333]. ML models capable of handling the large datasets collected in the financial world can often become black boxes expensive to run. In addition, their training often involve solving hard combinatorial problems such as the ones introduced in chapter 6. The optimisation techniques provided by the quantum computing paradigm may deliver competitive, faster and more interpretable models when combined with already efficient classical ML framework. Integrating quantum solutions as modules in a classical-quantum hybrid workflow is at the core of today efforts to find practicality for noisy intermediate-scale neutral atom quantum computers.

Quantum and quantum-inspired approaches have already shown many promising applications in financial problems [334–341]. In this work, we implement an hybrid algorithm in the form of a quantum-enhanced machine learning solution for the prediction of credit rating downgrades in the financial risk management field. The QUBO-solving block of the QBoost-based algorithm presented below utilises a neutral atom QPU with a native quantum sampling approach called Random Graph Sampling. Using the proposed quantum solution, we achieve a similar performance as with the classical counterpart algorithm used as benchmark using comparable training times and showcasing better

interpretability, *i.e.* a simpler and smaller predictive model. More details on this work can be found in this paper:

- [342] L. Leclerc et al. "Financial risk management on a neutral atom quantum processor". Phys. Rev. Res. 5, p. 043117, 2023.

8.1 - A financial risk management use case

In finance, an interesting and relevant problem consists in estimating the probability of debtors reimbursing their loans, which represents an essential quantitative problem for banks. Financial institutions generally attempt to estimate credit worthiness of debtors by sorting them in classes called *credit ratings*. These institutions can build their own credit rating model but can also rely on credit ratings provided by one or more of the three main rating agencies Fitch, Moody's and Standard & Poor's (S&P). Borrowers are generally grouped into two main categories according to their credit worthiness: investment grade borrowers with low credit risk, and sub-investment grade borrowers with higher credit risk. Should a borrower's rating downgrade from investment to sub-investment category, it will be referred to as a *fallen angel*.

8.1.1 - Fallen angels forecasting using machine learning

The early anticipation of these fallen angels is a problem of utmost importance for financial institutions and one that has gathered significant attention from the ML community in the past years. Indeed, these institutions usually have access to large amount of data accumulated over several decades. The wide variety of features gathered can be fed to advanced ML models tasked with solving a *classification* problem. As a reminder, a classification task involves categorising input data $\boldsymbol{x}$ into predefined classes or categories y. The goal is to train a model to learn patterns from already labelled training data [Fig. 7.1(a)]. This classifier can then accurately assign new data points to one of these classes based on their features [Fig. 7.1(c)]. For example, classifying a pet as dog or cat, a chemical compound as toxic or harmless or a client as likely or not to reimburse a loan are all binary classification tasks with only two possible output $y = \pm 1$. In the fallen angel forecasting context, the trained ML model needs to answer with good enough performance the following question: will a debtor have a high or low risk of becoming a fallen angel in the foreseeable future?

Proposals of binary classification methods, or *classifiers*, targeting such tasks have shown promising results using Random Forest (RF) and XGBoost algorithms [343, 344]. The RF algorithm is a well-known ensemble method based on bootstrap aggregation, also called *bagging*, applicable for regression and classification alike [121]. Training a RF of m trees on a given dataset of size n entails generating m new datasets with n elements each by sampling with replacement the original training set. To ensure low correlation between the trees, each tree is also trained on a different subset of randomly selected features. The trained classifiers are then collectively used to predict the class of unseen data through majority vote over the m decisions. Due to their feature extraction flexibility, those tree-based ensemble methods turned out to be more suitable for similar credit risk modelling tasks [345, 346], compared to deep learning approaches [347, 348]. However, these methods quickly become computationally demanding as the numbers of decision trees grow. Furthermore denser and denser forests usually become black boxes in terms of interpretability, i.e. hard to understand by their users.

8.1.2 - Benchmarking the use case with a classical solution

Dataset

The dataset used in this study originates from public data over a period of 20 years (2001-2020). It comprises of more than 90 000 instances characterised by around 150 features, representing the historical evolution of credit ratings as well as numerous financial variables. Predictors include rating, financial and equity market variables and their trends calculated on a bi-annual, quarterly and five-year basis. The examples considered are based on over 2000 companies from 10 different industrial fields (e.g. energy, healthcare, utility) and 100 sub-sectors (e.g. infrastructure, oil and gas exploration, mining), located in 70 different countries. Each of the records is labelled as either a fallen-angel (*i.e.* critical downgrade; class 1/positive) or a non fallen-angel (*i.e.* stable or upgraded credit score; class 0/negative) based on standard credit rating scales. The training set consists of around 65 000 examples from the 2001-2016 period. The testing set includes around 26 000 examples from the 2016-2020 period. The class distribution is highly unbalanced with only 9% of fallen angels in the training set and 12% in the test set, making it hard to find generalisable patterns.

Performance metrics

The metrics used are defined in Eq. 7.1 Our primary goal consists in increasing the precision P of predicting fallen angels while keeping the recall R over 80%. This goal amounts to increase the number of true positive, i.e. of successful detections of fallen angel, while keeping the number of errors under control. However, the impact of f_n is far worse than the one of f_p as missing a fallen angel is far more detrimental in terms of resources for a bank than remaining cautious by wrongly considering instances as potential fallen angels. It is therefore preferable to lower f_n at the expense of maybe lowering t_p too and improve the precision, at the expense of maybe increasing f_p.

Classical solution

The provided classical ML algorithm is a standard RF model using 1200 decision trees. To deal with the highly skewed distribution of classes, as mentioned above, both random under-sampling of majority class and over-sampling of minority class were tested to balance the training set. Training the RF model and optimising its hyperparameters through random search lasts more than 3 hours on a classical computer. The model achieves $R = 83\%$ and $P = 28\%$. This result, far from being optimal, especially in terms of precision, is due to several factors, representative of the complexity of the problem:

1. The dataset is highly unbalanced, which is a notoriously hard problem for classical machine learning models.

2. Processing a significant amount of features can be resource-consuming and it remains impossible to exhaustively search the space of solutions at too large sizes. Therefore, the classical method uses a suboptimal shortcut to select relevant features.

3. Finding the optimal weight for each predictor is an exponentially complex optimisation problem as the number of predictors increases. Hence, the RF model uses majority voting for classification, which is quite restrictive in terms of performance.

In the following, we will keep the RF model performance as a benchmark baseline and build a model aiming at enhanced precision while keeping the recall at this same value of $R = 83\%$. Accommodating this criterion of constant recall requires tuning the decision threshold, a parameter that governs conversion of class membership probabilities to the corresponding hard predictions (e.g. 0 or 1). Specifically, P and R will be computed for several decision thresholds and a linear interpolation enables to determine the precision value corresponding to a given R.

8.2 - Quantum-enhanced classifier

We address the limitations identified above with a proposed quantum-enhanced machine learning approach, taking advantage of quantum combinatorial optimisation to efficiently explore the space of solutions.

8.2.1 - QBoost framework

First proposed by Neven et al. [349], the QBoost algorithm is an ensemble model comprising of a set of weak, i.e. simple low-depth, Decision Tree (DT) classifiers, also called *learners*, optimally combined to build a strong classifier.

The workflow of the algorithm starts with a boosting procedure, based on the standard Adaboost algorithm [350, 351]. A set of N weak learners h_i is classically trained in a sequential fashion on the training set $T = \{\boldsymbol{x}_s, y_s\}_{s=1}^{S}$ where $y_s \in [-1, 1]$ is the classification label of the data point $\boldsymbol{x}_s$. Initially, the first weak learner is trained such that all the data points are weighted uniformly, using the constant distribution $D_{i=1}(s) = \frac{1}{S}$. Each weak learner h_i is then iteratively trained on the same training set where the data points are however weighted differently based on an updated distribution $D_i(s)$. This latter distribution is re-computed after the training of each weak learner. More precisely, it depends on the quantity ε_i that considers the misclassified points by the previous weak learner:

$$\varepsilon_i = \sum_s \mathbb{1}\left[h_i(\boldsymbol{x}_s) \neq y_s\right] D_i(s). \tag{8.1}$$

From ε_i, one computes the quantity $\alpha_i = \ln(1/\varepsilon_i - 1)/2$ used as exponential coefficients to update the data distribution

$$D_{i+1}(s) = \frac{1}{Z_i} D_i(s) \cdot e^{-\alpha_i y_s h_i(\boldsymbol{x}_s)}. \tag{8.2}$$

with a normalisation factor Z_i such that D_{i+1} remains a probability distribution.

After the entire ensemble of weak learners $\{h_i\}_{i=1}^{N}$ has been trained, a strong classifier $\mathcal{H}$ is built by selecting a subset of the weak learners. The optimal combination is obtained through the optimisation of binary weights $\boldsymbol{w} \in \mathbb{B}^N$ that minimise the following cost function

$$C(\boldsymbol{w}) = \sum_s \left(\frac{1}{N} \sum_i^N w_i h_i(\boldsymbol{x}_s) - y_s \right)^2 + \lambda \|\boldsymbol{w}\|_0, \tag{8.3}$$

where w_i is the i-th binary weight and $h_i(\boldsymbol{x}_s) \in [-1, 1]$ is the prediction of the i-th weak learner for the data point $\boldsymbol{x}_s$. A regularisation term parameterised by λ helps to favour better generalisation of the model on new data by penalising too complex ensembles with many weak learners.

Expanding the squared term in the above equation and dropping the constant terms, which are irrelevant to the minimisation problem, we can reformulate the cost function as

$$C(\boldsymbol{w}) = \sum_{i,j}^{N} w_i w_j \mathrm{Corr}(h_i, h_j) + \sum_{i}^{N} w_i(\lambda - 2\mathrm{Corr}(h_i, y)), \tag{8.4}$$

with $\mathrm{Corr}(h_i, h_j) = \sum_s h_i(\boldsymbol{x}_s)h_j(\boldsymbol{x}_s)$ and $\mathrm{Corr}(h_i, y) = \sum_s h_i(\boldsymbol{x}_s)y_s$. On the one hand, the weak classifiers whose outputs correlate well with the labels cause the second term to be lowered via $\mathrm{Corr}(h_i, y)$. On the other hand, via the quadratic part $\mathrm{Corr}(h_i, h_j)$ describing the correlations between the weak classifiers, pairs of strongly correlated classifiers increase the value of the cost function, thereby increasing the chance for one of them to be switched off. This is in line with the general paradigm of ensemble methods for promoting a diversification of the ensemble in order to improve the model generalisation on unseen data.

Once the optimisation of Eq.(8.4) is performed (see section 8.2.3), the strong classifier $\mathcal{H}$ can be built using the weights minimising C, w_i^{opt}. Given a new data point $\boldsymbol{x}$, we infer a classification prediction by:

$$\mathcal{H}(\boldsymbol{x}) = \mathrm{sign}\left(\frac{1}{N}\sum_{i}^{N} w_i^{\mathrm{opt}} h_i(\boldsymbol{x}) - y_{\mathrm{opt}}\right) \tag{8.5}$$

where y_{opt} is an optimal threshold that enhances results as proposed in [352] and is computed as a post-processing step

$$y_{\mathrm{opt}} = \left(\frac{1}{S}\sum_{s}^{S}\frac{1}{N}\sum_{i}^{N} w_i^{opt} h_i(\boldsymbol{x}_s)\right). \tag{8.6}$$

8.2.2 - QBoost-inspired classifier

An important challenge in designing a successful ensemble is to ensure that the base learners are highly diverse, *i.e.*, that their predictions do not correlate too much with each other. The initial idea of QBoost was to accomplish this by using weak learners of the same type, specifically Decision Trees (DT) classifiers and train them sequentially using the boosting procedure. Another way is to use different types of base learners [353], creating an heterogeneous ensemble with a mix of different learners including, e.g., DT, Logistic Regression (LR), k-Nearest Neighbours (kNN), and Gaussian Naive Bayes (GNB) [354]. Having inherently different mathematical foundations, these learners can exhibit significantly different views of the data landscape.

For this specific problem, we find that a classifier based on a heterogeneous ensemble comprising different types of learners can lead to better generalisation performance than the plain-vanilla model with DT only. The choice of type and mixing of such an heterogeneous ensemble is motivated by comparing results from extensive simulations, both with one type of learner as well as with combinations thereof. Each of these models are trained on the under-sampled version of the training set and the corresponding prediction performance are obtained on a separate cross-validation set, using Precision and Recall. As displayed in Fig. 8.1, DTs perform better in recall while kNNs perform better in precision. Heuristically, combining these two types of base learners results in the actual best performing model over any other tested combinations.

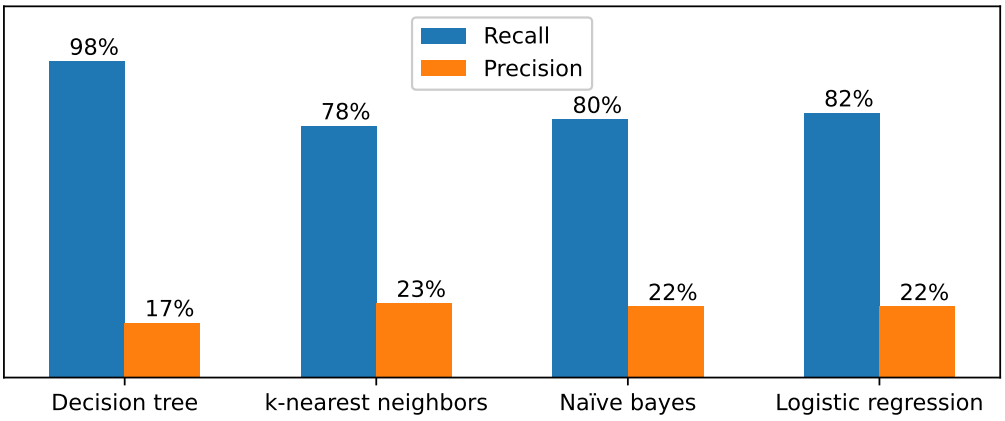


Figure 8.1: **QBoost performances as per type of base learners.** Comparison of Recall (blue) and Precision (orange), obtained using QBoost with different base learners including decision trees, k-nearest neighbours, gaussian naive bayes and logistic regression.

Furthermore, in order to take advantage of the historical structure of the data, where multiple historical data points for the same companies are available, we propose to train each of the learners of the heterogeneous ensemble on different historical periods of the dataset. Specifically, using dates features in the dataset, the raw training set is split into subsets and then subgroups of learners are trained on them. This ensemble-training procedure based on subsampling is expected to further diversify the ensemble, where the weak learners are trained independently on the different economic recession and expansion periods underlying the training dataset. Additionally, it reduces the training time significantly as each subgroup of learners is trained on a subset of data. The learners trained in this way can potentially capture different views of the data, resulting in a better diversification of the ensemble.

Here we propose this approach with two variations:

1. *Boosting.* Following [349], we train each of the ensembles on the different subsamples of data with the sequential boosting procedure described in section 8.2.1. Generally, the learners can exhibit negative correlations among each other.

2. *Subsampling.* We train each of the ensembles without sequential boosting, relying only on subsampling for diversification. Generally, the learners exhibit only positive correlations among each other.

8.2.3 - Optimisation of the ensemble via QUBO solving

As explained in section 8.2.1, the weak learners obtained during the ensemble training are then optimally combined to form a stronger classifier. Finding the best binary weights $\boldsymbol{w}$ for this combination amounts to the minimisation of the cost function given in Eq.(8.4). Since the weights w_i are binary variables, that is, $w_i^2 = w_i$, we reformulate the cost function C as C_Q:

$$C_Q(\boldsymbol{w}) = \boldsymbol{w}^T Q \boldsymbol{w} = \sum_{i,j}^{N} Q_{ij} w_i w_j, \tag{8.7}$$

where

$$Q_{ij} = \begin{cases} \text{Corr}(h_i, h_j) & \text{if } i \neq j \\ \frac{S}{N^2} + \lambda - 2\text{Corr}(h_i, y) & \text{else.} \end{cases} \tag{8.8}$$

This second formulation is written in the form of a *Quadratic Unconstrained Binary Optimisation* (QUBO) problem [355]. Solving a QUBO problem amounts to finding the minimum of a quadratic polynomial of bit variables, *i.e.*, the optimal bitstring minimising the cost function C_Q, with $Q \in M_N(\mathbb{R})$, the symmetric matrix encompassing the correlations between the weights to optimise.

As the number of learners grows, the classical optimisation of the weights is performed over an exponentially increasing space, thus opening the door to potentially more efficient quantum methods. Indeed, some results suggest that quantum computers could better approximate *sufficiently good* solutions (as defined in Section 8.3.3) in a shorter time compared to classical computers. This expectation partly stems from the fact that quantum computers may offer shortcuts through the optimisation landscape inaccessible to traditional classical simulated annealing methods [356]. In our case, if one can produce a state such that the probability amplitudes peak in low-cost bitstrings, sampling from it becomes an efficient way of optimising the weights.

8.2.4 - QUBO solving with VQAs

One crucial ingredient for solving QUBO-like combinatorial problems using neutral atom quantum processors lies in the ability to implement a custom cost Hamiltonian $\hat{C}_Q$ on the quantum processor which should be closely related to the cost function C_Q. When this Hamiltonian is generated exactly, the iterative methods mentioned in chapter 7 can ensure that after k iterations the evolution of a quantum system subjected to $\hat{H}_{ctrl}^{(k)}(t) + \hat{C}_Q$ will tend to produce low energy states $|\boldsymbol{w}\rangle$, *i.e.* solutions with low values $C_Q(\boldsymbol{w})$. There are ways to compute the evolution over $\hat{C}_Q$ using circuit-based quantum computers [254], or special-purpose superconducting processors like D-wave machines [357].

For analog neutral atom technology, innately replicating the cost Hamiltonian is tackled with the position dependent interaction part $\hat{H}_{\text{int}}$ of the resource Hamiltonian as explained in chapter 6. As an illustration, we examine a collection of QUBO problems, each comprising $N = 10$ variables, and explore the production of solution states using progressively more complex QAOA protocols. As we increase the depth p of the QAOA, we anticipate a corresponding enhancement in optimisation efficiency, with the normalised cost function $\langle \hat{C}_Q \rangle \in [0, 1]$ gradually converging towards 0. This trend arises as the optimised driving protocols generated by the deeper and deeper QAOAs tend to populate states that minimise more the associated QUBO. For each QUBO instance, we benchmark two distinct embedding techniques, as depicted in Fig. 8.2(a): the 10 qubits are either located at fixed positions on a rescaled triangular lattice or optimally positioned in free space using the Multidimensional

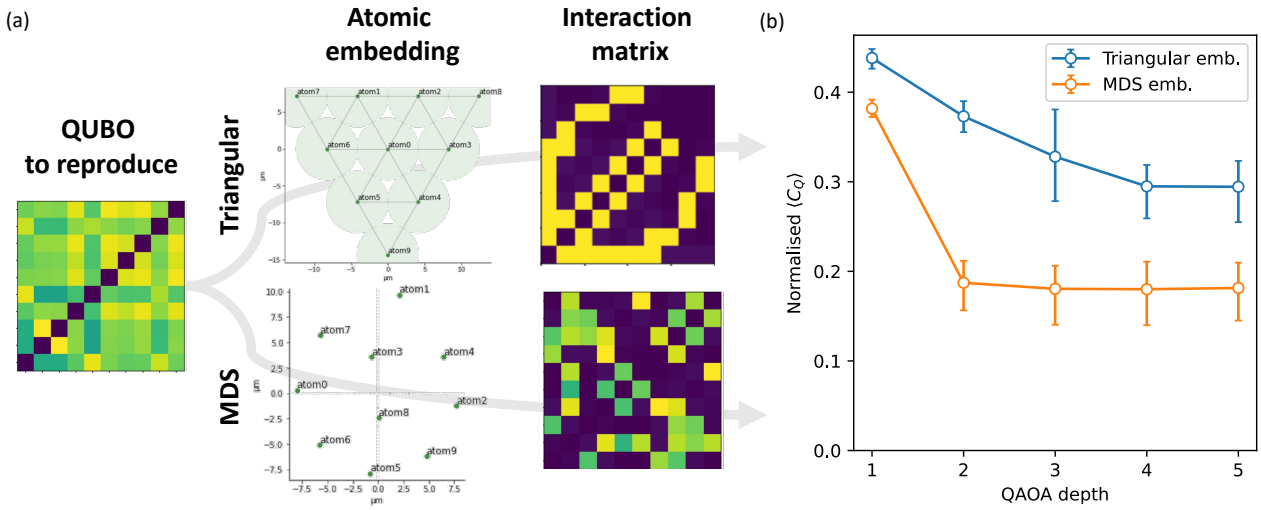


Figure 8.2: **Limitation of QAOA protocol on embedded QUBOs.** (a) A QUBO represented by a symmetric matrix is mapped to atomic registers using either triangular embedding or MDS optimised positioning. The resulting interaction matrices are also plotted with normalised colour bars. (b) Performance comparison with QAOA of increasing depth for the two types of embedding presented in (a).

Scaling (MDS) protocol described in Box. 27. Despite employing an optimal embedding strategy, the interaction matrix falls short of replicating all QUBO features as conveyed by the large discrepancies between the displayed QUBO matrix and interaction counterpart. A perfect match between the two necessitates satisfying a set of constraints that grows quadratically with variable count. Consequently, any embedding can only partially approximate $\hat{C}_Q$, typically leading to $\hat{H}_{\text{int}}$ providing an incomplete representation of the desired cost function.

Fig. 8.2(b) illustrates the distribution of scores obtained with both embedding when increasing the depth of QAOA. Although expanding the parameter space of QAOA improves distribution quality in both methods, convergence appears swifter with the MDS embedding. Furthermore, distributions obtained with MDS embedding demonstrate lower costs compared to those with triangular embedding. However,the mean of normalised $\langle \hat{C}_Q \rangle$ across the QUBO set eventually reaches a nonzero plateau in both cases. This plateau arises from the inherent discrepancy between the interaction Hamiltonian and the desired cost Hamiltonian, limiting further enhancements through additional layers of QAOA. Therefore, the quality of the solutions sampled from an optimised quantum state will be limited under a straightforward implementation of a variational quantum algorithm.

In addition, the optimisation of variational quantum algorithms [358] usually requires diagnosing the expressibility and trainability of several pulse sequences/circuits in order to trust that low energy states are being constructed. Moreover, at each iteration, to obtain a statistical resolution of the energy of the prepared state with given precision, one usually requires a number of samples scaling as the inverse of the square of said precision. Given the current low repetition rate of neutral atom based QPUs (of the order of $1-5$ Hz), running variational algorithms for each problem instance may require several tens of hours of operation on robust hardware. With current technology, it is therefore

Box 27: Free-space mapping of QUBO with Multidimensional Scaling

An alternative to the classical optimisation scheme described in Sec. 6.4.2 consists in relying on existing ML algorithms dedicated to non-linear dimensionality reduction, labelled as *manifold learning* procedures. These methods operates on datasets with artificially high dimensionality, i.e. a large number of features, some of which are not really descriptive. Reducing this number often eases the visualisation of the data and we can apply those algorithms, for example Multi Dimensional Scaling (MDS) [a], to map an artificial distance matrix obtained from a QUBO instance to a set of positions in a plane. More specifically, MDS seeks a low-dimensional representation of the data in which the new distances respect well the original distances in the high-dimensional space. From a QUBO $Q \in M_N(\mathbb{R})$, one can construct the original distance matrix D with terms $D_{ij} = (C_6/Q_{ij})^{1/6}$ for $i \neq j$. Feeding D to MDS with or without an initial guess of the positions outputs optimised coordinates of N points in a $2-$dimensional space, used to arrange in free-space the atoms in order to reproduce $\hat{C}_Q$ with $\hat{H}_{\text{int}}$ as much as possible. As no constraints are typically included in such algorithm, the positions can be rescaled afterwards such that the minimum distance between two atoms remains greater than d_{min} μm. If atoms stray too far from the centre, indicating they shouldn't interact with the rest of the register, they can be relocated within bounds, away from nearby atoms.

[a] See scikit-learn for details

crucial to employ methods to generalise optimised protocols as already highlighted in chapter 7 or methods involving only a small budget of cycles and that can quickly provide significant solutions to the QUBO problem.

8.3 - Random Graph Sampling

Stepping away from the variational paradigm of QAOA and QAA, we devise a sampling algorithm that exploits the stochastic loading probability of neutral atom QPU in order to probe efficiently the solution space of a QUBO. This algorithm is faster to implement as it does not require iterative processes such as atom rearrangement or closed loop communication between a classical optimiser and the quantum hardware. We thus introduce the Random Graph Sampling (RGS) method which builds up on the randomness of the atomic loading process. This procedure has allowed us to sample solutions of QUBOs of sizes up to 50 as demonstrated in the following section.

8.3.1 - Leveraging stochastic atomic loading

In neutral atom QPU, atoms are spatially arranged by combining the trapping capacity of optical tweezers with the programmability of a SLM (see Sec. 1.1.1). By the means of those two devices, atoms can be individually trapped in arbitrary geometries. Once an atomic cloud has been loaded, each of the N_t traps is randomly filled with success probability p_{fill} according to a binomial law $\mathcal{B}(N_t, p_{\text{fill}})$. Thus, one must set up around $N_t = N/p_{\text{fill}}$ traps to maximise the probability of trapping N atoms. A rearrangement algorithm then moves atoms one at a time to the wanted positions using a moving tweezer. The excess atoms are released mid-stroke to end up with a register of N correctly positioned

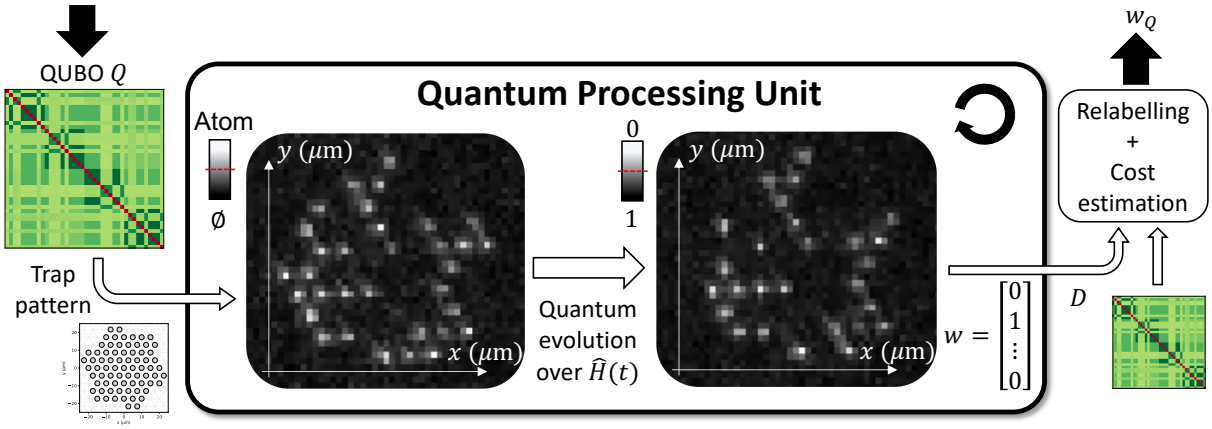


Figure 8.3: **Random Graph Sampling pipeline for solving a QUBO Q on a neutral atom based QPU.** First, a QUBO, here with negative weights on the diagonal (red) and positive weights outside (green scale), is taken as input. From this QUBO, a trap pattern is devised and sent to the QPU, as well as the wanted number of repetitions and the pulse sequence used. At the beginning of each cycle, a first fluorescence picture enables to identify which traps were filled by an atom (bright spot). The system evolves according to $\hat{H}(t)$ and a final picture is taken to measure the collapsed state of the system, outputting a bitstring $\boldsymbol{w}$. Using Q, we select w_Q among D, the distributions of bitstrings obtained from repeating this process several times.

atoms. However, by skipping the rearrangement step, we obtain samples from an essentially random sub configuration of the underlying pattern of traps. For $N_t = 2N$, the number of possible configuration of size N scales as $\binom{2N}{N} \sim 4^N/\sqrt{N}$, offering a large variety of $\hat{H}_{\text{int}}$ for a devised pattern. In order to produce the quantum distributions from which we sample the QUBO solutions, we repeatedly apply a fixed sequence of constant pulses to the atoms. The latter evolve under $\hat{H}(t)$ according to their interactions, which are set at each cycle by the atom random positions.

The QUBO to solve, Q, first acts as a resource to design the trap pattern (N_t, shape, spacing) sent to the QPU (see Fig. 8.3). Once a chosen budget of samples has been acquired on the QPU, we are left with a bitstring distribution D. Using again the QUBO, we apply a relabelling procedure described in the following paragraph to each bitstring according to both Q and the related atom positions. This optimisation procedure is designed to scale only linearly with N and is tasked to search for a way of labelling the atoms from 1 to N which minimises for each repetition the difference between $\hat{C}_Q$ and $\hat{H}_{\text{int}}$. Finally, we compute the bitstring corresponding to the optimised weights for the ensemble of learners considered.

8.3.2 - Optimised relabelling process

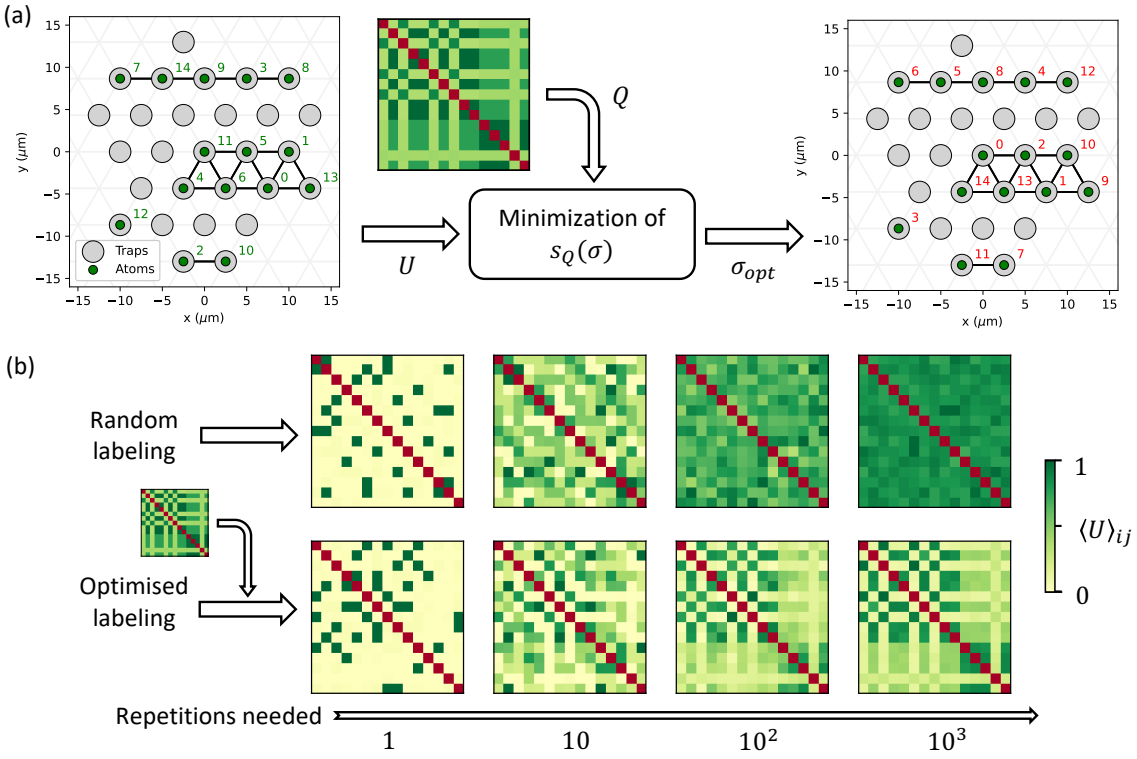


Figure 8.4: **Classically optimised relabelling used in RGS.** (a) 15 atoms (green dots) are filling a fraction of a triangular trap layout (gray circles). Each atom is randomly labelled (green) and from their positions an interaction matrix U is derived. Using the QUBO to solve, Q, and Eq. 8.9, a permutation of the labels σ_{opt} is found. The atoms are relabelled (red) such as the resulting interaction matrix replicates better Q. (b) Normalised interaction matrices obtained when averaging over many repetitions of traps loading. While the random labelling (top line) produces a uniform matrix, the optimised labelling (bottom line) enables to access some features of the QUBO at each cycle, producing an average matrix resembling Q.

We describe here the relabelling process used in RGS. For a given cycle where N traps out of N_t are filled with atoms, a first measure of the system before the quantum processing part enables to locate the atoms, as shown in Fig. 8.4(a). The latter are randomly labelled and this arbitrary labelling usually orders the bitstring measured after the quantum processing. However, we can choose another labelling more specific to the QUBO we want to solve. This post-processing step determines a labelling σ_{opt} of the atoms such that the resulting interaction matrix better reproduces the QUBO matrix than the one obtained from the randomly generated graph. For each way of labelling the atoms from 1 to N, *i.e.* each permutation of length N, we compute the separation

$$s_Q(\sigma) = \sum_{i<j} \| U(r_{\sigma(i)\sigma(j)}) - Q_{ij} \|, \tag{8.9}$$

where Q is the QUBO matrix, σ, a permutation of length N and $U(r_{\sigma(i)\sigma(j)})$ the interaction term

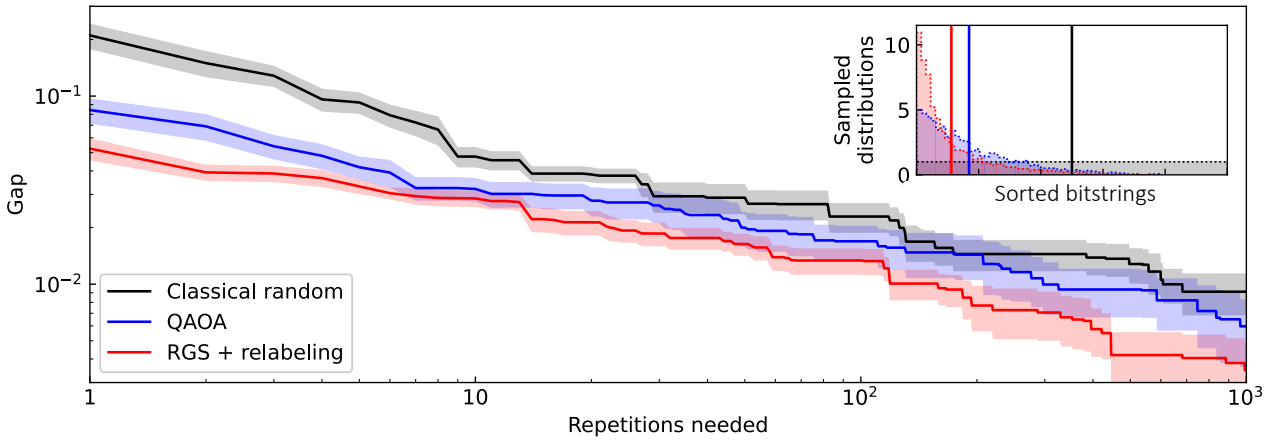


Figure 8.5: **Benchmark of RGS method**. Evolution of the gap obtained after a number of repetitions with a classical random search (black), a numerically simulated QAOA (blue) and numerically simulated RGS (red) with QUBO dependent labelling (averaged over 10 random QUBO instances of size $N = 15$). The inset shows the distributions (histograms) and the mean (solid vertical lines) of the solutions found with each method, ranked by their cost function value (better bitstrings being on the left).

between atoms originally named i and j. The two matrices are normalised to allow a proper comparison. We perform a random search with fixed budget n_{iter} over the $N!$ possible labelling permutations. The permutation minimising s_Q is then used to read out the measured bitstring. Searching for such a permutation is reasonably fast for the sizes that can be loaded in the QPU. We have checked that for $N \leq 100$, this takes less than $n_{\text{iter}} \times 2$ ms to converge. In the following, we set $n_{\text{iter}} = 10N$ so as to scale only linearly with the number of qubits and not as $N!$. This may not be sufficient to identify the best permutation at larger sizes, but it remains enough to reproduce some of the features of the QUBO at each cycle as shown in Fig. 8.4(b). Furthermore, on average, the whole QUBO is much better represented with this optimised relabelling step than simply using a random permutation. It is worth pointing out that this optimisation step can be done retrospectively, after the quantum data has been acquired, as long as we have access to the initial traps filling. Thus, its execution time does not limit the duration of a cycle, and this can effectively become a post-processing step done on a classical computer.

8.3.3 - Benchmark against numerically simulated QAOA

While RGS offers no theoretical guarantee of sampling a global minimum of the cost function, we can still expect to output bitstrings with low function value. To characterise a solution, we introduce the *gap* of a bitstring $\boldsymbol{w}$ defined as:

$$\text{gap} = \left| \frac{C_Q(\boldsymbol{w}) - C_Q(\boldsymbol{w}_Q^0)}{C_Q(\boldsymbol{w}_Q^0)} \right| \tag{8.10}$$

where $\boldsymbol{w}_Q^0$ is the best solution found by a classical algorithm. This solution $\boldsymbol{w}_Q^0$ is not guaranteed to be the best possible, but acts as such for benchmark purposes. Reaching a gap of 0 amounts to having

found the best solution provided by the benchmark. Note that for small sizes of N, it is possible to use an exhaustive search as a benchmark and $\boldsymbol{w}_Q^0$ is in that case the theoretical best solution. For larger cases, $\boldsymbol{w}_Q^0$ is returned by a state-of-the-art simulated annealing (SA) algorithm which is given a large amount of repetitions (200 000 in this work). Finding a bitstring with a gap below 1% for instance, means finding a solution with a cost close to 1% of the optimal one, which, in many operational problems such as our case study, is often considered *sufficient*. We check that for the various sizes considered, the difference between selecting a 1% solution and the optimal one, *i.e.* with a gap of 0) is reflected in the classification model with variation of precision P smaller than the standard deviation obtained on the QUBO set. We thus consider as a good enough solution a bitstring with gap below 1%.

We benchmark the RGS approach on a set of randomly generated QUBOs of size 15 and compare it to a classical uniform sampling of $\mathbb{B}^N$ and to a numerical simulation of QAOA using Pulser, all using a similar budget of 1000 cycles, or measurements. Getting into the detail, the QAOA algorithm is allowed 10 iterations with 100 cycles each in order to optimise the duration of 3 pulses. The cost function evaluated at each iteration is $\langle C_Q \rangle$ averaged over the 100 measurements. The atoms are positioned using the MDS embedding method, meaning that an experimental implementation would require to rearrange them after the loading, lengthening the duration of each cycle. In contrast, for RGS, the positions are random at each cycle while the pulse sequence remains the same, 3 pulses with fixed durations independent of the QUBOs. This sequence can be pretrained classically on numerical simulations. We show the results of these three methods in Fig. 8.5 with both the convergence of each one with respect to the number of cycles performed and the aggregated bitstring distributions sorted by increasing values of C_Q. Not only does RGS converge faster, achieving a gap of less than 1% with three times fewer cycles than QAOA, it also produces, on average, sampled distributions with greater concentration on bitstrings with low value. For this set, a bitstring sampled using RGS+relabelling is on average in the 11 best % of $\mathbb{B}^N$ while one sampled with QAOA is in the 17 best %.

8.4 - Experimental implementation

8.4.1 - QUBO solving results

We experimentally implement the RGS method described in section 8.3 to solve 5 sets of QUBOs ranging in size from $N = 12$ to 50 qubits and compare its performance with the SA algorithm used to define $\boldsymbol{w}_Q^0$. The methods are always compared for similar budget of cycle repetitions and we can assess the quality of the bitstring distribution or the scalability of each approach.

QUBOs of one set of fixed size being produced by repeatedly applying the QBoost subsampling approach on the same dataset, they only exhibit minor discrepancies between their structure and range of values. Therefore, we can, within reasonable approximation and for faster implementation on the quantum hardware, only use one trap pattern per set. Considering a loading probability $p_{\text{fill}} = 0.55$, we design a triangular pattern with $N_t = 40/0.55 \approx 73$ traps for QUBOs of size 40 (see Fig. 8.3) and similarly with 91 traps for QUBOs of size 50. This choice is motivated by both available trapping laser power and maximisation of number of samples at sizes 40 and 50. The spacing of the regular pattern and thus the atomic interaction in the array is chosen in combination with the maximum value of $\Omega(t)$ reached during the pulse sequence. Having Hamiltonian terms Ω and U of comparable magnitude in

$\hat{H}(t)$ enables to explore the interacting regime.

In addition, we can reuse statistics acquired at large sizes and extract distributions of bitstring of smaller size as explained in Box. 28. In essence, by neglecting the interaction between pairs of atoms separated by more than the lattice spacing, a N-atom regular array can be divided into smaller clusters with similar regular shape and regions seemingly isolated from each other. Applying this method to the original distribution of $\sim 65\,000$ measurements, ranging in size from 34 to 66 atoms, outputs a wider distribution of $\sim 334\,000$ bitstrings ranging in size from 1 to 66 as depicted on Fig. 8.6. In this implementation, since we only consider QUBOs output by the subsampling approach detailed in Sec. 8.2.2, all of them are similar in structure, being produced by the same dataset and with the same hyperparameters for weak learners ensemble generation. We apply the relabelling step to the extracted distribution in order to solve the considered sets of QUBOs.

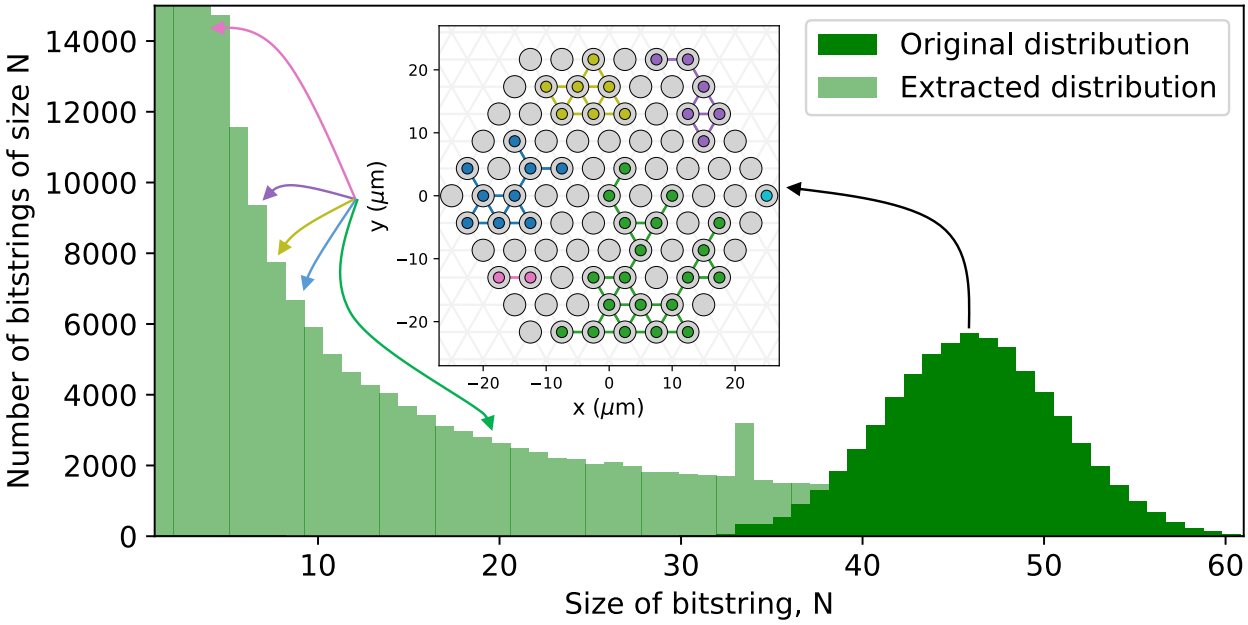


Figure 8.6: **Clustering of atomic configurations** to extract n-sized bitstrings from N-sized ones with $N > n$. From an original distribution of 65 000 bitstrings (dark green), we construct a larger distribution of 334 000 bitstrings (light green). A bitstring of size 45 has been measured with the atomic configuration displayed in the inset. Atoms are sorted between clusters (various colors) of sizes 2, 6, 7, 9, 20. and the initial bitstring is cut into 5 smaller bitstrings, usable to solve QUBOs of corresponding sizes.

The results obtained by RGS with relabelling are showcased both in terms of convergence to low cost value $C_Q(\boldsymbol{w})$ solutions (see Fig. 8.7(a)) and scalability of the method with respect to the complexity of the problem, *i.e.* the QUBO size (see Fig. 8.7(b)). The classical random method consisting in uniformly sampling with replacement bitstrings from $\mathbb{B}^N$, it scales exponentially with N. In contrast, the RGS algorithm shows better performances, already finding solutions with a gap smaller than 10% after only a few repetitions. Looking at the number of repetitions needed to go below 1% with respect to N, a log-log linear fit returns a scaling in $0.2 \times N^{1.55}$. Since the QPU run-time

Box 28: Configuration clustering by NN approximation

We elaborate on how to extract usable bitstrings of size n from ones of size $N > n$ measured on the quantum set up. Those smaller bitstrings can in specific cases be used to solve QUBOs of size n. At each computation cycle, a pattern of N_t traps is filled by $N \sim \mathcal{B}(N_t, p_{\text{fill}})$ atoms. Many cycles would then produce bitstrings whose sizes follow a Gaussian distribution centred around $N_t\, p_{\text{fill}}$ as shown in Fig. 8.6. For each cycle, knowing which traps were filled (as shown in inset), we can isolate cluster of atoms with the following rule: two atoms belong to different clusters if their pair distance exceeds the pattern spacing. Therefore, due to the rapid decay of the interaction with the distance, *i.e.* $U(r_{ij}) \propto r_{ij}^{-6}$, we can consider that clusters do not interact between them. Indeed, here, two non neighbouring atoms are interacting at least 27 times weaker than a pair of neighbouring ones. Segmenting a N-sized bitstring leads to the extraction of s smaller bitstrings with sizes n_i such that $\sum_i^s n_i = N$. This method produces bitstrings obtained from fully nearest neighbour interacting systems, as no atom remains isolated. However, it can reduce the number of measurements made at a large size N. The resulting bitstrings can only be used to solve QUBOs of corresponding sizes and which would have produced the same trap pattern as the one used to acquire the original distribution.

scales linearly with the number of cycles, the quantum optimisation actual duration is also expected to scale as $N^{1.55}$. Comparing RGS to the SA algorithm, we observe better performance of the latter at small sizes but more and more comparable performance at increasing sizes. In the case of $N = 40$, this specific implementation of RGS finds on average a gap below 0.2% after 150 repetitions while SA needs around 4 times more cycles. For $N = 50$, the mean gap achieved after hundreds of cycle is around 1.%. Overall, RGS with relabelling applied to QUBOs produced by the subsampling-based classifier exhibits similar behaviour with the state-of-the-art SA algorithm.

8.4.2 - Classification results

In this section, we present the classification results obtained using the quantum classifier based on subsampling (see Section 8.2.2), trained using the quantum optimiser implemented on the QPU up to 50 qubits. This quantum classifier, leveraging the subsampling approach without boosting, is based on the optimisation of QUBOs with positive off-diagonal values, amenable to efficient optimisation with the current quantum hardware (see Section 8.4.1). We find the best results for 50 qubits, corresponding to an initial weak ensemble of 50 learners, whose percentages of kNNs and DTs have been optimally chosen through a hyperparameter optimisation procedure. For this hyperparameter optimisation, the training set was split into 80% training and 20% cross-validation sets using stratified-shuffled splitting.

Our proposed classifier is able to achieve very similar performances to the classical RF algorithm as conveyed by comparing the confusion matrices depicted in Fig. 8.8(a). Using bitstrings with gap below 1%, our model (a2) reaches $P = 2585/(2585 + 6681) = 27.9 \pm 0.09\%$, closely approaching the benchmark threshold (a1) $P = 28.0 \pm 0.07\%$ for the same recall value of $R = 83\%$. The proposed model reduces a bit the number of f_n by increasing its tendency to label instances as positive, i.e. fallen angel. Very interestingly, this result is obtained with only 50 initial learners compared to the RF ensemble of 1200 learners. The difference in the number of learners employed is of great relevance

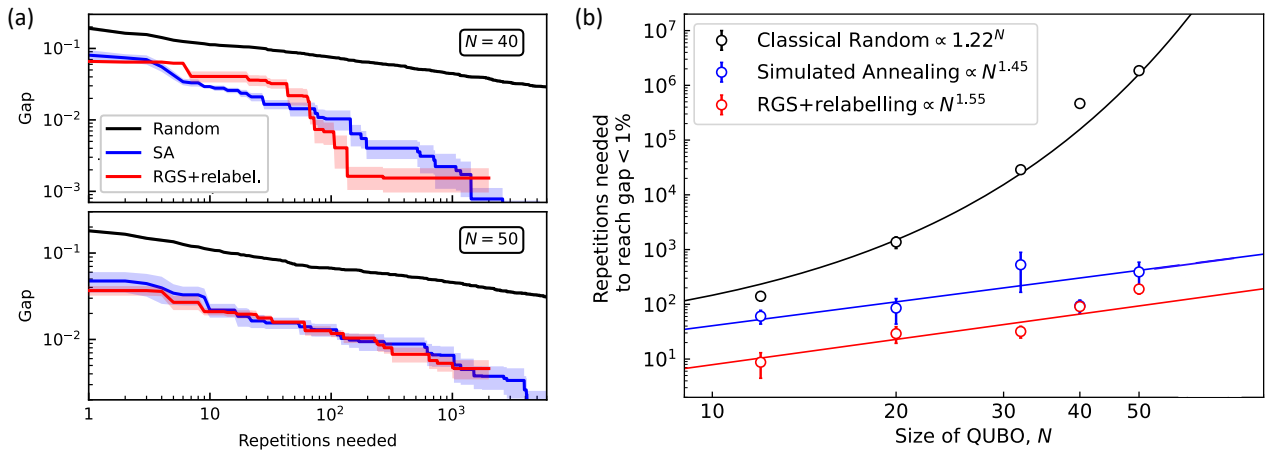


Figure 8.7: **Experimental results for QUBO solving.** (a) Gap convergence obtained with classical uniform sampling (black), Simulated Annealing sampling (blue) and RGS sampling with optimised relabelling (red) for increasing size of QUBOs. The best gap found after some cycle repetitions is averaged over sets of 5 QUBOs (plain line). (b) Scaling analysis of the number of repetitions needed to reach a gap below a threshold of 1% with respect to problem size. The results obtained by the three mentioned methods at sizes $N = 12$, 20, 32, 40 and 50 (dots) are fitted either exponentially or polynomially (line) depending on the best match.

for the interpretability of the model. Indeed, the decision for a new unseen point outputted by the model can be traced back more easily and better understood by the user. The best results for 50 qubits were obtained with a total runtime of around 50 minutes, against a total runtime of more than 3 hours for the classical benchmark, representing a relevant practical speed-up.

Fig. 8.8(b) exhibits the scaling of the mean classification performance for the subsampling approach performed on the QPU and the boosting approach simulated with TN. This latter model, being based on the boosting procedure, leverages the optimisation of QUBOs with negative off-diagonal values which cannot be currently directly optimised on neutral atom QPU. It can be seen that even at low values of qubits/learners, the proposed model based on boosting, already showed the same level of performance as the RF with 1200 trees. With 90 learners, it shows a mean precision score of about $P = 29\%$ (reducing the false positives by 1%) corresponding to the recall of 83%. The best results for 90 qubits/learners present a total runtime of the order of 20 minutes against the total runtime of more than 3 hours for the classical benchmark, attaining also in this case a relevant practical speed-up.

Based on the scaling projections, it can be argued that this type of model is expected to remain the best performing one. It can be seen in the inset of Fig. 8.8 that a crossing with the experimentally realised classifier based on subsampling could occur for a large number of qubits, around 380, although it is difficult to assess the reliability of the extrapolation done at those limited number of qubits.

The algorithm proposed in this chapter comprises a hybrid classical-quantum classification model based on QBoost, tested on a neutral atom quantum platform and benchmarked against Random Forest, a classical ML model used in the Finance industry. These results were obtained leveraging the

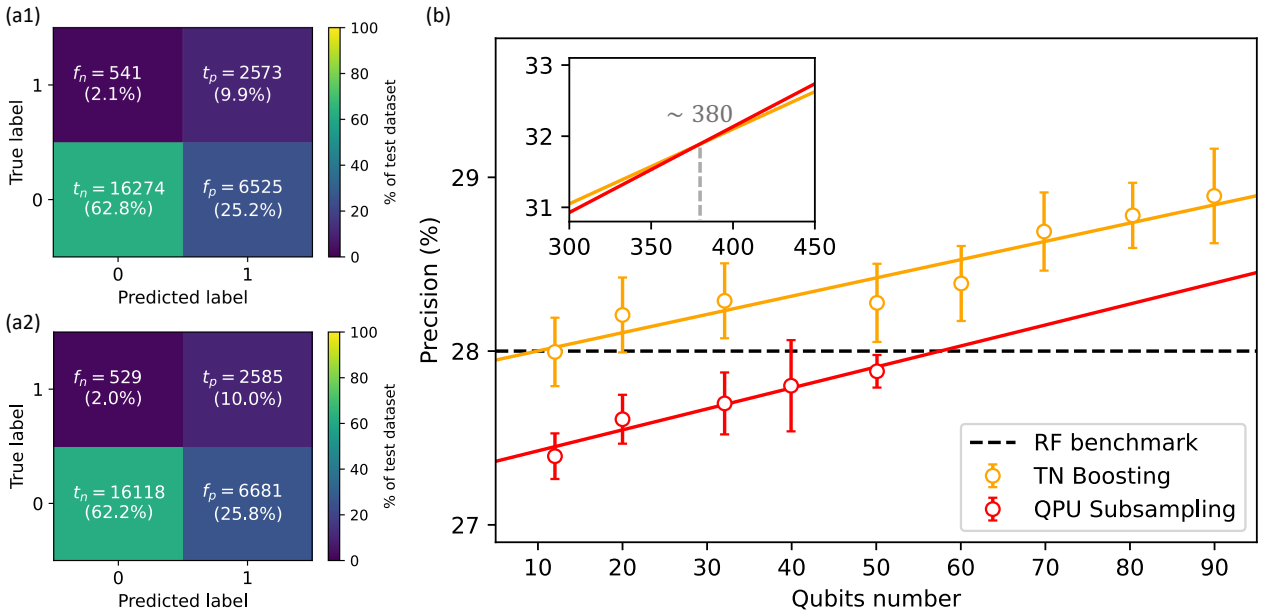


Figure 8.8: **Classification results.** (a) Confusion matrices obtained with (a1) the RF benchmark algorithm and (a2) the QPU subsampling approach at 50 qubits. Each proportion of $t/f_{p/n}$ is normalised. (b) Scaling of precision P of various proposed quantum classifiers with respect to the number of qubits, keeping $R = 83\%$. The subsampling approach (red) is implemented on QPU (filled dot) between 12 and 50 qubits. The boosting approach (yellow) is implemented using Tensor Networks between 12 and 90 qubits. The best performance of the RF classifier acts as threshold (dashed). The error bars represent the variability in corresponding performance across 5 iterations/QUBOs. Scaling projections are obtained by linear extrapolation (solid lines).

hardware-tailored Random Graph Sampling method to optimise QUBOs up to size 50. This method showed similar performances with Simulated Annealing approach and was able to provide solutions to QUBO within acceptable repetitions budget. We report that the proposed classifier trained on QPU achieved competitive performance with 27.9% precision against the benchmarked 28% precision for the same recall of approximately 83%. However, the proposed approach outperformed its classical counterpart with respect to interpretability with only 50 learners employed versus 1200 for the Random Forest and comparable runtimes. A variety of similar problems appearing in the finance industry can be tackled by quantum-enhanced algorithms as interpretability and performance improvements for real case scenarios with complex and highly imbalanced datasets are pressing issues. In order to gain practical advantage over fully classical methods, one need to make the best possible use of the available neutral atom technology, including quantum blocks into well known classical frameworks, in hybrid approaches such as the ones presented below.

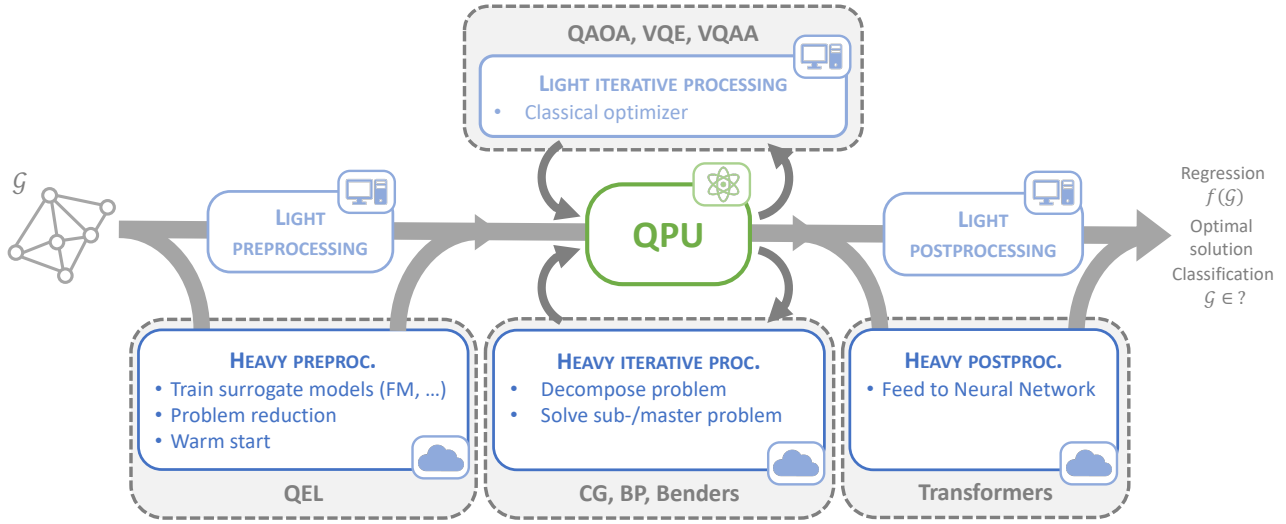


Figure 8.9: Schematics for possible hybrid algorithms. These algorithms require heavy classical computation, either at preprocessing (left) in quantum extremal learning framework, in iterative processes (center) as detailed in 8.5.1 or at post-processing (right) as detailed in 8.5.2.

8.5 - Review of hybrid quantum-classical algorithms

The prevailing belief that "classical algorithms" exhibit in general exponential resource consumption with problem size overlooks many efficient methods developed to simplify large-scale instances. State-of-the-art classical solvers employ extensive pre-processing and heuristics rather than brute force methods to efficiently tackle such problems. For example, branch-and-bound (B&B) methods [359] divide the input problem into smaller problems and employ bounds to eliminate sub-problems that cannot lead to an optimal solution. *Kernelisation*, on the other hand, reduces the size of the input problem while preserving the solution and helps to distil the essential features into a more manageable size [360]. Methods such as B&B and kernelisation therefore offer an elegant framework to limited quantum processors. They enable quantum processors to handle limited-size components of a problem while delegating the pre-processing and the aggregation of sub-solutions to classical processors. Implementing "non-native" hybrid schemes has recently been proposed on neutral-atom quantum computers [361]. In the following, we present how hybrid approaches can help tackle optimisation and ML tasks.

8.5.1 - Hybrid optimisation approaches

In the realm of classical optimisation problems, the quantum optimisation framework is seldom a standalone solution but rather a module integrated in a hybrid workflow. Solving a combinatorial problem already embedded into an atomic register amounts to an optimal control task and can be tackled by variational quantum algorithms [358] as explained in chapter 6 or quantum neural networks [362]. The algorithm involved may rely on a set of parameters, such as pulse shape or circuit layer, which when optimised gives an efficient protocol to drive the quantum system to the correct solution state. In this context, as shown on top of Fig. 8.9, a light classical routine performs an optimisation

process over the parameters of the controls, forming a closed loop with the QPU. Hybrid classical-quantum methods involves decomposing the optimisation problem at hand into two or more sub-problems, each linked to specific decision-making processes of the original problem. While certain sub-problems can be tackled by a quantum algorithm such as a VQA, others are resolved using classical resources. The overall solution to the original problem is then a composite of the solutions to each sub-problem. Over the years, several advanced methods have been proposed, each with significant potential and notable limitations in terms of universality.

A widely applied classical method is the *Column Generation algorithm* [363], which aims to constrain the solution space to a few viable options. It iteratively adds new options during the process only if they can enhance the current best solution for the problem. However, identifying such solutions involves addressing several smaller, yet challenging, combinatorial sub-problems. Therefore, designing an efficient approach to solve these sub-problems, commonly known as pricing sub-problems, is crucial to ensuring high-quality solutions for a given problem.

Researchers in [364] demonstrated that addressing the pricing sub-problems with a quantum algorithm could not only enhance the quality of the final solutions but also reduce the number of iterations needed to find them. Utilising a neutral atom-based quantum computer as a sampler, one can design a quantum algorithm to identify not only one but several viable options to be added to the original problem. The authors presented a proof-of-concept for solving the well-known Vertex Colouring problem, which has immediate real-world applications such as resource allocation, scheduling, and planning-related problems [365].

Benders' Decomposition (BD) is another well known method in mathematical optimisation for solving large-scale problems, especially mixed-integer linear programs (MILPs) with continuous and integer variables. It breaks down the problem into a master problem and subproblems. The master problem deals with only a subset of decision variables, while the subproblems handle the rest, and usually becomes easier to solve. Solutions from subproblems generate additional constraints, called Benders' cuts, enhancing by addition the master problem's solution. While the subproblems are in general manageable on classical computers, the master problem, containing discrete variables, constitutes the computational bottleneck. Authors in [366] propose a framework combining classical BD with neutral atom-based computation where they automate the conversion of the master problem into a QUBO formulation.

8.5.2 - Hybrid GML approaches

Similar to the quantum optimisation approach, significant advancements in both performance and insights can be obtained through the hybridisation of QPUs with classical computers in GML. Utilising the unique expressiveness of quantum feature maps in conjunction with state-of-the-art graph machine learning techniques can lead to fruitful developments [317, 367, 368]. In [369, 370], the authors employ quantum correlation matrices as positional encoding for a transformer model. Empirical evidence is provided to show that the use of quantum features, particularly in the case of $k = 2$ particle quantum walks (k-QWs), exhibits superior expressiveness. When tested on arbitrarily chosen families of non-isomorphic strongly regular graphs, the k-QWs can distinguish between them, whereas classical walk kernels and k-Weisfeiler Lehman (k-WL) tests with $k < 3$ (at least) fail. It's worth noting that classical k-WL involves comparisons between k-tuples of nodes, rendering its complexity at least $\mathcal{O}(N^k)$, where N is the size of the graph. In contrast, for k-QWs, the algorithm's complexity is determined by

the number of shots needed to measure and reconstruct the distribution of the desired observable. The number of shots is $\mathcal{O}(\frac{1}{\epsilon^2})$ for a precision up to ϵ [301]. This potential advantage arises if the expressiveness of k-QWs is provably comparable to k-WL, as observed empirically for k equal to 1&2 but not proven in the general case.

This approach is not limited to transformer architectures alone. Quantum feature maps can be integrated into a generic encoder, which undergoes classical post-processing in any ML framework, as long as it ensures superior expressiveness for the specific task. In the realm of GML, this translates to the distinguishability of graphs by the WL test and its variants [323].

Another scheme called *quantum extremal learning* involves first to learn a surrogate function based on a given dataset, and second to optimise this function to find its extremal value. Previous methods [371, 372] relied on classically fitting QUBOs to datasets first, using for instance Factorisation Machines, and then optimising with a quantum QUBO solver. The authors of [373] uses a quantum ML model as a surrogate before applying the quantum optimisation procedure, greatly extending the class of models that can be studied and optimised with a notable example of applying this method on an industrially-relevant dataset of molecular structures.

9 - Conclusion and outlook

In this manuscript, we presented how Rydberg-based analog quantum computers can be supplemented by classical routines of numerical modelling and optimal control in order to enhance their overall performance in the context of quantum simulation and quantum algorithms. Furthermore, we emphasised that the optimal utilisation of existing quantum resources, still limited by noise and slow repetition rate, is achieved by integrating them within broader classical frameworks of optimisation and machine learning.

Throughout the course of this thesis, we wanted to convey the idea that performing successful experiments on current Rydberg atom platforms still requires a deep knowledge of the underlying physics. The quantum resources at our disposition being limited by hardware constraints and noisy processes, making the most out of them requires careful benchmarks and well designed control protocols. In chapter 1, we provided a step-by-step explanation of the inner workings of a quantum processor using Rydberg atoms, characterising the available possibilities, the kind of noise happening at each stage and the future features we can expect from the technological developments.

Empowering a quantum computer with classical tools

In chapter 2, we presented Pulser, an open source Python package dedicated to the emulation of the dynamics happening on Rydberg atom processors and enabling a direct communication with Pasqal QPUs using a cloud pipeline. We also presented several tensor network methods used throughout the thesis and gave insights on how to incorporate various physical noise such as decoherence or control fluctuations in emulation routines. We applied the latter to build a digital twin of a quantum processor and performed a first benchmark on a STIRAP protocol. Finally, we provided a summary table of the typical values of the encountered constraints and noises as we think it can prove useful for the community.

In chapter 3, after having presented several quantum optimal control methods and demonstrated their use to find time optimal and more robust quantum gates, we applied the GRAPE framework to the STIRAP protocol benchmarked in chapter 2 and showed that a faster and less power consuming protocol can slightly improve the state initialisation in the XY mode. Then, we introduced the concept of variational quantum algorithms with several parameterisations, including QAOA-like and an adiabatic version, and motivated the use in the following chapters of the Bayesian optimisation method as an efficient gradient-free classical optimiser.

Quantum many-body dynamics for quantum simulation

In chapter 4, we used the constructed variational quantum adiabatic algorithm to optimise a path towards the antiferromagnetic phase of the 1D Ising model, showing that the Bayesian optimiser helps locate regions with high structure factor while avoiding vanishing gaps, responsible for diabatic errors. Then, we report that implementing an adiabatic schedule optimised on a tensor network emulated version of the QPU does not yield the expected performance due to decoherence effects. We carefully benchmarked the noisy dynamics of small hexagonal systems to understand the hardware limitations on the correlation length and construct a variational closed loop to improve hardware results. Finally,

we performed a numerical study on the many body localisation phenomenon that one can probe using disordered antiferromagnetic chains of atoms on the processor and elaborated on the feasibility of implementing such an experiment.

In chapter 5, we switched to the dipolar XY model to first investigate how the preparation of its ground state can highlight a continuous symmetry breaking. After a detailed benchmark of the complete protocol on a small plaquette to identify the various noise contributions, we adiabatically prepare long range order ferromagnetic states on a 10×10 square lattice. Then using quenches, we experimentally probe the phase diagram of both ferromagnetic and antiferromagnetic version of the dipolar XY model and showed how thermal and quantum fluctuations can destroy order. In a second part, after introducing a new way of performing multi-basis measurement on Rydberg-Rydberg encoded qubits, we experimentally prepare and measure chiral states showcasing with a digital twin that the hardware errors are well understand. Using local rotations, we performed full quantum state tomography and studied topological properties of frustrated geometries. Using either ground-Rydberg or Rydberg-Rydberg encodings, Rydberg platforms can thus probe a number of interesting phenomena in quantum magnetism utilising adiabatic protocols and local addressing schemes.

Quantum algorithms : exploring innovative applications with graphs

We have also demonstrated in the last part of this thesis, in addition to the traditional use for simulating intractable quantum dynamics, that Rydberg atom QPUs also possess the capability to address combinatorial optimisation problems and machine learning tasks in innovative hybrid algorithms.

In chapter 6, we presented how to map combinatorial problems such as the Maximum Independent Set to spin systems and to embed the related graph instances to atomic arrays using optimised positioning and batching techniques. We then solved specific unit disk instances using variational quantum algorithms, aided by classical Bayesian optimiser as a clever navigator of the landscape to optimise. We showed that the Rydberg version of Quantum Approximate Optimisation Algorithm suffers from technical limitations while the Variational Quantum Adiabatic Algorithm can not only provide more robust optimised drives but also protocols working on a family of similarly structured graphs. We selected an optimised adiabatic protocol for its generalisability on triangular graphs and experimentally applied it to a dataset of graphs representing smart charging tasks provided by EDF, reporting that a solution was found each time after only few repetitions.

In chapter 7, we illustrated how quantum dynamics across an ensemble of Rydberg atoms can serve as a quantum feature map for graphs. We extended the latter to a Quantum Evolution Kernel, and we highlighted that the geometric structures produced by this new kernel are beyond the reach of classical learning methods. Although no clear improvement of the classification score was obtained on the experimentally implemented PTC-FM dataset, we numerically estimated that a gain in performance could be obtained on a synthetic dataset created from underlying lattices that yield distinct frustrations during quantum dynamics.

Finally, in chapter 8, we introduced a quantum-enhanced classification algorithm tasked with predicting credit rating downgrades in the finance sector. We combined the already existing QBoost framework with a novel QUBO solving method utilising Random Graph Sampling on Rydberg atom processors and obtained results similar in performance with classical methods such as simulated annealing and random forest. However, the quantum-enhanced machine learning model, once trained,

was using significantly fewer weak learners, improving on the interpretability side. We ended the chapter by reviewing recent techniques to combine in an optimal way quantum blocks within classical optimisation and machine learning frameworks in an hybrid fashion.

Outlook

Reaching practicality in industrial applications for the kind of analog algorithms presented above will first require an increase of the size of the instances being tackled. Fortunately, Rydberg platforms are constantly increasing the number of atoms they can trap, with a groundbreaking record of 6100 highly coherent atomic qubits [374], approaching the threshold values mentioned by [267] for practical advantage of UD-MIS solving for instance. However, scaling the instance size must not constitute the only target as integrating the hypothesis of finite fidelity of operations inside adapted tensor networks can accelerate the classical emulation of large quantum systems [375] and close once again a leap quantum computers could take.

Building on the results of chapter 6, it would be interesting to produce generalisable adiabatic protocols on more complex lattices, extending the families of unit disk graphs tackled, or on other kind of structural similarities, notably for QUBO problems generated by the same methods as viewed in chapter 8 . Moreover, one can also consider the potential of optimising more complex protocols using counter-diabatic local terms to mitigate diabatic errors, as explored in [376], or leveraging Bayesian optimisation to craft efficient reverse annealing protocols, as demonstrated in [377]. Furthermore, training neural networks with optimised controls executing specific gates can lead to the discovery of a suite of protocols realising a parameterised family of gates, such as two-qubit controlled rotations, as introduced in [378].

On another note, the results of chapter 7 conveys the idea that while the geometry induced by quantum kernel might not be useful to tackle classical data, it could prove fruitful handling quantum datasets. However, to be able to not only compare measured observables but the quantum states themselves, one would require more advances methods of manipulation and transfer through quantum networks [379] and storage with quantum memories [380].

Apart from the optimised gates exemplified in chapter 3, the entirety of the research showcased in this thesis relies heavily on the analog computation paradigm. Nonetheless, the digital paradigm is currently catching up quickly in the Rydberg community, characterised by significant enthusiasm for achieving higher entanglement fidelities and more resilient error-correction codes [53]. Given the familiarity of a substantial portion of the quantum computing community with this paradigm, particularly prevalent in superconducting and ion qubits, this momentum could propel Rydberg atom technology to the forefront of efforts toward building a fault-tolerant quantum computer [381].

A - Classical kernels

A.1 - SVM-ϑ kernel

The Lovasz-ϑ kernel is a graph kernel that is based on the Lovasz extension of graph isomorphisms. It measures the similarity between graphs by comparing their induced subgraphs. However, computing the Lovasz-ϑ kernel can be computationally intensive, especially for large graphs, which may limit its practicality in some scenarios. The SVM-ϑ kernel was proposed as an alternative to the more computationally intensive Lovasz-ϑ kernel. Both ϑ kernels leverage the so-called orthogonal representation of a graph. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the orthogonal representation is an assignment of unit vectors $\{\mathbf{u}_i\}$ to each node of the graph, subject to the constraint that unit vectors associated to vertices that are not joined by an edge are orthogonal: $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ if $\{i, j\} \notin \mathcal{E}$.

Orthogonal representations are not unique, but there is a particular representation associated with the ϑ number [382] of a graph. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n vertices, denote $U_{\mathcal{G}}$ an orthogonal representation of $\mathcal{G}$, and C the space of unit vectors in $\mathbb{R}^n$. The ϑ number is defined as:

$$\vartheta(\mathcal{G}) := \min_{\mathbf{c} \in C} \min_{U_{\mathcal{G}}} \max_{\mathbf{u}_i \in U_{\mathcal{G}}} \frac{1}{\langle \mathbf{c}, \mathbf{u}_i \rangle^2}. \tag{A.1}$$

From now on, we will always be referring to the particular orthogonal representation $U_{\mathcal{G}}$ that minimizes (A.1).

Now consider a subset of vertices $B \subset \mathcal{V}$, and call $U_{\mathcal{G}|B}$ the orthogonal representation obtained from $U_{\mathcal{G}}$ by removing the vectors that are not in B:

$$U_{\mathcal{G}|B} := \{\mathbf{u}_i \in U_{\mathcal{G}} : \ i \in B\}. \tag{A.2}$$

Note that $U_{\mathcal{G}|B}$ preserves the global properties encoded in $U_{\mathcal{G}}$ through the orthogonal constraint, and that $U_{\mathcal{G}|B}$ is not in general the orthogonal representation of the subgraph of $\mathcal{G}$ containing only the vertices in B. Define the ϑ_B number:

$$\vartheta_B(\mathcal{G}) := \min_{\mathbf{c} \in C} \max_{\mathbf{u}_i \in U_{\mathcal{G}|B}} \frac{1}{\langle \mathbf{c}, \mathbf{u}_i \rangle^2}. \tag{A.3}$$

We are ready now to give the definition of the Lovasz-ϑ kernel. Given two graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$, $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, define:

$$K_{\text{Lo}}(\mathcal{G}_1, \mathcal{G}_2) := \sum_{B_1 \subset V_1} \sum_{B_2 \subset V_2} \delta_{|B_1|,|B_2|} \ \frac{1}{Z} \ k\left(\vartheta_{B_1}, \vartheta_{B_2}\right), \tag{A.4}$$

where $Z = \binom{|V_1|}{|B_1|}\binom{|V_2|}{|B_2|}$, δ is the Kronecker delta, and k is a freely specifiable kernel (called base kernel) from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$.

The SVM-ϑ kernel is defined as (A.4), but it uses an approximation for the ϑ numbers. Consider a graph $\mathcal{G}$ with n vertices and adjacency matrix A, and let $\rho \geq -\lambda$, where λ is the minimum eigenvalue of A. The matrix

$$\kappa := \frac{1}{\rho} A + I, \tag{A.5}$$

is positive semi-definite. Define the maximization problem:

$$\max_{\alpha_i \geq 0} \; 2\sum_{i=1}^{n} \alpha_i - \sum_{i,j=1}^{n} \alpha_i \alpha_j \kappa_{ij}. \tag{A.6}$$

If $\{\alpha_i^*\}$ are the maximizers of (A.6), then it can be proven that on certain families of graphs the quantity $\sum_i \alpha_i^*$ is with high probability a constant factor approximation to $\vartheta(\mathcal{G})$:

$$\vartheta(\mathcal{G}) \leq \sum_{i=1}^{n} \alpha_i^* \leq \gamma \vartheta(\mathcal{G}), \tag{A.7}$$

for some γ. The SVM-ϑ kernel then replaces the ϑ_B numbers on subgraphs with:

$$\vartheta_B(\mathcal{G}) \to \sum_{j \in B} \alpha_j^*. \tag{A.8}$$

The SVM-ϑ kernel requires a choice of base kernel $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. We choose a translation invariant universal kernel [383] $k(x, y) = (\beta + ||x - y||^2)^{-\alpha}$, where α and β are two trainable hyperparameters.

A.2 - Graphlet Sampling kernel

This kernel is designed to capture the local structural patterns, known as graphlets, within a graph. Graphlets are small connected subgraphs that can be used as building blocks to represent the structural properties of a larger graph. The Graphlet Sampling kernel measures the similarity between two graphs by counting the occurrences of different graphlets in both graphs and comparing their frequencies.

Formally, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $H = (\mathcal{V}_H, \mathcal{E}_H)$ be two graphs. We say that $\mathcal{H}$ is a subgraph of $\mathcal{G}$ if there exists an injective map $\alpha : \mathcal{V}_\mathcal{H} \to \mathcal{V}$ such that $(u, v) \in \mathcal{E}_H \iff (\alpha(u), \alpha(v)) \in \mathcal{E}$. In general it might be possible to map $\mathcal{H}$ into $\mathcal{G}$ in several different ways, *i.e.* the mapping α, if it exists, is not necessarily unique.

Given two graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, the idea behind the Graphlet kernel is to pick an integer $k < \min\{|\mathcal{V}_1|, |\mathcal{V}_2|\}$, enumerate all possible graphs of size k and find the number of ways they can be mapped to $\mathcal{G}_1$ and $\mathcal{G}_2$. Denote by $f_{\mathcal{G}_i}^{(k)}$ the vector where each entry counts the way a specific graph of size k can be mapped as a subgraph of $\mathcal{G}_i$. A kernel can then be defined as the dot product $f_{\mathcal{G}_1}^{(k)} \cdot f_{\mathcal{G}_2}^{(k)}$ between the two vectors.

The complexity of computing such a kernel scales as $O(n^k)$, as there are $\binom{n}{k}$ size-k subgraphs in a graph of size n. For this reason it is preferable to resort to sampling rather than complete enumeration [384]. Given a choice of integer N, graphs $g_1, \ldots, g_N$ of size between 3 and k are randomly sampled. The number of ways each g_i can be mapped as a subgraph of $\mathcal{G}_j$ is computed and stored in a vector $f_{\mathcal{G}_j}$, and the Graphlet Sampling kernel is defined as the dot product:

$$K_{\text{GS}}(\mathcal{G}_1, \mathcal{G}_2) := f_{\mathcal{G}_1} \cdot f_{\mathcal{G}_2}. \tag{A.9}$$

To account for the different size of $\mathcal{G}_1$ and $\mathcal{G}_2$, each vector can be normalized by the total number of its subgraphs.

A.3 - Random Walk kernel

The Random Walk kernel is one of the oldest and most studied graph kernels [385]. Given two graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, the idea is to measure the probability of simultaneous random walks of a certain length between two vertices in $\mathcal{G}_1$ and $\mathcal{G}_2$.

Simultaneous random walks can be conveniently encoded in powers of the adjacency matrix on the product graph. The product graph $\mathcal{G}_1 \times \mathcal{G}_2 = \mathcal{G}_\times = (\mathcal{V}_\times, \mathcal{E}_\times)$ is defined as follows:

$$\mathcal{V}_\times := \{(u_i, u_r) \mid u_i \in \mathcal{V}_1,\ u_r \in \mathcal{V}_2\}, \tag{A.10}$$

$$\mathcal{E}_\times := \{((u_i, u_r), (v_j, v_s)) \mid (u_i, v_j) \in \mathcal{E}_1, \\ (u_r, v_s) \in \mathcal{E}_2\}. \tag{A.11}$$

In other words, an edge in the product graph indicates that an edge exists between the endpoints in both $\mathcal{G}_1$ and $\mathcal{G}_2$. If $A_\times$ is the adjacency matrix of the product graph, then the entries of $A_\times^k$ indicate the probability of a simultaneous random walk of length k between two vertices $u_i, v_j \in \mathcal{V}_1$ and $u_r, v_s \in \mathcal{V}_2$.

If $p, q \in \mathbb{R}^{|\mathcal{V}_\times|}$ are vectors representing the probability distribution of respectively starting or stopping the walk at a certain node of $\mathcal{V}_\times$, the first idea for a kernel would be to compute the sum $\sum_k q^{\mathrm{T}} A_\times^k p$, which however may fail to converge. A simple modification to make the sum convergent is to choose an appropriate length-dependent weight $\mu(k)$:

$$K(\mathcal{G}_1, \mathcal{G}_2) := \sum_{k=0}^{\infty} \mu(k)\ q^{\mathrm{T}} A_\times^k p. \tag{A.12}$$

The Geometric Random Walk kernel is obtained by choosing the weights to be the coefficients of a geometric series $\mu(k) = \lambda^k$, and p, q to be uniform. If λ is tuned in such a way as to make the series convergent, the kernel reads:

$$K_{\mathrm{RW}}(\mathcal{G}_1, \mathcal{G}_2) := \sum_{k=0}^{\infty} \lambda^k\ e^{\mathrm{T}} A_\times^k e = e^{\mathrm{T}} \left(I - \lambda A_\times\right)^{-1} e, \tag{A.13}$$

where e denote vectors with all the entries equal to 1.

The cost of matrix inversion scales as the cube of the matrix size. If $|\mathcal{V}_1| = |\mathcal{V}_2| = n$, then the cost of the algorithm scales as $O(n^6)$, as it involves the inversion of an adjacency matrix of size $n^2 \times n^2$. Several methods are proposed in [386] to make the computation faster. The Spectral Decomposition method in particular allows to reduce the complexity for unlabeled graphs to $O(n^3)$. Essentially, one exploits the fact that the adjacency matrix of the product graph can be decomposed in the tensor product of the individual adjacency matrices:

$$A_\times = A_1 \otimes A_2, \tag{A.14}$$

which allows to diagonalize each $n \times n$ adjacency matrix in $O(n^3)$ time and perform the inversion only on the diagonal components.

A.4 - Shortest Path kernel

Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, an edge path between two vertices $u, v \in \mathcal{V}$ is a sequence of edges $(e_1, \ldots, e_n)$ such that $u \in e_1$, $v \in e_n$, e_i and e_{i+1} are contiguous (*i.e.* they have one of the endpoints in common) and $e_i \neq e_j$ for $i \neq j$. Computing the shortest edge path between any two nodes of a graph can be done in polynomial time with the Dijkstra [387] or Floyd-Warshall [388] algorithms, which makes it a viable feature to be probed by a graph kernel.

The first step of the Shortest Path kernel is to transform the graphs into shortest path graphs. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the shortest path graph $\mathcal{G}^S = (\mathcal{V}^S, \mathcal{E}^S)$ associated to $\mathcal{G}$ is defined as:

$$\mathcal{V}^S = \mathcal{V}, \tag{A.15}$$

$$\mathcal{E}^S = \{(u, v) \mid \exists \text{ an edge path } (e_1, \ldots, e_n) \text{ between } u \text{ and } v \text{ in } \mathcal{G}\}. \tag{A.16}$$

In addition, to each edge $e \in \mathcal{E}^S$ a label $l(e)$ is assigned given by the length of the shortest path in $\mathcal{G}$ between its endpoints. The Shortest Path kernel is then defined as:

$$K_{\text{SP}}(\mathcal{G}_1, \mathcal{G}_2) := \sum_{e \in \mathcal{E}_1^S} \sum_{p \in \mathcal{E}_2^S} k(e, p), \tag{A.17}$$

with k being a kernel between edge paths such as the Brownian bridge kernel:

$$k(e, p) := \max\{0,\ c - |l(e) - l(p)|\}, \tag{A.18}$$

for a choice of c.

In the following section, we will detail our quantum evolution kernel. A main interest is to understand if it can capture graph features that are not accessible to classical graph kernels.

Bibliography

[1] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of bell's inequalities using time-varying analyzers," *Phys. Rev. Lett.*, vol. 49, pp. 1804–1807, 25 1982-12. DOI: 10.1103/PhysRevLett.49.1804 (cit. on p. 1).

[2] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018-8, ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. [Online]. Available: https://doi.org/10.22331/q-2018-08-06-79 (cit. on p. 1).

[3] L. Beguin, "Measurement of the van der Waals interaction between two Rydberg atoms," Theses, Institut d'Optique Graduate School, 2013-12. [Online]. Available: https://pastel.hal.science/pastel-00936072 (cit. on p. 5).

[4] S. d. Leseleuc de kerouara, "Quantum simulation of spin models with assembled arrays of Rydberg atoms," Theses, Université Paris Saclay (COmUE), 2018-12. [Online]. Available: https://pastel.archives-ouvertes.fr/tel-02088297 (cit. on pp. 5, 28).

[5] P. Scholl, "Simulation quantique de modèles de spins avec des grandes matrices d'atomes de Rydberg," Theses, Université Paris-Saclay, 2021-12. [Online]. Available: https://pastel.archives-ouvertes.fr/tel-03523082 (cit. on pp. 5, 79).

[6] S. Aaronson and S.-H. Hung, *Certified randomness from quantum supremacy*, 2023. DOI: 10.1145/3564246.3585145. arXiv: 2303.01625 [quant-ph] (cit. on p. 6).

[7] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019-10, ISSN: 1476-4687. DOI: 10.1038/s41586-019-1666-5 (cit. on p. 6).

[8] J. Schoser *et al.*, "Intense source of cold rb atoms from a pure two-dimensional magneto-optical trap," *Phys. Rev. A*, vol. 66, p. 023 410, 2 2002-8. DOI: 10.1103/PhysRevA.66.023410 (cit. on p. 6).

[9] N. Schlosser, G. Reymond, and P. Grangier, "Collisional blockade in microscopic optical dipole traps," *Phys. Rev. Lett.*, vol. 89, p. 023 005, 2 2002-6. DOI: 10.1103/PhysRevLett.89.023005 (cit. on p. 6).

[10] M. O. Brown, T. Thiele, C. Kiehl, T.-W. Hsu, and C. A. Regal, "Gray-molasses optical-tweezer loading: Controlling collisions for scaling atom-array assembly," *Phys. Rev. X*, vol. 9, p. 011 057, 1 2019-3. DOI: 10.1103/PhysRevX.9.011057 (cit. on pp. 7, 26).

[11] A. M. Kaufman, B. J. Lester, and C. A. Regal, "Cooling a single atom in an optical tweezer to its quantum ground state," *Phys. Rev. X*, vol. 2, p. 041 014, 4 2012-11. DOI: 10.1103/PhysRevX.2.041014 (cit. on p. 7).

[12] C. Tuchendler, A. M. Lance, A. Browaeys, Y. R. P. Sortais, and P. Grangier, "Energy distribution and cooling of a single atom in an optical tweezer," *Phys. Rev. A*, vol. 78, no. 3, 2008-9, ISSN: 1094-1622. DOI: 10.1103/physreva.78.033425 (cit. on p. 7).

[13] K.-N. Schymik *et al.*, "Single atoms with 6000-second trapping lifetimes in optical-tweezer arrays at cryogenic temperatures," *Phys. Rev. Appl.*, vol. 16, p. 034 013, 3 2021-9. DOI: 10.1103/PhysRevApplied.16.034013 (cit. on p. 7).

[14] R. Grimm, M. Weidemüller, and Y. B. Ovchinnikov, "Optical dipole traps for neutral atoms," in *Advances in atomic, molecular, and optical physics*, vol. 42, Elsevier, 2000, pp. 95–170. DOI: 10.1016/S1049-250X(08)60186-X. arXiv: physics/9902072 [physics.atom-ph] (cit. on p. 8).

[15] D. Schäffner *et al.*, "Arrays of individually controllable optical tweezers based on 3d-printed microlens arrays," *Optics Express*, vol. 28, no. 6, p. 8640, 2020-3, ISSN: 1094-4087. DOI: 10.1364/oe.386243 (cit. on p. 7).

[16] M. Endres *et al.*, "Atom-by-atom assembly of defect-free one-dimensional cold atom arrays," *Science*, vol. 354, no. 6315, pp. 1024–1027, 2016. DOI: 10.1126/science.aah3752. eprint: https://www.science.org/doi/pdf/10.1126/science.aah3752. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.aah3752 (cit. on p. 7).

[17] S. Bergamini, B. Darquié, M. Jones, L. Jacubowiez, A. Browaeys, and P. Grangier, "Holographic generation of microtrap arrays for single atoms by use of a programmable phase modulator," *Journal of the Optical Society of America B*, vol. 21, no. 11, p. 1889, 2004-11, ISSN: 1520-8540. DOI: 10.1364/josab.21.001889 (cit. on p. 7).

[18] D. Barredo, V. Lienhard, S. de Léséleuc, T. Lahaye, and A. Browaeys, "Synthetic three-dimensional atomic structures assembled atom by atom," en, *Nature*, vol. 561, no. 7721, pp. 79–82, 2018-9, ISSN: 1476-4687. DOI: 10.1038/s41586-018-0450-2. arXiv: 1712.02727 [quant-ph] (cit. on pp. 8, 28).

[19] T. M. Graham, E. Oh, and M. Saffman, "Multiscale architecture for fast optical addressing and control of large-scale qubit arrays," *Applied Optics*, vol. 62, no. 12, p. 3242, 2023-4, ISSN: 2155-3165. DOI: 10.1364/ao.484367 (cit. on p. 8).

[20] D. F. Crouse, "On implementing 2d rectangular assignment algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 4, pp. 1679–1696, 2016. DOI: 10.1109/TAES.2016.140952 (cit. on p. 10).

[21] H. Kim, M. Kim, W. Lee, and J. Ahn, "Gerchberg-saxton algorithm for fast and efficient atom rearrangement in optical tweezer traps," *Opt. Express*, vol. 27, no. 3, pp. 2184–2196, 2019-2. DOI: 10.1364/OE.27.002184. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-27-3-2184 (cit. on p. 10).

[22] K.-N. Schymik *et al.*, "Enhanced atom-by-atom assembly of arbitrary tweezer arrays," *Phys. Rev. A*, vol. 102, p. 063 107, 6 2020-12. DOI: 10.1103/PhysRevA.102.063107 (cit. on p. 10).

[23] R. El Sabeh *et al.*, "Efficient algorithms to solve atom reconfiguration problems. ii. assignment-rerouting-ordering algorithm," *Phys. Rev. A*, vol. 108, no. 2, 2023-8, ISSN: 2469-9934. DOI: 10.1103/physreva.108.023108 (cit. on p. 10).

[24] B. Cimring *et al.*, "Efficient algorithms to solve atom reconfiguration problems. i. redistribution-reconfiguration algorithm," *Phys. Rev. A*, vol. 108, no. 2, 2023-8, ISSN: 2469-9934. DOI: 10.1103/physreva.108.023107 (cit. on pp. 10, 26).

[25] K.-N. Schymik *et al.*, "In situ equalization of single-atom loading in large-scale optical tweezer arrays," *Phys. Rev. A*, vol. 106, no. 2, p. 022 611, 2022. DOI: 10.1103/PhysRevA.106.022611 (cit. on pp. 10, 28).

[26] M. A. Norcia *et al.*, *Iterative assembly of* ^{171}*yb atom arrays in cavity-enhanced optical lattices*, 2024. arXiv: 2401.16177 [quant-ph] (cit. on pp. 10, 26).

[27] Y. Song, M. Kim, H. Hwang, W. Lee, and J. Ahn, "Quantum simulation of cayley-tree ising hamiltonians with three-dimensional rydberg atoms," *Phys. Rev. Res.*, vol. 3, p. 013 286, 1 2021-3. DOI: 10.1103/PhysRevResearch.3.013286 (cit. on p. 10).

[28] S. Ebadi *et al.*, "Quantum phases of matter on a 256-atom programmable quantum simulator," *Nature*, vol. 595, no. 7866, pp. 227–232, 2021-7, ISSN: 1476-4687. DOI: 10.1038/s41586-021-03582-4 (cit. on p. 10).

[29] D. Bluvstein *et al.*, "A quantum processor based on coherent transport of entangled atom arrays," *Nature*, vol. 604, no. 7906, pp. 451–456, 2022-4, ISSN: 1476-4687. DOI: 10.1038/s41586-022-04592-6 (cit. on pp. 10, 26, 133).

[30] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, "Efficient Z gates for quantum computing," *Phys. Rev. A*, vol. 96, p. 022 330, 2 2017-8. DOI: 10.1103/PhysRevA.96.022330 (cit. on p. 11).

[31] E. D. Black, "An introduction to Pound–Drever–Hall laser frequency stabilization," en, *Am. J. Phys.*, vol. 69, no. 1, pp. 79–87, 2001-1. DOI: 10.1119/1.1286663 (cit. on p. 13).

[32] P. Cladé, "Oscillations de bloch d'atomes ultrafroids et mesure de la constante de structure fine," Thèse de doctorat dirigée par Himbert, Marc Physique Paris, CNAM 2005, Ph.D. dissertation, CNAM, 2005, 1 vol. (221 p.) [Online]. Available: http://www.theses.fr/2005CNAM0492 (cit. on p. 13).

[33] R. Löw *et al.*, "An experimental and theoretical guide to strongly interacting rydberg gases," *Journal of Physics B: Atomic, Molecular and Optical Physics*, vol. 45, no. 11, p. 113 001, 2012-5, ISSN: 1361-6455. DOI: 10.1088/0953-4075/45/11/113001 (cit. on p. 17).

[34] H. Wu, R. Richaud, J.-M. Raimond, M. Brune, and S. Gleyzes, "Millisecond-lived circular rydberg atoms in a room-temperature experiment," *Phys. Rev. Lett.*, vol. 130, no. 2, 2023-1, ISSN: 1079-7114. DOI: 10.1103/physrevlett.130.023202 (cit. on p. 18).

[35] S. Weber *et al.*, "Tutorial: Calculation of Rydberg interaction potentials," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 50, no. 13, p. 133 001, 2017. DOI: 10.1088/1361-6455/aa743a (cit. on p. 20).

[36] S. de Léséleuc *et al.*, "Observation of a symmetry-protected topological phase of interacting bosons with rydberg atoms," *Science*, vol. 365, no. 6455, pp. 775–780, 2019-8, ISSN: 1095-9203. DOI: 10.1126/science.aav9105. arXiv: 1810.13286 [cond-mat, physics:physics, physics:quant-ph] (cit. on p. 18).

[37] C. Chen *et al.*, "Continuous symmetry breaking in a two-dimensional rydberg array," *Nature*, vol. 616, no. 7958, pp. 691–695, 2023-4, ISSN: 1476-4687. DOI: 10.1038/s41586-023-05859-2 (cit. on pp. 18, 29, 98, 101, 108, 110, 114, 116).

[38] G. Günter *et al.*, "Observing the dynamics of dipole-mediated energy transport by interaction-enhanced imaging," *Science (New York, N.Y.)*, vol. 342, 2013-11. DOI: 10.1126/science.1244843 (cit. on p. 18).

[39] E. Collini, "Spectroscopic signatures of quantum-coherent energy transfer," *Chemical Society Reviews*, vol. 42, no. 12, p. 4932, 2013, ISSN: 1460-4744. DOI: 10.1039/c3cs35444j (cit. on p. 18).

[40] P. M. Preiss *et al.*, "Strongly correlated quantum walks in optical lattices," *Science*, vol. 347, no. 6227, pp. 1229–1233, 2015-3, ISSN: 1095-9203. DOI: 10.1126/science.1260364 (cit. on p. 18).

[41] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, "Exponential algorithmic speedup by a quantum walk," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, ser. STOC03, San Diego, CA, USA: ACM, 2003-6, pp. 59–68, ISBN: 1581136749. DOI: 10.1145/780542.780552 (cit. on p. 18).

[42] L.-P. Henry, "Classical and quantum two-dimensional ice : Coulomb and ordered phases," Theses, Ecole normale supérieure de lyon - ENS LYON, 2013-11. [Online]. Available: https://theses.hal.science/tel-00932367 (cit. on p. 21).

[43] A. Mitra, "Quantum quench dynamics," *Annual Review of Condensed Matter Physics*, vol. 9, no. 1, pp. 245–259, 2018-3, ISSN: 1947-5462. DOI: 10.1146/annurev-conmatphys-031016-025451 (cit. on p. 22).

[44] P. W. Anderson, "Absence of diffusion in certain random lattices," *Phys. Rev.*, vol. 109, pp. 1492–1505, 5 1958-3. DOI: 10.1103/PhysRev.109.1492 (cit. on p. 22).

[45] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte der Physik*, vol. 48, no. 9–11, pp. 771–783, 2000-9, ISSN: 1521-3978. DOI: 10.1002/1521-3978(200009)48:9/11<771::aid-prop771>3.0.co;2-e (cit. on p. 22).

[46] F. Gyger *et al.*, *Continuous operation of large-scale atom arrays in optical lattices*, 2024. arXiv: 2402.04994 [quant-ph] (cit. on p. 26).

[47] A. Bergschneider *et al.*, "Spin-resolved single-atom imaging of 6li in free space," *Phys. Rev. A*, vol. 97, no. 6, 2018-6, ISSN: 2469-9934. DOI: 10.1103/physreva.97.063613 (cit. on p. 26).

[48] A. Buzulutskov, "Advances in cryogenic avalanche detectors," *Journal of Instrumentation*, vol. 7, no. 02, pp. C02025–C02025, 2012-2, ISSN: 1748-0221. DOI: 10.1088/1748-0221/7/02/c02025 (cit. on p. 26).

[49] E. Deist *et al.*, "Mid-circuit cavity measurement in a neutral atom array," *Phys. Rev. Lett.*, vol. 129, no. 20, 2022-11, ISSN: 1079-7114. DOI: 10.1103/physrevlett.129.203602 (cit. on p. 26).

[50] H. Levine *et al.*, "Dispersive optical systems for scalable raman driving of hyperfine qubits," *Phys. Rev. A*, vol. 105, no. 3, 2022-3, ISSN: 2469-9934. DOI: 10.1103/physreva.105.032618 (cit. on p. 27).

[51] S. J. Evered *et al.*, "High-fidelity parallel entangling gates on a neutral atom quantum computer," *arXiv e-prints*, arXiv:2304.05420, arXiv:2304.05420, 2023-4. DOI: 10.48550/arXiv.2304.05420. arXiv: 2304.05420 [quant-ph] (cit. on pp. 27, 53).

[52] S. Jandura and G. Pupillo, "Time-optimal two- and three-qubit gates for rydberg atoms," *Quantum*, vol. 6, p. 712, 2022-5, ISSN: 2521-327X. DOI: 10.22331/q-2022-05-13-712 (cit. on pp. 27, 53, 54, 56).

[53] D. Bluvstein *et al.*, "Logical quantum processor based on reconfigurable atom arrays," *Nature*, vol. 626, no. 7997, pp. 58–65, 2023-12, ISSN: 1476-4687. DOI: 10.1038/s41586-023-06927-3 (cit. on pp. 27, 28, 203).

[54] I. S. Madjarov *et al.*, "High-fidelity entanglement and detection of alkaline-earth rydberg atoms," *Nat. Phys.*, vol. 16, no. 8, pp. 857–861, 2020. DOI: 10.1038/s41567-020-0903-z (cit. on p. 28).

[55] K. Singh, S. Anand, A. Pocklington, J. T. Kemp, and H. Bernien, "Dual-element, two-dimensional atom array with continuous-mode operation," *Phys. Rev. X*, vol. 12, no. 1, p. 011 040, 2022. DOI: 10.1103/PhysRevX.12.011040 (cit. on p. 28).

[56] T. Nguyen *et al.*, "Towards quantum simulation with circular rydberg atoms," *Phys. Rev. X*, vol. 8, no. 1, 2018-2, ISSN: 2160-3308. DOI: 10.1103/physrevx.8.011032 (cit. on p. 28).

[57] B. Ravon *et al.*, "Array of individual circular rydberg atoms trapped in optical tweezers," *arXiv preprint arXiv:2304.04831*, 2023. DOI: 10.1103/PhysRevLett.131.093401 (cit. on p. 28).

[58] G. Bornet *et al.*, *Enhancing a many-body dipolar rydberg tweezer array with arbitrary local controls*, 2024. arXiv: 2402.11056 [quant-ph] (cit. on pp. 28, 29, 98, 123).

[59] T. M. Graham *et al.*, *Mid-circuit measurements on a single species neutral alkali atom quantum processor*, 2023. arXiv: 2303.10051 [quant-ph] (cit. on p. 28).

[60] P. Scholl *et al.*, "Microwave Engineering of Programmable X X Z Hamiltonians in Arrays of Rydberg Atoms," *PRX Quantum*, vol. 3, no. 2, 020303, p. 020 303, 2022-4. DOI: 10.1103/PRXQuantum.3.020303. arXiv: 2107.14459 [quant-ph] (cit. on p. 29).

[61] K. Wright *et al.*, "Benchmarking an 11-qubit quantum computer," *Nat. Commun.*, vol. 10, no. 1, p. 5464, 2019-11, ISSN: 2041-1723. DOI: 10.1038/s41467-019-13534-2 (cit. on p. 29).

[62] A. J. Menssen *et al.*, *Scalable photonic integrated circuits for programmable control of atomic systems*, 2022-10. DOI: 10.48550/arXiv.2210.03100. arXiv: 2210.03100 [physics, physics:quant-ph] (cit. on p. 29).

[63] I. Christen *et al.*, *An integrated photonic engine for programmable atomic control*, 2022-8. DOI: 10.48550/arXiv.2208.06732. arXiv: 2208.06732 [physics, physics:quant-ph] (cit. on p. 29).

[64] H. Silvério *et al.*, "Pulser: An open-source package for the design of pulse sequences in programmable neutral-atom arrays," *Quantum*, vol. 6, p. 629, 2022. DOI: 10.22331/q-2022-01-24-629 (cit. on p. 34).

[65] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Annals of physics*, vol. 349, pp. 117–158, 2014-10, ISSN: 0003-4916. DOI: 10.1016/j.aop.2014.06.013. arXiv: 1306.2164 [cond-mat.str-el] (cit. on p. 39).

[66] F. Pan and P. Zhang, "Simulating the Sycamore quantum supremacy circuits," *arXiv e-prints*, arXiv:2103.03074, arXiv:2103.03074, 2021-3. arXiv: 2103.03074 [quant-ph] (cit. on p. 39).

[67] A. Klümper, A. Schadschneider, and J. Zittartz, "Matrix product ground states for one-dimensional spin-1 quantum antiferromagnets," *Europhysics Letters (EPL)*, vol. 24, no. 4, pp. 293–297, 1993-11, ISSN: 1286-4854. DOI: 10.1209/0295-5075/24/4/010 (cit. on p. 40).

[68] F. Verstraete and J. I. Cirac, *Renormalization algorithms for quantum-many body systems in two and higher dimensions*, 2004. arXiv: cond-mat/0407066 [cond-mat.str-el] (cit. on p. 40).

[69] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.*, vol. 69, no. 19, p. 2863, 1992. DOI: 10.1103/PhysRevLett.69.2863 (cit. on p. 40).

[70] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of physics*, vol. 326, no. 1, pp. 96–192, 2011. DOI: 10.1016/j.aop.2010.09.012 (cit. on p. 40).

[71] I. Cirac, D. Perez-Garcia, N. Schuch, and F. Verstraete, "Matrix product states and projected entangled pair states: Concepts, symmetries, and theorems," *arXiv preprint arXiv:2011.12127*, 2020. DOI: 10.1103/RevModPhys.93.045003 (cit. on p. 40).

[72] K. Hyatt and E. M. Stoudenmire, "Dmrg approach to optimizing two-dimensional tensor networks," 2020. arXiv: 1908.08833 [cond-mat.str-el] (cit. on p. 40).

[73] E. M. Stoudenmire and S. R. White, "Studying two-dimensional systems with the density matrix renormalization group," *Annu. Rev. Condens. Matter Phys.*, vol. 3, no. 1, pp. 111–128, 2012. DOI: 10.1146/annurev-conmatphys-020911-125018 (cit. on p. 40).

[74] P. Scholl *et al.*, "Quantum simulation of 2d antiferromagnets with hundreds of rydberg atoms," *Nature*, vol. 595, no. 7866, pp. 233–238, 2021-7, ISSN: 1476-4687. DOI: 10.1038/s41586-021-03585-1 (cit. on pp. 40, 76, 98).

[75] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, "Time-dependent variational principle for quantum lattices," *Phys. Rev. Lett.*, vol. 107, no. 7, 2011-8, ISSN: 1079-7114. DOI: 10.1103/physrevlett.107.070601 (cit. on p. 41).

[76] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, "Unifying time evolution and optimization with matrix product states," *Phys. Rev. B*, vol. 94, no. 16, 2016-10, ISSN: 2469-9969. DOI: 10.1103/physrevb.94.165116 (cit. on p. 41).

[77] M. Fishman, S. R. White, and E. M. Stoudenmire, "The itensor software library for tensor network calculations," *arXiv preprint arXiv:2007.14822*, 2020. DOI: 10.21468/SciPostPhysCodeb.4 (cit. on p. 41).

[78] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman, *Julia: A fast dynamic language for technical computing*, 2012. arXiv: 1209.5145 [cs.PL] (cit. on p. 41).

[79] K. Bidzhiev *et al.*, *Cloud on-demand emulation of quantum dynamics with tensor networks*, 2023. arXiv: 2302.05253 [quant-ph] (cit. on p. 41).

[80] S. de Lé séleuc, D. Barredo, V. Lienhard, A. Browaeys, and T. Lahaye, "Analysis of imperfections in the coherent optical excitation of single atoms to rydberg states," *Phys. Rev. A*, vol. 97, no. 5, 2018-5. DOI: 10.1103/physreva.97.053803 (cit. on p. 43).

[81] K. Mølmer, Y. Castin, and J. Dalibard, "Monte carlo wave-function method in quantum optics," *Journal of the Optical Society of America B*, vol. 10, no. 3, p. 524, 1993-3, ISSN: 1520-8540. DOI: 10.1364/josab.10.000524 (cit. on pp. 46, 88).

[82] U. Gaubatz, P. Rudecki, S. Schiemann, and K. Bergmann, "Population transfer between molecular vibrational levels by stimulated raman scattering with partially overlapping laser fields. a new concept and experimental results," *The Journal of Chemical Physics*, vol. 92, no. 9, pp. 5363–5376, 1990-5, ISSN: 1089-7690. DOI: 10.1063/1.458514 (cit. on p. 47).

[83] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, "Stimulated raman adiabatic passage in physics, chemistry, and beyond," *Rev. Mod. Phys.*, vol. 89, p. 015 006, 1 2017-3. DOI: 10.1103/RevModPhys.89.015006 (cit. on p. 47).

[84] J. Werschnik and E. K. U. Gross, *Quantum optimal control theory*, 2007. DOI: 10.1088/0953-4075/40/18/R01. arXiv: 0707.1883 [quant-ph] (cit. on p. 51).

[85] C. P. Koch *et al.*, "Quantum optimal control in quantum technologies. strategic report on current status, visions and goals for research in europe," *EPJ Quantum Technology*, vol. 9, no. 1, 2022-7, ISSN: 2196-0763. DOI: 10.1140/epjqt/s40507-022-00138-x (cit. on p. 51).

[86] P. Gross, D. Neuhauser, and H. Rabitz, "Optimal control of curve-crossing systems," *The Journal of Chemical Physics*, vol. 96, no. 4, pp. 2834–2845, 1992-2, ISSN: 1089-7690. DOI: 10.1063/1.461980 (cit. on p. 51).

[87] J. B. Murdoch, A. H. Lent, and M. R. Kritzer, "Computer-optimized narrowband pulses for multislice imaging," *Journal of Magnetic Resonance (1969)*, vol. 74, no. 2, pp. 226–263, 1987-9, ISSN: 0022-2364. DOI: 10.1016/0022-2364(87)90336-2 (cit. on p. 51).

[88] G. A. Paz-Silva and L. Viola, "General transfer-function approach to noise filtering in open-loop quantum control," *Phys. Rev. Lett.*, vol. 113, p. 250 501, 25 2014-12. DOI: 10.1103/PhysRevLett.113.250501 (cit. on pp. 52, 57).

[89] I. Serban, J. Werschnik, and E. K. U. Gross, "Optimal control of time-dependent targets," *Phys. Rev. A*, vol. 71, p. 053 810, 5 2005-5. DOI: 10.1103/PhysRevA.71.053810 (cit. on p. 52).

[90] J. P. Palao, R. Kosloff, and C. P. Koch, "Protecting coherence in optimal control theory: State-dependent constraint approach," *Phys. Rev. A*, vol. 77, p. 063 412, 6 2008-6. DOI: 10.1103/PhysRevA.77.063412 (cit. on p. 52).

[91] V. F. Krotov, "Global methods in optimal control theory," in *Advances in Nonlinear Dynamics and Control: A Report from Russia*. Birkhäuser Boston, 1993, pp. 74–121, ISBN: 9781461203490. DOI: 10.1007/978-1-4612-0349-0_3 (cit. on p. 53).

[92] S. E. Sklarz and D. J. Tannor, "Loading a bose-einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear schrödinger equation," *Phys. Rev. A*, vol. 66, no. 5, 2002-11, ISSN: 1094-1622. DOI: 10.1103/physreva.66.053619 (cit. on p. 53).

[93] O. V. Morzhin and A. N. Pechen, "Krotov method for optimal control of closed quantum systems," *Russian Mathematical Surveys*, vol. 74, no. 5, pp. 851–908, 2019-10, ISSN: 1468-4829. DOI: 10.1070/rm9835 (cit. on p. 53).

[94] M. H. Goerz *et al.*, "Krotov: A Python implementation of Krotov's method for quantum optimal control," *SciPost Phys.*, vol. 7, p. 80, 2019. DOI: 10.21468/SciPostPhys.7.6.080 (cit. on pp. 53, 54).

[95] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, "Fast Quantum Gates for Neutral Atoms," en, *Phys. Rev. Lett.*, vol. 85, no. 10, pp. 2208–2211, 10 2000-9, ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.85.2208 (cit. on p. 53).

[96] H. Levine *et al.*, "Parallel implementation of high-fidelity multiqubit gates with neutral atoms," *Phys. Rev. Lett.*, vol. 123, arXiv:1908.06101, p. 170 503, 17 2019-10. DOI: 10.1103/PhysRevLett.123.170503. arXiv: 1908.06101 [quant-ph] (cit. on pp. 53, 54).

[97] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, "Optimal control of coupled spin dynamics: Design of nmr pulse sequences by gradient ascent algorithms," *Journal of Magnetic Resonance*, vol. 172, no. 2, pp. 296–305, 2005-2, ISSN: 1090-7807. DOI: 10.1016/j.jmr.2004.11.004 (cit. on p. 55).

[98] A. Smith, B. E. Anderson, H. Sosa-Martinez, C. A. Riofrío, I. H. Deutsch, and P. S. Jessen, "Quantum control in the cs $6S_{1/2}$ ground manifold using radio-frequency and microwave magnetic fields," *Phys. Rev. Lett.*, vol. 111, p. 170 502, 17 2013-10. DOI: 10.1103/PhysRevLett.111.170502 (cit. on p. 55).

[99] S. Machnes *et al.*, "Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework," *Phys. Rev. A*, vol. 84, no. 2, 2011-8, ISSN: 1094-1622. DOI: 10.1103/physreva.84.022305 (cit. on p. 55).

[100] H. Ball *et al.*, *Software tools for quantum control: Improving quantum computer performance through noise and error suppression*, 2020. DOI: 10.1088/2058-9565/abdca6. arXiv: 2001.04060 [quant-ph] (cit. on pp. 55, 57).

[101] P. S. Mundada *et al.*, "Experimental benchmarking of an automated deterministic error-suppression workflow for quantum algorithms," *Phys. Rev. Appl.*, vol. 20, p. 024 034, 2 2023-8. DOI: 10.1103/PhysRevApplied.20.024034 (cit. on p. 55).

[102] S. Jandura, J. D. Thompson, and G. Pupillo, "Optimizing rydberg gates for logical-qubit performance," *PRX Quantum*, vol. 4, no. 2, p. 020 336, 2 2023-6, ISSN: 2691-3399. DOI: 10.1103/PRXQuantum.4.020336 (cit. on p. 57).

[103] T. Caneva, T. Calarco, and S. Montangero, "Chopped random-basis quantum optimization," *Phys. Rev. A*, vol. 84, p. 022 326, 2 2011-8. DOI: 10.1103/PhysRevA.84.022326 (cit. on p. 57).

[104] H. Geen and R. Freeman, "Band-selective radiofrequency pulses," *Journal of Magnetic Resonance (1969)*, vol. 93, no. 1, pp. 93–141, 1991-6, ISSN: 0022-2364. DOI: 10.1016/0022-2364(91)90034-q (cit. on p. 58).

[105] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nat. Commun.*, vol. 9, no. 1, 2018-11, ISSN: 2041-1723. DOI: 10.1038/s41467-018-07090-4 (cit. on p. 63).

[106] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, "Connecting ansatz expressibility to gradient magnitudes and barren plateaus," *PRX Quantum*, vol. 3, no. 1, 2022-1, ISSN: 2691-3399. DOI: 10.1103/prxquantum.3.010313 (cit. on p. 63).

[107] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, "Diagnosing barren plateaus with tools from quantum optimal control," *Quantum*, vol. 6, p. 824, 2022-9, ISSN: 2521-327X. DOI: 10.22331/q-2022-09-29-824 (cit. on p. 63).

[108] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Phys. Rev. A*, vol. 99, no. 3, 2019-3, ISSN: 2469-9934. DOI: 10.1103/physreva.99.032331 (cit. on p. 64).

[109] A. F. Izmaylov, R. A. Lang, and T.-C. Yen, "Analytic gradients in variational quantum algorithms: Algebraic extensions of the parameter-shift rule to general unitary transformations," *Phys. Rev. A*, vol. 104, no. 6, 2021-12, ISSN: 2469-9934. DOI: 10.1103/physreva.104.062443 (cit. on p. 64).

[110] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965-1, ISSN: 1460-2067. DOI: 10.1093/comjnl/7.4.308 (cit. on p. 64).

[111] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: Past, present, and future," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091–8126, 2020-10, ISSN: 1573-7721. DOI: 10.1007/s11042-020-10139-6 (cit. on p. 64).

[112] A. G. Gad, "Particle swarm optimization algorithm and its applications: A systematic review," *Archives of Computational Methods in Engineering*, vol. 29, no. 5, pp. 2531–2561, 2022-4, ISSN: 1886-1784. DOI: 10.1007/s11831-021-09694-4 (cit. on p. 64).

[113] F. J. Schreiber, J. Eisert, and J. J. Meyer, "Classical surrogates for quantum learning models," *Phys. Rev. Lett.*, vol. 131, p. 100 803, 10 2023-9. DOI: 10.1103/PhysRevLett.131.100803 (cit. on p. 64).

[114] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, p. 012 020, 2021-10, ISSN: 1573-0565. DOI: 10.1023/A:1010933404324 (cit. on p. 64).

[115] G. James, D. Witten, T. Hastie, and R. Tibshirani, "Statistical learning," in *An Introduction to Statistical Learning: with Applications in R*. New York, NY: Springer New York, 2013, pp. 15–57, ISBN: 978-1-4614-7138-7. DOI: 10.1007/978-1-4614-7138-7_2 (cit. on p. 64).

[116] E. Brochu, V. M. Cora, and N. de Freitas, *A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning*, 2010-12. arXiv: 1012.2599 [cs.LG] (cit. on p. 65).

[117] F. Sauvage and F. Mintert, "Optimal quantum control with poor statistics," *PRX Quantum*, vol. 1, p. 020 322, 2 2020-12. DOI: 10.1103/PRXQuantum.1.020322. arXiv: 1909.01229 [quant-ph] (cit. on p. 65).

[118] C. E. Rasmussen, "Covariance Functions," in *Gaussian Processes for Machine Learning*, O. Bousquet, U. von Luxburg, and G. Rätsch, Eds., Springer Berlin Heidelberg, 2004, pp. 79–104, ISBN: 978-3-540-28650-9. DOI: 10.1007/978-3-540-28650-9_4 (cit. on p. 66).

[119] M. Claesen and B. D. Moor, *Hyperparameter search in machine learning*, 2015. arXiv: 1502.02127 [cs.LG] (cit. on p. 67).

[120] R. H. Byrd, S. L. Hansen, J. Nocedal, and Y. Singer, *A stochastic quasi-newton method for large-scale optimization*, 2014. DOI: 10.1137/140954362. arXiv: 1401.7020 [math.OC] (cit. on p. 69).

[121] L. Breiman, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, ISSN: 0885-6125. DOI: 10.1023/a:1010933404324 (cit. on pp. 70, 182).

[122] E. Ising, "Beitrag zur theorie des ferromagnetismus," *Zeitschrift für Physik*, vol. 31, no. 1, pp. 253–258, 1925-2, ISSN: 0044-3328. DOI: 10.1007/bf02980577 (cit. on p. 75).

[123] S. Blundell, *Magnetism in Condensed Matter* (Oxford Master Series in Condensed Matter Physics), en. London, England: Oxford University Press, 2001-8. DOI: 10.1093/oso/9780198505921.001.0001 (cit. on p. 75).

[124] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6–7, pp. 467–488, 1982-6, ISSN: 1572-9575. DOI: 10.1007/bf02650179 (cit. on p. 75).

[125] C. Monroe *et al.*, "Programmable quantum simulations of spin systems with trapped ions," *Reviews of Modern Physics*, vol. 93, no. 2, p. 025 001, 2 2021-4, ISSN: 1539-0756. DOI: 10.1103/revmodphys.93.025001 (cit. on p. 75).

[126] C. Kokail *et al.*, "Self-verifying variational quantum simulation of lattice models," en, *Nature*, vol. 569, no. 7756, pp. 355–360, 2019-5, ISSN: 1476-4687. DOI: 10.1038/s41586-019-1177-4 (cit. on pp. 75, 123).

[127] R. Barends *et al.*, "Digital quantum simulation of fermionic models with a superconducting circuit," en, *Nat. Commun.*, vol. 6, no. 1, p. 7654, 2015-7. DOI: 10.1038/ncomms8654 (cit. on p. 75).

[128] C. W. Kim, J. M. Nichol, A. N. Jordan, and I. Franco, "Analog quantum simulation of the dynamics of open quantum systems with quantum dots and microelectronic circuits," *PRX Quantum*, vol. 3, p. 040 308, 4 2022-10. DOI: 10.1103/PRXQuantum.3.040308 (cit. on p. 75).

[129] I. Bloch, J. Dalibard, and S. Nascimbène, "Quantum simulations with ultracold quantum gases," en, *Nat. Phys.*, vol. 8, no. 4, pp. 267–276, 2012-4, ISSN: 1745-2481. DOI: 10.1038/nphys2259 (cit. on p. 75).

[130] B. Yan *et al.*, "Observation of dipolar spin-exchange interactions with lattice-confined polar molecules," en, *Nature*, vol. 501, no. 7468, pp. 521–525, 2013-9. DOI: 10.1038/nature12483 (cit. on p. 76).

[131] C. Chen *et al.*, *Spectroscopy of elementary excitations from quench dynamics in a dipolar xy rydberg simulator*, 2023. arXiv: 2311.11726 [cond-mat.quant-gas] (cit. on p. 76).

[132] F. Alet and N. Laflorencie, "Many-body localization: An introduction and selected topics," *Comptes Rendus. Physique*, vol. 19, no. 6, pp. 498–525, 2018-4, ISSN: 1878-1535. DOI: 10.1016/j.crhy.2018.03.003 (cit. on pp. 76, 94).

[133] E. Lieb, T. Schultz, and D. Mattis, "Two soluble models of an antiferromagnetic chain," en, *Annals of Physics*, vol. 16, no. 3, pp. 407–466, 1961-12, ISSN: 0003-4916. DOI: 10.1016/0003-4916(61)90115-4 (cit. on pp. 76, 79).

[134] S. Sachdev, *Quantum Phase Transitions*. Cambridge University Press, 2011-4, ISBN: 9780511973765. DOI: 10.1017/cbo9780511973765 (cit. on pp. 76, 79, 86).

[135] M. H. Amin, "Consistency of the adiabatic theorem," *Phys. Rev. Lett.*, vol. 102, no. 22, p. 220 401, 2009. DOI: 10.1103/PhysRevLett.102.220401 (cit. on p. 80).

[136] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, "The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective," *Physics Reports*, vol. 523, no. 3, pp. 127–205, 2013. DOI: 10.1016/j.physrep.2012.10.002 (cit. on p. 80).

[137] A. Keesling *et al.*, "Quantum kibble–zurek mechanism and critical dynamics on a programmable rydberg simulator," *Nature*, vol. 568, no. 7751, pp. 207–211, 2019-4, ISSN: 1476-4687. DOI: 10.1038/s41586-019-1070-1 (cit. on p. 80).

[138] V. Lienhard *et al.*, "Observing the space- and time-dependent growth of correlations in dynamically tuned synthetic ising models with antiferromagnetic interactions," *Phys. Rev. X*, vol. 8, p. 021 070, 2 2018-6. DOI: 10.1103/PhysRevX.8.021070 (cit. on p. 81).

[139] D. Bluvstein *et al.*, "Controlling quantum many-body dynamics in driven rydberg atom arrays," *Science*, vol. 371, no. 6536, pp. 1355–1359, 2021-3, ISSN: 1095-9203. DOI: 10.1126/science.abg2530 (cit. on p. 91).

[140] J. M. Deutsch, "Eigenstate thermalization hypothesis," *Reports on Progress in Physics*, vol. 81, no. 8, p. 082 001, 2018-7, ISSN: 1361-6633. DOI: 10.1088/1361-6633/aac9f1 (cit. on p. 91).

[141] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, "Colloquium : Many-body localization, thermalization, and entanglement," *Reviews of Modern Physics*, vol. 91, no. 2, 2019-5, ISSN: 1539-0756. DOI: 10.1103/revmodphys.91.021001 (cit. on pp. 91, 95).

[142] M. Schreiber *et al.*, "Observation of many-body localization of interacting fermions in a quasirandom optical lattice," *Science*, vol. 349, no. 6250, pp. 842–845, 2015-8, ISSN: 1095-9203. DOI: 10.1126/science.aaa7432 (cit. on p. 91).

[143] J. Smith *et al.*, "Many-body localization in a quantum simulator with programmable random disorder," *Nat. Phys.*, vol. 12, no. 10, pp. 907–911, 2016-6, ISSN: 1745-2481. DOI: 10.1038/nphys3783 (cit. on pp. 91, 93, 94).

[144] P. Hauke and M. Heyl, "Many-body localization and quantum ergodicity in disordered long-range ising models," *Phys. Rev. B*, vol. 92, no. 13, 2015-10, ISSN: 1550-235X. DOI: 10.1103/physrevb.92.134204 (cit. on p. 93).

[145] H. Li, J. Wang, X.-J. Liu, and H. Hu, "Many-body localization in ising models with random long-range interactions," *Phys. Rev. A*, vol. 94, no. 6, 2016-12, ISSN: 2469-9934. DOI: 10.1103/physreva.94.063625 (cit. on p. 93).

[146] T. B. Wahl, A. Pal, and S. H. Simon, "Signatures of the many-body localized regime in two dimensions," *Nat. Phys.*, vol. 15, no. 2, pp. 164–169, 2018-11, ISSN: 1745-2481. DOI: 10.1038/s41567-018-0339-x (cit. on p. 95).

[147] J. Knolle and R. Moessner, "A field guide to spin liquids," *Annual Review of Condensed Matter Physics*, vol. 10, no. 1, pp. 451–472, 2019-3, ISSN: 1947-5462. DOI: 10.1146/annurev-conmatphys-031218-013401 (cit. on p. 97).

[148] L. Pollet, J. D. Picon, H. P. Büchler, and M. Troyer, "Supersolid phase with cold polar molecules on a triangular lattice," *Phys. Rev. Lett.*, vol. 104, p. 125 302, 12 2010-3. DOI: 10.1103/PhysRevLett.104.125302 (cit. on p. 97).

[149] B. Capogrosso-Sansone, C. Trefzger, M. Lewenstein, P. Zoller, and G. Pupillo, "Quantum phases of cold polar molecules in 2d optical lattices," *Phys. Rev. Lett.*, vol. 104, p. 125 301, 12 2010-3. DOI: 10.1103/PhysRevLett.104.125301 (cit. on p. 97).

[150] E. G. Dalla Torre, E. Berg, and E. Altman, "Hidden order in 1d bose insulators," *Phys. Rev. Lett.*, vol. 97, p. 260 401, 26 2006-12. DOI: 10.1103/PhysRevLett.97.260401 (cit. on p. 97).

[151] C. Castelnovo, R. Moessner, and S. L. Sondhi, "Magnetic monopoles in spin ice," *Nature*, vol. 451, no. 7174, pp. 42–45, 2008-1, ISSN: 1476-4687. DOI: 10.1038/nature06433 (cit. on p. 97).

[152] B. Sbierski, M. Bintz, S. Chatterjee, M. Schuler, N. Y. Yao, and L. Pollet, *Magnetism in the two-dimensional dipolar xy model*, 2023. DOI: 10.1103/PhysRevB.109.144411. arXiv: 2305.03673 [cond-mat.quant-gas] (cit. on pp. 97, 114).

[153] pp. 193–216, 1965. DOI: 10.1016/b978-0-08-010586-4.50034-1 (cit. on p. 98).

[154] L. D. Landau and V. L. Ginzburg, "On the theory of superconductivity," *Zh. Eksp. Teor. Fiz.*, vol. 20, p. 1064, 1950. DOI: 10.1007/978-3-540-68008-6_4 (cit. on p. 98).

[155] J. Kepler, *De Nive Sexangula*. Frankfurt: Gottfried Tampach, 1611 (cit. on p. 98).

[156] J. Goldstone, "Field theories with "Superconductor" solutions," *Il Nuovo Cimento*, vol. 19, pp. 154–164, 1961. DOI: 10.1007/BF02812722 (cit. on p. 98).

[157] H. Tasaki, *Physics and Mathematics of Quantum Many-Body Systems*. Cham: Springer International Publishing, 2020. DOI: 10.1007/978-3-030-41265-4 (cit. on pp. 98, 112).

[158] F. Bloch, "Zur Theorie des Ferromagnetismus," *Zeitschrift für Physik*, vol. 61, pp. 206–219, 1930. DOI: 10.1007/BF01339661 (cit. on p. 98).

[159] R. Peierls, "Quelques propriétés typiques des corps solides," *Annales de l'institut Henri Poincaré*, vol. 5, pp. 177–122, 1935 (cit. on p. 98).

[160] N. D. Mermin and H. Wagner, "Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models," *Phys. Rev. Lett.*, vol. 17, pp. 1133–1136, 1966. DOI: 10.1103/PhysRevLett.17.1133 (cit. on p. 98).

[161] P. C. Hohenberg, "Existence of Long-Range Order in One and Two Dimensions," *Phys. Rev.*, vol. 158, pp. 383–386, 1967. DOI: 10.1103/PhysRev.158.383 (cit. on p. 98).

[162] P. Bruno, "Absence of spontaneous magnetic order at nonzero temperature in one- and two-dimensional heisenberg and XY systems with long-range interactions," *Phys. Rev. Lett.*, vol. 87, p. 137 203, 13 2001-9. DOI: 10.1103/PhysRevLett.87.137203 (cit. on pp. 98, 100, 110).

[163] F. J. Dyson, "Existence of a phase-transition in a one-dimensional Ising ferromagnet," *Communications in Mathematical Physics*, vol. 12, pp. 91–107, 1969. DOI: 10.1007/BF01645907 (cit. on p. 98).

[164] H. Kunz and C. E. Pfister, "First order phase transition in the plane rotator ferromagnetic model in two dimensions," *Communications in Mathematical Physics*, vol. 46, pp. 245–251, 1976. DOI: 10.1007/BF01609121 (cit. on pp. 98, 100, 113).

[165] S. V. Maleev, "Dipole forces in two-dimensional and layered ferromagnets," *Soviet Journal of Experimental and Theoretical Physics*, vol. 43, p. 1240, 1976 (cit. on pp. 98, 100).

[166] J. Fröhlich, R. Israel, E. H. Lieb, and B. Simon, "Phase transitions and reflection positivity. I. General theory and long range lattice models," *Communications in Mathematical Physics*, vol. 62, pp. 1–34, 1978. DOI: 10.1007/BF01940327 (cit. on p. 98).

[167] H. T. Diep, Ed., *Frustrated Spin Systems*, 2nd. New Jersey: World Scientific, 2013. DOI: 10.1142/8676 (cit. on p. 98).

[168] C. Castelnovo, R. Moessner, and S. L. Sondhi, "Magnetic monopoles in spin ice," *Nature*, vol. 451, pp. 42–45, 2008. DOI: 10.1038/nature06433 (cit. on p. 98).

[169] N. Y. Yao, M. P. Zaletel, D. M. Stamper-Kurn, and A. Vishwanath, "A quantum dipolar spin liquid," *Nat. Phys.*, vol. 14, pp. 405–410, 2018. DOI: 10.1038/s41567-017-0030-7 (cit. on p. 98).

[170] A. Kele ş and E. Zhao, "Absence of long-range order in a triangular spin system with dipolar interactions," *Phys. Rev. Lett.*, vol. 120, p. 187 202, 18 2018-5. DOI: 10.1103/PhysRevLett.120.187202 (cit. on p. 98).

[171] A. Kele ş and E. Zhao, "Renormalization group analysis of dipolar heisenberg model on square lattice," *Phys. Rev. B*, vol. 97, p. 245 105, 24 2018-6. DOI: 10.1103/PhysRevB.97.245105 (cit. on p. 98).

[172] N. Defenu *et al.*, "Long-range interacting quantum systems," *arXiv:2109.01063*, 2021. DOI: 10.1103/RevModPhys.95.035002 (cit. on pp. 98, 108, 110).

[173] K. De'Bell, A. B. MacIsaac, and J. P. Whitehead, "Dipolar effects in magnetic thin films and quasi-two-dimensional systems," *Rev. Mod. Phys.*, vol. 72, pp. 225–257, 2000. DOI: 10.1103/RevModPhys.72.225 (cit. on p. 98).

[174] A. Taroni, S. T. Bramwell, and P. C. W. Holdsworth, "Universal window for two-dimensional critical exponents," *Journal of Physics: Condensed Matter*, vol. 20, p. 275 233, 2008. DOI: 10.1088/0953-8984/20/27/275233 (cit. on p. 98).

[175] A. Mazurenko *et al.*, "A cold-atom fermi–hubbard antiferromagnet," *Nature*, vol. 545, no. 7655, pp. 462–466, 2017-5, ISSN: 1476-4687. DOI: 10.1038/nature22362 (cit. on p. 98).

[176] B. Yan *et al.*, "Observation of dipolar spin-exchange interactions with lattice-confined polar molecules," *Nature*, vol. 501, pp. 521–525, 2013. DOI: 10.1038/nature12483 (cit. on p. 98).

[177] L. Christakis *et al.*, "Probing site-resolved correlations in a spin system of ultracold molecules," *Nature*, vol. 614, pp. 64–69, 2023. DOI: 10.1038/s41586-022-05558-4 (cit. on p. 98).

[178] L. Chomaz *et al.*, "Dipolar physics: A review of experiments with magnetic quantum gases," *arXiv:2201.02672*, 2022. DOI: 10.1088/1361-6633/aca814 (cit. on p. 98).

[179] P. Richerme *et al.*, "Non-local propagation of correlations in quantum systems with long-range interactions," *Nature*, vol. 511, pp. 198–201, 2014. DOI: 10.1038/nature13450 (cit. on p. 98).

[180] P. Jurcevic *et al.*, "Quasiparticle engineering and entanglement propagation in a quantum many-body system," *Nature*, vol. 511, pp. 202–205, 2014. DOI: 10.1038/nature13461 (cit. on p. 98).

[181] M. F. Maghrebi, Z.-X. Gong, and A. V. Gorshkov, "Continuous symmetry breaking in 1d long-range interacting quantum systems," *Phys. Rev. Lett.*, vol. 119, p. 023 001, 2017. DOI: 10.1103/PhysRevLett.119.023001 (cit. on p. 98).

[182] L. Feng *et al.*, "Continuous symmetry breaking in a trapped-ion spin chain," *Nature*, vol. 623, no. 7988, pp. 713–717, 2023-11, ISSN: 1476-4687. DOI: 10.1038/s41586-023-06656-7 (cit. on p. 98).

[183] P. Schauß *et al.*, "Crystallization in ising quantum magnets," *Science*, vol. 347, no. 6229, pp. 1455–1458, 2015-3, ISSN: 1095-9203. DOI: 10.1126/science.1258351 (cit. on p. 98).

[184] E. Guardado-Sanchez *et al.*, "Probing the quench dynamics of antiferromagnetic correlations in a 2d quantum ising spin system," *Phys. Rev. X*, vol. 8, p. 021 069, 2 2018-6. DOI: 10.1103/PhysRevX.8.021069 (cit. on p. 98).

[185] S. Ebadi *et al.*, "Quantum phases of matter on a 256-atom programmable quantum simulator," *Nature*, vol. 595, pp. 227–232, 2021. DOI: 10.1038/s41586-021-03582-4 (cit. on p. 98).

[186] C. N. Yang, "Concept of off-diagonal long-range order and the quantum phases of liquid He and of superconductors," *Reviews of Modern Physics*, vol. 34, pp. 694–704, 1962. DOI: 10.1103/RevModPhys.34.694 (cit. on pp. 99, 107).

[187] H. Tasaki, "Long-range order, "tower" of states, and symmetry breaking in lattice quantum systems," *Journal of Statistical Physics*, vol. 174, pp. 735–761, 2019. DOI: 10.1007/s10955-018-2193-8 (cit. on pp. 100, 114).

[188] P. W. Anderson, "An approximate quantum theory of the antiferromagnetic ground state," *Phys. Rev.*, vol. 86, pp. 694–701, 1952. DOI: 10.1103/PhysRev.86.694 (cit. on pp. 100, 114).

[189] P. Anderson, *Basic Notions of Condensed Matter Physics*, 10. pr. Cambridge, Mass: Perseus Publ, 2010. DOI: 10.4324/9780429494116 (cit. on pp. 100, 114).

[190] A. Beekman, L. Rademaker, and J. van Wezel, "An introduction to spontaneous symmetry breaking," *SciPost Physics Lecture Notes*, 2019-12, ISSN: 2590-1990. DOI: 10.21468/scipostphyslectnotes.11. [Online]. Available: http://dx.doi.org/10.21468/SciPostPhysLectNotes.11 (cit. on pp. 100, 114).

[191] B. I. Halperin, "On the hohenberg–mermin–wagner theorem and its limitations," *Journal of Statistical Physics*, vol. 175, no. 3–4, pp. 521–529, 2018-12, ISSN: 1572-9613. DOI: 10.1007/s10955-018-2202-y (cit. on p. 100).

[192] V. L. Berezinskiĭ, "Destruction of Long-range Order in One-dimensional and Two-dimensional Systems having a Continuous Symmetry Group I. Classical Systems," *Soviet Journal of Experimental and Theoretical Physics*, vol. 32, p. 493, 1971 (cit. on p. 100).

[193] V. L. Berezinskiĭ, "Destruction of Long-range Order in One-dimensional and Two-dimensional Systems Possessing a Continuous Symmetry Group. II. Quantum Systems," *Soviet Journal of Experimental and Theoretical Physics*, vol. 34, p. 610, 1972 (cit. on p. 100).

[194] J. M. Kosterlitz and D. J. Thouless, "Ordering, metastability and phase transitions in two-dimensional systems," *Journal of Physics C: Solid State Physics*, vol. 6, pp. 1181–1203, 1973. DOI: 10.4324/9780429494116-24 (cit. on p. 100).

[195] J. M. Kosterlitz, "The critical properties of the two-dimensional XY model," *Journal of Physics C: Solid State Physics*, vol. 7, pp. 1046–1060, 1974. DOI: 10.1088/0022-3719/7/6/005 (cit. on p. 100).

[196] G. Giachetti, N. Defenu, S. Ruffo, and A. Trombettoni, "Berezinskii-kosterlitz-thouless phase transitions with long-range couplings," *Phys. Rev. Lett.*, vol. 127, p. 156 801, 15 2021-10. DOI: 10.1103/PhysRevLett.127.156801 (cit. on p. 100).

[197] A. S. Sørensen and et al., "Adiabatic preparation of many-body states in optical lattices," *Phys. Rev. A*, vol. 81, p. 061 603, 2010. DOI: 10.1103/PhysRevA.81.061603 (cit. on p. 101).

[198] D. Barredo, S. de Léséleuc, V. Lienhard, T. Lahaye, and A. Browaeys, "An atom-by-atom assembler of defect-free arbitrary two-dimensional atomic arrays," *Science*, vol. 354, no. 6315, pp. 1021–1023, 2016-11, ISSN: 1095-9203. DOI: 10.1126/science.aah3778 (cit. on p. 102).

[199] S. de Léséleuc, D. Barredo, V. Lienhard, A. Browaeys, and T. Lahaye, "Analysis of imperfections in the coherent optical excitation of single atoms to rydberg states," *Phys. Rev. A*, vol. 97, p. 053 803, 2018. DOI: 10.1103/PhysRevA.97.053803 (cit. on p. 105).

[200] A. W. Sandvik and C. J. Hamer, "Ground-state parameters, finite-size scaling, and low-temperature properties of the two-dimensional $S = \frac{1}{2}$ XY model," *Phys. Rev. B*, vol. 60, pp. 6588–6593, 9 1999-9. DOI: 10.1103/PhysRevB.60.6588 (cit. on p. 108).

[201] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.*, vol. 69, pp. 2863–2866, 1992. DOI: 10.1103/PhysRevLett.69.2863 (cit. on p. 109).

[202] J. Hauschild and F. Pollmann, "Efficient numerical simulations with tensor networks: Tensor network python (tenpy)," *SciPost Physics Lecture Notes*, p. 5, 2018. DOI: 10.21468/SciPostPhysLectNotes.5 (cit. on p. 109).

[203] E. M. Stoudenmire and S. R. White, "Minimally entangled typical thermal state algorithms," *New Journal of Physics*, vol. 12, p. 055 026, 2010. DOI: 10.1088/1367-2630/12/5/055026 (cit. on p. 113).

[204] T. Comparin, F. Mezzacapo, and T. Roscilde, "Robust spin squeezing from the tower of states of u(1)-symmetric spin hamiltonians," *Phys. Rev. A*, vol. 105, p. 022 625, 2022. DOI: 10.1103/PhysRevA.105.022625 (cit. on p. 114).

[205] G. Bornet *et al.*, "Scalable spin squeezing in a dipolar rydberg atom array," *Nature*, vol. 621, no. 7980, pp. 728–733, 2023-8, ISSN: 1476-4687. DOI: 10.1038/s41586-023-06414-9 (cit. on pp. 114, 116).

[206] T. Fukuhara *et al.*, "Microscopic observation of magnon bound states and their dynamics," *Nature*, vol. 502, no. 7469, pp. 76–79, 2013-10, ISSN: 1476-4687. DOI: 10.1038/nature12541 (cit. on p. 114).

[207] P. T. Dumitrescu *et al.*, "Dynamical topological phase realized in a trapped-ion quantum simulator," *Nature*, vol. 607, no. 7919, pp. 463–467, 2022-7, ISSN: 1476-4687. DOI: 10.1038/s41586-022-04853-4 (cit. on p. 114).

[208] P. Roushan *et al.*, "Chiral ground-state currents of interacting photons in a synthetic magnetic field," *Nat. Phys.*, vol. 13, no. 2, pp. 146–151, 2017-2, ISSN: 1745-2481. DOI: 10.1038/nphys3930 (cit. on pp. 114, 119, 123).

[209] J. Zhang *et al.*, "Observation of a Discrete Time Crystal," *Nature*, vol. 543, no. 7644, pp. 217–220, 2017-3, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature21413. arXiv: 1609.08684 [cond-mat] (cit. on p. 114).

[210] E. Knill *et al.*, "Randomized benchmarking of quantum gates," *Phys. Rev. A*, vol. 77, p. 012 307, 1 2008-1. DOI: 10.1103/PhysRevA.77.012307 (cit. on p. 114).

[211] J. M. Gambetta *et al.*, "Characterization of addressability by simultaneous randomized benchmarking," *Phys. Rev. Lett.*, vol. 109, p. 240 504, 24 2012-12. DOI: 10.1103/PhysRevLett.109.240504 (cit. on p. 114).

[212] J. P. Gaebler *et al.*, "Randomized benchmarking of multiqubit gates," *Phys. Rev. Lett.*, vol. 108, p. 260 503, 26 2012-6. DOI: 10.1103/PhysRevLett.108.260503 (cit. on p. 114).

[213] G. Pagano *et al.*, "Quantum approximate optimization of the long-range ising model with a trapped-ion quantum simulator," *Proceedings of the National Academy of Sciences*, vol. 117, no. 41, pp. 25 396–25 401, 2020-10. DOI: 10.1073/pnas.2006373117 (cit. on p. 114).

[214] D. D. Yavuz *et al.*, "Fast ground state manipulation of neutral atoms in microscopic optical traps," *Phys. Rev. Lett.*, vol. 96, p. 063 001, 6 2006-2. DOI: 10.1103/PhysRevLett.96.063001 (cit. on p. 114).

[215] M. P. A. Jones *et al.*, "Fast quantum state control of a single trapped neutral atom," *Phys. Rev. A*, vol. 75, p. 040 301, 4 2007-4. DOI: 10.1103/PhysRevA.75.040301(R) (cit. on p. 114).

[216] L. Isenhower *et al.*, "Demonstration of a neutral atom controlled-not quantum gate," *Phys. Rev. Lett.*, vol. 104, p. 010 503, 1 2010-1. DOI: 10.1103/PhysRevLett.104.010503 (cit. on p. 114).

[217] T. Xia *et al.*, "Randomized benchmarking of single-qubit gates in a 2d array of neutral-atom qubits," *Phys. Rev. Lett.*, vol. 114, p. 100 503, 10 2015-3. DOI: 10.1103/PhysRevLett.114.100503 (cit. on p. 114).

[218] G. Birkl and J. Fortágh, "Micro traps for quantum information processing and precision force sensing," *Laser & Photonics Reviews*, vol. 1, no. 1, pp. 12–23, 2007. DOI: https://doi.org/10.1002/lpor.200610002. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.200610002 (cit. on p. 114).

[219] K. Takeda, A. Noiri, T. Nakajima, J. Yoneda, T. Kobayashi, and S. Tarucha, "Quantum tomography of an entangled three-qubit state in silicon," *Nature Nanotechnology*, vol. 16, no. 9, pp. 965–969, 2021-9, ISSN: 1748-3395. DOI: 10.1038/s41565-021-00925-0 (cit. on p. 118).

[220] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera, "Classification of mixed three-qubit states," *Phys. Rev. Lett.*, vol. 87, p. 040 401, 4 2001-7. DOI: 10.1103/PhysRevLett.87.040401 (cit. on p. 118).

[221] O. Gühne *et al.*, "Experimental detection of entanglement via witness operators and local measurements," *Journal of Modern Optics*, vol. 50, no. 6–7, pp. 1079–1102, 2003-4, ISSN: 1362-3044. DOI: 10.1080/09500340308234554 (cit. on p. 118).

[222] M. Neeley *et al.*, "Generation of three-qubit entangled states using superconducting phase qubits," *Nature*, vol. 467, no. 7315, pp. 570–573, 2010-9, ISSN: 1476-4687. DOI: 10.1038/nature09418 (cit. on p. 118).

[223] N. D. Mermin, "Extreme quantum entanglement in a superposition of macroscopically distinct states," *Phys. Rev. Lett.*, vol. 65, pp. 1838–1840, 15 1990-10. DOI: 10.1103/PhysRevLett.65.1838 (cit. on p. 118).

[224] D. I. Tsomokos, J. J. García-Ripoll, N. R. Cooper, and J. K. Pachos, "Chiral entanglement in triangular lattice models," *Phys. Rev. A*, vol. 77, p. 012 106, 1 2008-1. DOI: 10.1103/PhysRevA.77.012106 (cit. on p. 119).

[225] X. G. Wen, F. Wilczek, and A. Zee, "Chiral spin states and superconductivity," *Phys. Rev. B*, vol. 39, pp. 11 413–11 423, 16 1989-6. DOI: 10.1103/PhysRevB.39.11413 (cit. on p. 123).

[226] S. Notarnicola, A. Elben, T. Lahaye, A. Browaeys, S. Montangero, and B. Vermersch, "A randomized measurement toolbox for an interacting rydberg-atom quantum simulator," *New Journal of Physics*, vol. 25, no. 10, p. 103 006, 2023-10. DOI: 10.1088/1367-2630/acfcd3 (cit. on p. 123).

[227] T. Brydges *et al.*, "Probing rényi entanglement entropy via randomized measurements," *Science*, vol. 364, no. 6437, pp. 260–263, 2019. DOI: 10.1126/science.aau4963. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.aau4963 (cit. on p. 123).

[228] L. C. Freeman, "Visualizing social networks," *J. Soc. Struct.*, vol. 1, 2000. [Online]. Available: https://www.cmu.edu/joss/content/articles/volume1/Freeman.html (cit. on p. 127).

[229] A. Theocharidis, S. van Dongen, A. J. Enright, and T. C. Freeman, "Network visualization and analysis of gene expression data using biolayout express3d," *Nature Protocols*, vol. 4, pp. 1535–1550, 2009. DOI: 10.1038/nprot.2009.177 (cit. on p. 127).

[230] R. F. i. Cancho and R. V. Solé, "The small world of human language," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 268, no. 1482, pp. 2261–2265, 2001-11, ISSN: 1471-2954. DOI: 10.1098/rspb.2001.1800 (cit. on p. 127).

[231] J. Wurtz, P. L. S. Lopes, N. Gemelke, A. Keesling, and S. Wang, "Industry applications of neutral-atom quantum computing solving independent set problems," *arXiv e-prints*, arXiv:2205.08500, arXiv:2205.08500, 2022-5. DOI: 10.48550/arXiv.2205.08500. arXiv: 2205.08500 [quant-ph] (cit. on p. 127).

[232] C. Dalyac *et al.*, *Graph algorithms with neutral atom quantum processors*, 2024. eprint: arXiv:2403.11931 (cit. on p. 128).

[233] M. R. Garey and D. S. Johnson, ""strong"np-completeness results: Motivation, examples, and implications," *Journal of the ACM (JACM)*, vol. 25, no. 3, pp. 499–508, 1978. DOI: 10.1145/322077.322090 (cit. on p. 129).

[234] G. A. Kochenberger, F. Glover, B. Alidaee, and C. Rego, "A unified modeling and solution framework for combinatorial optimization problems," *Or Spectrum*, vol. 26, no. 2, pp. 237–250, 2004. DOI: 10.1007/s00291-003-0153-3 (cit. on p. 129).

[235] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, "Optimization by simulated annealing," en, *Science*, vol. 220, no. 4598, pp. 671–680, 1983-5. DOI: 10.1126/science.220.4598.671 (cit. on p. 130).

[236] M. Mezard and A. Montanari, *Information, physics, and computation*. Oxford University Press, 2009. DOI: 10.1093/acprof:oso/9780198570837.001.0001 (cit. on p. 130).

[237] B. Berger and T. Leighton, "Protein folding in the hydrophobic-hydrophilic (hp) is np-complete," in *Proceedings of the second annual international conference on Computational molecular biology*, 1998, pp. 30–39. DOI: 10.1145/279069.279080 (cit. on p. 130).

[238] A. Lucas, "Ising formulations of many np problems," *Frontiers in physics*, vol. 2, p. 5, 2014. DOI: 10.3389/fphy.2014.00005 (cit. on p. 130).

[239] A. Coja-Oghlan and C. Efthymiou, "On independent sets in random graphs," *Random Structures and Algorithms*, vol. 47, no. 3, pp. 436–486, 2014-5, ISSN: 1042-9832. DOI: 10.1002/rsa.20550 (cit. on p. 131).

[240] E. N. Gilbert, "Random plane networks," *Journal of the society for industrial and applied mathematics*, vol. 9, no. 4, pp. 533–543, 1961. DOI: 10.1137/0109045 (cit. on p. 132).

[241] J. Xu and B. Berger, "Fast and accurate algorithms for protein side-chain packing," *Journal of the ACM (JACM)*, vol. 53, no. 4, pp. 533–557, 2006. DOI: 10.1145/1162349.1162350 (cit. on p. 132).

[242] T. Erlebach, T. Hagerup, K. Jansen, M. Minzlaff, and A. Wolff, "Trimming of graphs, with application to point labeling," *Theory of Computing Systems*, vol. 47, pp. 613–636, 2010. DOI: 10.1007/s00224-009-9184-8 (cit. on p. 132).

[243] D. S. Hochbaum and W. Maass, "Approximation schemes for covering and packing problems in image processing and vlsi," *Journal of the ACM (JACM)*, vol. 32, no. 1, pp. 130–136, 1985. DOI: 10.1145/2455.214106 (cit. on p. 132).

[244] L. M. K. Vandersypen and I. L. Chuang, "NMR techniques for quantum control and computation," *Reviews of Modern Physics*, vol. 76, no. 4, pp. 1037–1069, 2004-10. DOI: 10.1103/RevModPhys.76.1037. arXiv: quant-ph/0404064 [quant-ph] (cit. on p. 133).

[245] J. Zeiher *et al.*, "Many-body interferometry of a Rydberg-dressed spin lattice," *Nat. Phys.*, vol. 12, no. 12, pp. 1095–1099, 2016-12. DOI: 10.1038/nphys3835. arXiv: 1602.06313 [cond-mat.quant-gas] (cit. on p. 133).

[246] S. Hollerith *et al.*, "Realizing Distance-Selective Interactions in a Rydberg-Dressed Atom Array," *Phys. Rev. Lett.*, vol. 128, no. 11, 113602, p. 113 602, 2022-3. DOI: 10.1103/PhysRevLett.128.113602. arXiv: 2110.10125 [cond-mat.quant-gas] (cit. on p. 133).

[247] H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin, "Quantum optimization for maximum independent set using rydberg atom arrays," *ArXiv:1808.10816*, arXiv:1808.10816, arXiv:1808.10816, 2018-8. DOI: 10.48550/ARXIV.1808.10816. arXiv: 1808.10816 [quant-ph] (cit. on pp. 134, 153).

[248] C. Dalyac *et al.*, "Qualifying quantum approaches for hard industrial optimization problems. a case study in the field of smart-charging of electric vehicles," *EPJ Quantum Technology*, vol. 8, no. 1, p. 12, 2021-5, ISSN: 2196-0763. DOI: 10.1140/epjqt/s40507-021-00100-3 (cit. on pp. 136, 148).

[249] M. Kim, K. Kim, J. Hwang, E.-G. Moon, and J. Ahn, "Rydberg quantum wires for maximum independent set problems," *Nat. Phys.*, vol. 18, no. 7, pp. 755–759, 2022-7, ISSN: 1745-2481. DOI: 10.1038/s41567-022-01629-5 (cit. on pp. 136, 154).

[250] S. Ebadi *et al.*, "Quantum optimization of maximum independent set using Rydberg atom arrays," *Science*, vol. 376, no. 6598, pp. 1209–1215, 2022-6. DOI: 10.1126/science.abo6587. arXiv: 2202.09372 [quant-ph]. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.abo6587 (cit. on p. 136).

[251] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices," *Phys. Rev. X*, vol. 10, p. 021 067, 2 2020-6. DOI: 10.1103/PhysRevX.10.021067 (cit. on pp. 136, 138).

[252] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, "A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem," *Science*, vol. 292, no. 5516, pp. 472–475, 2001-4. DOI: 10.1126/science.1057726 (cit. on p. 137).

[253] K. Blekos *et al.*, "A review on quantum approximate optimization algorithm and its variants," *Physics Reports*, vol. 1068, pp. 1–66, 2024-6, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2024.03.002 (cit. on p. 138).

[254] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *ArXiv:1411.4028*, arXiv:1411.4028, arXiv:1411.4028, 2014-11. DOI: 10.48550/ARXIV.1411.4028. arXiv: 1411.4028 [quant-ph] (cit. on pp. 138, 187).

[255] H.-D. Liu, Y.-H. Chen, H.-F. Lin, H.-S. Tao, and W.-M. Liu, "Antiferromagnetic metal and mott transition on shastry-sutherland lattice," *Scientific Reports*, vol. 4, no. 1, 2014-4, ISSN: 2045-2322. DOI: 10.1038/srep04829 (cit. on p. 143).

[256] R. Boppana and M. M. Halldórsson, "Approximating maximum independent sets by excluding subgraphs," *BIT*, vol. 32, no. 2, pp. 180–196, 1992-6, ISSN: 1572-9125. DOI: 10.1007/bf01994876 (cit. on p. 148).

[257] M. A. H. Rafi and J. Bauman, "A comprehensive review of dc fast-charging stations with energy storage: Architectures, power converters, and analysis," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 2, pp. 345–368, 2020. DOI: 10.1109/TTE.2020.3015743 (cit. on p. 149).

[258] C. Gschwendtner, S. R. Sinsel, and A. Stephan, "Vehicle-to-x (v2x) implementation: An overview of predominate trial configurations and technical, social and regulatory challenges," *Renewable and Sustainable Energy Reviews*, vol. 145, p. 110 977, 2021. DOI: 10.1016/j.rser.2021.110977 (cit. on p. 149).

[259] A. W. Kolen, J. K. Lenstra, C. H. Papadimitriou, and F. C. Spieksma, *Interval scheduling: A survey*, www.interscience.wiley.com, Wiley InterScience, 2007. DOI: 10.1002/nav.20231 (cit. on p. 149).

[260] F. C. R. Spieksma, "On the approximability of an interval scheduling problem," *J. of Scheduling*, vol. 2(5), pp. 215–227, 1999. DOI: 10.1002/(SICI)1099-1425(199909/10)2:5%3C215::AID-JOS27%3E3.0.CO;2-Y (cit. on p. 149).

[261] *Belib réseau parisien de bornes de recharges pour véhicules électriques*, www.data.gouv.fr/fr/datasets/belib-reseau-parisien-de-bornes-de-recharges-accelerees-22-kw-ac-dc-pour-vehicules-electriques/, 2017 (cit. on p. 149).

[262] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991. DOI: https://doi.org/10.1002/spe.4380211102. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/spe.4380211102. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.4380211102 (cit. on pp. 149, 150).

[263] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization: theoretical and practical aspects*. Springer Science & Business Media, 2006. DOI: 10.1007/978-3-662-05078-1 (cit. on p. 150).

[264] M. D. J. C., *An Example Inference Task: Clustering*. Cambridge University press, 2004. [Online]. Available: https://www.inference.org.uk/mackay/itprnn/ps/284.292.pdf (cit. on p. 151).

[265] M. C. Golumbic, "Foreword 2004: The annals edition," in *Algorithmic Graph Theory and Perfect Graphs*, ser. Annals of discrete mathematics, Elsevier, 2004, pp. xiii–xiv. DOI: 10.1016/S0167-5060(04)80043-8 (cit. on p. 153).

[266] H. B. Hunt III, M. V. Marathe, V. Radhakrishnan, S. S. Ravi, D. J. Rosenkrantz, and R. E. Stearns, "Nc-approximation schemes for np-and pspace-hard problems for geometric graphs," *Journal of algorithms*, vol. 26, no. 2, pp. 238–274, 1998. DOI: 10.1006/jagm.1997.0903 (cit. on p. 153).

[267] M. F. Serret, B. Marchand, and T. Ayral, "Solving optimization problems with Rydberg analog quantum computers: Realistic requirements for quantum advantage using noisy simulation and classical benchmarks," *Phys. Rev. A*, vol. 102, no. 5, 052617, p. 052 617, 2020-11. DOI: 10.1103/PhysRevA.102.052617. arXiv: 2006.11190 [quant-ph] (cit. on pp. 153, 203).

[268] M. Cain, S. Chattopadhyay, J.-G. Liu, R. Samajdar, H. Pichler, and M. D. Lukin, "Quantum speedup for combinatorial optimization with flat energy landscapes," 2023. arXiv: 2306.13123 [quant-ph] (cit. on p. 153).

[269] H. Wang, H.-C. Yeh, and A. Kamenev, "Many-body localization enables iterative quantum optimization," *Nat. Commun.*, vol. 13, no. 1, p. 5503, 2022. DOI: 10.1038/s41467-022-33179-y (cit. on p. 153).

[270] A. D. King *et al.*, "Quantum critical dynamics in a 5,000-qubit programmable spin glass," *Nature*, pp. 1–6, 2023. DOI: 10.1038/s41586-023-05867-2 (cit. on p. 153).

[271] B. F. Schiffer, D. S. Wild, N. Maskara, M. Cain, M. D. Lukin, and R. Samajdar, "Circumventing superexponential runtimes for hard instances of quantum adiabatic optimization," *arXiv preprint arXiv:2306.13131*, 2023. DOI: 10.1103/PhysRevResearch.6.013271 (cit. on p. 153).

[272] R. S. Andrist *et al.*, "Hardness of the maximum independent set problem on unit-disk graphs and prospects for quantum speedups," *arXiv preprint arXiv:2307.09442*, 2023. DOI: 10.1103/PhysRevResearch.5.043277 (cit. on p. 153).

[273] A. Byun, M. Kim, and J. Ahn, "Finding the maximum independent sets of platonic graphs using rydberg atoms," *PRX Quantum*, vol. 3, p. 030 305, 3 2022-7. DOI: 10.1103/PRXQuantum.3.030305 (cit. on p. 154).

[274] C. Dalyac, L.-P. Henry, M. Kim, J. Ahn, and L. Henriet, "Exploring the impact of graph locality for the resolution of MIS with neutral atom devices," *arXiv e-prints*, arXiv:2306.13373, arXiv:2306.13373, 2023-6. DOI: 10.48550/arXiv.2306.13373. arXiv: 2306.13373 [quant-ph] (cit. on p. 154).

[275] S. Stastny, H. P. Büchler, and N. Lang, "Functional completeness of planar rydberg blockade structures," *Phys. Rev. B*, vol. 108, p. 085 138, 8 2023-8. DOI: 10.1103/PhysRevB.108.085138 (cit. on p. 155).

[276] M.-T. Nguyen, J.-G. Liu, J. Wurtz, M. D. Lukin, S.-T. Wang, and H. Pichler, "Quantum optimization with arbitrary connectivity using rydberg atom arrays," *PRX Quantum*, vol. 4, arXiv:2209.03965, p. 010 316, 1 2023-2. DOI: 10.1103/PRXQuantum.4.010316. arXiv: 2209.03965 [quant-ph] (cit. on p. 155).

[277] M. Lanthaler, C. Dlaska, K. Ender, and W. Lechner, "Rydberg-blockade-based parity quantum optimization," *Phys. Rev. Lett.*, vol. 130, no. 22, p. 220 601, 2023. DOI: 10.1103/PhysRevLett.130.220601 (cit. on p. 155).

[278] A. Byun *et al.*, *Rydberg-atom graphs for quadratic unconstrained binary optimization problems*, 2023. arXiv: 2309.14847 [quant-ph] (cit. on p. 155).

[279] S. Jeong, M. Kim, M. Hhan, and J. Ahn, "Quantum Programming of the Satisfiability Problem with Rydberg Atom Graphs," *arXiv e-prints*, arXiv:2302.14369, arXiv:2302.14369, 2023-2. DOI: 10.48550/arXiv.2302.14369. arXiv: 2302.14369 [quant-ph] (cit. on p. 155).

[280] J. Park, S. Jeong, M. Kim, K. Kim, *et al.*, *A rydberg-atom approach to the integer factorization problem*, 2024. arXiv: 2312.08703 [quant-ph] (cit. on p. 155).

[281] Z. Zeng, G. Giudici, and H. Pichler, *Quantum dimer models with rydberg gadgets*, 2024. arXiv: 2402.10651 [quant-ph] (cit. on p. 155).

[282] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-commerce recommendation applications," *Data Min. Knowl. Discov.*, vol. 5, no. 1–2, pp. 115–153, 2001-1, ISSN: 1384-5810. DOI: 10.1023/A:1009804230409 (cit. on p. 157).

[283] T. Pourhabibi, K.-L. Ong, B. H. Kam, and Y. L. Boo, "Fraud detection: A systematic literature review of graph-based anomaly detection approaches," *Decision Support Systems*, vol. 133, p. 113 303, 2020, ISSN: 0167-9236. DOI: https://doi.org/10.1016/j.dss.2020.113303. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167923620300580 (cit. on p. 157).

[284] G. Muzio, L. O'Bray, and K. Borgwardt, "Biological network analysis with deep learning," *Briefings in Bioinformatics*, vol. 22, no. 2, pp. 1515–1530, 2020-11, ISSN: 1477-4054. DOI: 10.1093/bib/bbaa257. eprint: https://academic.oup.com/bib/article-pdf/22/2/1515/36655309/bbaa257.pdf (cit. on p. 157).

[285] M. Schuld and N. Killoran, "Quantum machine learning in feature hilbert spaces," *Phys. Rev. Lett.*, vol. 122, arXiv:1803.07128, p. 040 504, 4 2019-2. DOI: 10.1103/PhysRevLett.122.040504. arXiv: 1803.07128 [quant-ph] (cit. on pp. 157, 164).

[286] V. Havlíček *et al.*, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, 2019-3, ISSN: 1476-4687. DOI: 10.1038/s41586-019-0980-2 (cit. on pp. 157, 164).

[287] M. Schuld, K. Brádler, R. Israel, D. Su, and B. Gupt, "Measuring the similarity of graphs with a gaussian boson sampler," *Phys. Rev. A*, vol. 101, no. 3, p. 032 314, 2020. DOI: 10.1103/PhysRevA.101.032314 (cit. on p. 157).

[288] K. Kishi, T. Satoh, R. Raymond, N. Yamamoto, and Y. Sakakibara, *Graph kernels encoding features of all subgraphs by quantum superposition*, 2021. DOI: 10.1109/JETCAS.2022.3200837. arXiv: 2103.16093 [quant-ph] (cit. on p. 157).

[289] C. Helma, R. D. King, S. Kramer, and A. Srinivasan, "The Predictive Toxicology Challenge 2000–2001," *Bioinformatics*, vol. 17, no. 1, pp. 107–108, 2001-01, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/17.1.107. eprint: https://academic.oup.com/bioinformatics/article-pdf/17/1/107/576297/170107.pdf (cit. on pp. 157, 169).

[290] H.-Y. Huang *et al.*, "Power of data in quantum machine learning," *Nat. Commun.*, vol. 12, 2631, p. 2631, 2021-1. DOI: 10.1038/s41467-021-22539-9. arXiv: 2011.01938 [quant-ph] (cit. on pp. 157, 175–177).

[291] B. Albrecht *et al.*, "Quantum feature maps for graph machine learning on a neutral atom quantum processor," *Phys. Rev. A*, vol. 107, no. 4, p. 042 615, 2023. DOI: 10.1103/PhysRevA.107.042615 (cit. on p. 158).

[292] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, no. 61, pp. 2121–2159, 2011. [Online]. Available: http://jmlr.org/papers/v12/duchi11a.html (cit. on p. 160).

[293] C. Lemaréchal, "Cauchy and the gradient method," *Doc Math Extra*, vol. 251, no. 254, p. 10, 2012. DOI: 10.4171/dms/6/27 (cit. on p. 160).

[294] P. Goyal and E. Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowledge-Based Systems*, vol. 151, pp. 78–94, 2018. DOI: 10.1016/j.knosys.2018.03.022 (cit. on p. 162).

[295] S. Bhagat, G. Cormode, and S. Muthukrishnan, "Node classification in social networks," in *Social Network Data Analytics*, C. C. Aggarwal, Ed. Boston, MA: Springer US, 2011, pp. 115–148, ISBN: 978-1-4419-8462-3. DOI: 10.1007/978-1-4419-8462-3_5 (cit. on p. 162).

[296] D. Liben-Nowell and J. Kleinberg, "The link prediction problem for social networks," in *Proceedings of the Twelfth International Conference on Information and Knowledge Management*, ser. CIKM '03, New Orleans, LA, USA: Association for Computing Machinery, 2003, pp. 556–559, ISBN: 1581137230. DOI: 10.1145/956863.956972 (cit. on p. 162).

[297] M. Schuld and F. Petruccione, *Machine learning with quantum computers*. Springer, 2021. DOI: 10.1007/978-3-030-83098-4 (cit. on p. 162).

[298] I. Glasser, N. Pancotti, and J. I. Cirac, "From probabilistic graphical models to generalized tensor networks for supervised learning," 2019. arXiv: 1806.05964 [quant-ph] (cit. on p. 162).

[299] J. M. Arrazola, A. Delgado, B. R. Bardhan, and S. Lloyd, "Quantum-inspired algorithms in practice," *arXiv preprint arXiv:1905.10415*, 2019. DOI: 10.22331/q-2020-08-13-307 (cit. on p. 162).

[300] I. Cong, S. Choi, and M. D. Lukin, "Quantum convolutional neural networks," *Nat. Phys.*, vol. 15, no. 12, pp. 1273–1278, 2019-8, ISSN: 1745-2481. DOI: 10.1038/s41567-019-0648-8. arXiv: 1810.03787 [quant-ph] (cit. on pp. 162, 164).

[301] H.-Y. Huang, R. Kueng, and J. Preskill, "Predicting many properties of a quantum system from very few measurements," *Nat. Phys.*, vol. 16, no. 10, pp. 1050–1057, 2020. DOI: 10.1038/s41567-020-0932-7 (cit. on pp. 163, 200).

[302] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, "Provably efficient machine learning for quantum many-body problems," *arXiv preprint arXiv:2106.12627*, 2021. DOI: 10.1126/science.abk3333 (cit. on p. 163).

[303] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, 1994, pp. 124–134. DOI: 10.1109/SFCS.1994.365700 (cit. on p. 163).

[304] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum algorithm for linear systems of equations," *Phys. Rev. Lett.*, vol. 103, no. 15, p. 150 502, 2009. DOI: 10.1007/978-3-642-27848-8_771-1 (cit. on p. 163).

[305] S. Aaronson, "Read the fine print," *Nat. Phys.*, vol. 11, no. 4, pp. 291–293, 2015. DOI: 10.1038/nphys3272 (cit. on p. 163).

[306] J. Zheng, Q. Gao, and Y. Lv, "Quantum Graph Convolutional Neural Networks," *arXiv e-prints*, arXiv:2107.03257, arXiv:2107.03257, 2021-7. DOI: 10.48550/arXiv.2107.03257. arXiv: 2107.03257 [eess.SP] (cit. on p. 164).

[307] G. Verdon, T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary, "Quantum Graph Neural Networks," *arXiv e-prints*, arXiv:1909.12264, arXiv:1909.12264, 2019-9. DOI: 10.48550/arXiv.1909.12264. arXiv: 1909.12264 [quant-ph] (cit. on pp. 164, 168).

[308] E. H. Lieb and D. W. Robinson, "The finite group velocity of quantum spin systems," *Communications in Mathematical Physics*, vol. 28, no. 3, pp. 251–257, 1972-9, ISSN: 1432-0916. DOI: 10.1007/bf01645779 (cit. on p. 165).

[309] M. C. Tran *et al.*, "Locality and digital quantum simulation of power-law interactions," *Phys. Rev. X*, vol. 9, p. 031 006, 3 2019-7. DOI: 10.1103/PhysRevX.9.031006 (cit. on p. 165).

[310] B. Weisfeiler and A. Leman, "The reduction of a graph to canonical form and the algebra which appears therein," *nti, Series*, vol. 2, no. 9, pp. 12–16, 1968. [Online]. Available: https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf (cit. on p. 165).

[311] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," 2021. arXiv: 2104.13478 [cs.LG] (cit. on p. 166).

[312] S. Maskey, R. Levie, Y. Lee, and G. Kutyniok, "Generalization analysis of message passing neural networks on large random graphs," 2022. arXiv: 2202.00645 [cs.LG] (cit. on p. 166).

[313] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" In *International Conference on Learning Representations*, 2019. [Online]. Available: https://openreview.net/forum?id=ryGs6iA5Km (cit. on pp. 166, 178).

[314] J. SHAO and D. TU, *The Jackknife and Bootstrap*. SPRINGER, 1995. DOI: 10.1007/978-1-4612-0795-5 (cit. on p. 166).

[315] L. Bai and E. R. Hancock, "Graph kernels from the jensen-shannon divergence," *Journal of mathematical imaging and vision*, vol. 47, no. 1, pp. 60–69, 2013. DOI: 10.1007/s10851-012-0383-6 (cit. on p. 167).

[316] P. Mernyei, K. Meichanetzidis, and İ. İ. Ceylan, *Equivariant quantum graph circuits*, 2022. arXiv: 2112.05261 [cs.LG]. [Online]. Available: https://arxiv.org/pdf/2112.05261.pdf (cit. on p. 168).

[317] S. Thabet, R. Fouilland, and L. Henriet, *Extending graph transformers with quantum computed aggregation*, 2022-10. arXiv: 2210.10610 [quant-ph] (cit. on pp. 168, 178, 199).

[318] L.-P. Henry, S. Thabet, C. Dalyac, and L. Henriet, "Quantum evolution kernel: Machine learning on graphs with programmable arrays of qubits," *Phys. Rev. A*, vol. 104, p. 032 416, 3 2021-9. DOI: 10.1103/PhysRevA.104.032416 (cit. on pp. 168, 178).

[319] G. Siglidis, G. Nikolentzos, S. Limnios, C. Giatsidis, K. Skianis, and M. Vazirgiannis, *Grakel: A graph kernel library in python*, 2018. DOI: 10.48550/ARXIV.1806.02193 (cit. on p. 169).

[320] L. Boon, E. Ugarte-Berzal, J. Vandooren, and G. Opdenakker, "Protease propeptide structures, mechanisms of activation, and functions," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 55, no. 2, pp. 111–165, 2020, PMID: 32290726. DOI: 10.1080/10409238.2020.1742090. eprint: https://doi.org/10.1080/10409238.2020.1742090 (cit. on p. 169).

[321] R. Shapiro and B. L. Vallee, "Interaction of human placental ribonuclease with placental ribonuclease inhibitor," *Biochemistry*, vol. 30, no. 8, pp. 2246–2255, 1991-2, ISSN: 1520-4995. DOI: 10.1021/bi00222a030 (cit. on p. 169).

[322] N. M. Kriege, F. D. Johansson, and C. Morris, "A survey on graph kernels," *Applied Network Science*, vol. 5, no. 6, 2020. DOI: 10.1007/s41109-019-0195-3 (cit. on p. 169).

[323] C. Morris *et al.*, "Weisfeiler and leman go neural: Higher-order graph neural networks," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 4602–4609, 2019-7. DOI: 10.1609/aaai.v33i01.33014602. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/4384 (cit. on pp. 178, 200).

[324] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural Message Passing for Quantum Chemistry," *arXiv e-prints*, Proceedings of Machine Learning Research, vol. 70, arXiv:1704.01212, D. Precup and Y. W. Teh, Eds., arXiv:1704.01212, 2017-4. DOI: 10.48550/arXiv.1704.01212. arXiv: 1704.01212 [cs.LG]. [Online]. Available: https://proceedings.mlr.press/v70/gilmer17a.html (cit. on p. 179).

[325] C. Ying *et al.*, *Do transformers really perform bad for graph representation?* 2021. arXiv: 2106.05234 [cs.LG] (cit. on p. 179).

[326] L. Rampášek, M. Galkin, V. P. Dwivedi, A. T. Luu, G. Wolf, and D. Beaini, *Recipe for a general, powerful, scalable graph transformer*, 2022. DOI: 10.48550/ARXIV.2205.12454 (cit. on p. 179).

[327] D. Kreuzer, D. Beaini, W. Hamilton, V. Létourneau, and P. Tossou, "Rethinking graph transformers with spectral attention," *Advances in Neural Information Processing Systems*, vol. 34, pp. 21 618–21 629, 2021. arXiv: 2106.03893 (cit. on p. 179).

[328] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, "Group-invariant quantum machine learning," *PRX Quantum*, vol. 3, p. 030 341, 3 2022-9. DOI: 10.1103/PRXQuantum.3.030341 (cit. on p. 179).

[329] A. Skolik, M. Cattelan, S. Yarkoni, T. Bäck, and V. Dunjko, "Equivariant quantum circuits for learning on weighted graphs," *npj Quantum Information*, vol. 9, no. 1, p. 47, 2023. DOI: 10.48550/ARXIV.2205.06109 (cit. on p. 179).

[330] M. Ragone *et al.*, "Representation theory for geometric quantum machine learning," *arXiv preprint arXiv:2210.07980*, 2022. DOI: 10.48550/ARXIV.2210.07980 (cit. on p. 179).

[331] Q. T. Nguyen *et al.*, "Theory for equivariant quantum neural networks," *arXiv preprint arXiv:2210.08566*, 2022. DOI: 10.48550/ARXIV.2210.08566. arXiv: 2210.08566 [quant-ph] (cit. on p. 179).

[332] L. Schatzki, M. Larocca, Q. T. Nguyen, F. Sauvage, and M. Cerezo, "Theoretical guarantees for permutation-equivariant quantum neural networks," *npj Quantum Information*, vol. 10, no. 1, 2024-1, ISSN: 2056-6387. DOI: 10.1038/s41534-024-00804-1 (cit. on p. 179).

[333] R. Orús, S. Mugel, and E. Lizaso, "Quantum Computing for Finance: Overview and Prospects," *Reviews in Physics*, 2019-11. DOI: 10.1016/j.revip.2019.100028 (cit. on p. 181).

[334] R. Patel *et al.*, "Quantum-inspired tensor neural networks for partial differential equations," *ArXiv:2208.02235*, 2022. [Online]. Available: https://arxiv.org/abs/2208.02235 (cit. on p. 181).

[335] S. Mugel *et al.*, "Dynamic portfolio optimization with real datasets using quantum processors and quantum-inspired tensor networks," *Phys. Rev. R*, vol. 4, no. 1, p. 013 006, 2022. DOI: 10.1103/PhysRevResearch.4.013006 (cit. on p. 181).

[336] P. Bermejo and R. Orús, "Variational quantum and quantum-inspired clustering," *Scientific Reports*, vol. 13, no. 1, 2023-8, ISSN: 2045-2322. DOI: 10.1038/s41598-023-39771-6 (cit. on p. 181).

[337] A. Borle, V. E. Elfving, and S. J. Lomonaco, "Quantum approximate optimization for hard problems in linear algebra," *ArXiv:2006.15438*, 2020. DOI: 10.48550/ARXIV.2006.15438 (cit. on p. 181).

[338] O. Kyriienko, A. E. Paine, and V. E. Elfving, "Solving nonlinear differential equations with differentiable quantum circuits," *Phys. Rev. A*, vol. 103, p. 052 416, 5 2021-5. DOI: 10.1103/PhysRevA.103.052416 (cit. on p. 181).

[339] A. E. Paine, V. E. Elfving, and O. Kyriienko, "Quantum quantile mechanics: Solving stochastic differential equations for generating time-series," *ArXiv:2108.03190*, 2021. DOI: 10.48550/ARXIV.2108.03190 (cit. on p. 181).

[340] O. Kyriienko, A. E. Paine, and V. E. Elfving, "Protocols for trainable and differentiable quantum generative modelling," *ArXiv:2202.08253*, 2022. DOI: 10.48550/ARXIV.2202.08253 (cit. on p. 181).

[341] O. Kyriienko and E. B. Magnusson, "Unsupervised quantum machine learning for fraud detection," *ArXiv:2208.01203*, 2022. DOI: 10.48550/ARXIV.2208.01203 (cit. on p. 181).

[342] L. Leclerc *et al.*, "Financial risk management on a neutral atom quantum processor," *Phys. Rev. Res.*, vol. 5, p. 043 117, 4 2023-11. DOI: 10.1103/PhysRevResearch.5.043117 (cit. on p. 182).

[343] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *European Journal of Operational Research*, vol. 247, no. 1, pp. 124–136, 2015. DOI: 10.1016/j.ejor.2015.05.030. [Online]. Available: https://EconPapers.repec.org/RePEc:eee:ejores:v:247:y:2015:i:1:p:124-136 (cit. on p. 182).

[344] X. Xie, S. Pang, and J. Chen, "Hybrid recommendation model based on deep learning and stacking integration strategy," *Intell. Data Anal.*, vol. 24, no. 6, pp. 1329–1344, 2020-1, ISSN: 1088-467X. DOI: 10.3233/IDA-194961 (cit. on p. 182).

[345] S. Chen, Z. Guo, and X. Zhao, "Predicting mortgage early delinquency with machine learning methods," *European Journal of Operational Research*, vol. 290, no. 1, pp. 358–372, 2021. DOI: 10.1016/j.ejor.2020.07.05 (cit. on p. 182).

[346] E. Kambal, I. M. Osman, M. Taha, N. Mohammed, and S. Mohammed, "Credit scoring using data mining techniques with particular reference to sudanese banks," 2013, pp. 378–383. [Online]. Available: https://api.semanticscholar.org/CorpusID:9678462 (cit. on p. 182).

[347] B. R. Gunnarsson, S. Vanden Broucke, B. Baesens, M. Óskarsdóttir, and W. Lemahieu, "Deep learning for credit scoring: Do or don't?" *European Journal of Operational Research*, vol. 295, no. 1, pp. 292–305, 2021. DOI: 10.1016/j.ejor.2021.03.006 (cit. on p. 182).

[348] X. Dastile and T. Celik, "Making deep learning-based predictions for credit scoring explainable," *IEEE Access*, vol. 9, pp. 50 426–50 440, 2021. DOI: 10.1109/ACCESS.2021.3068854 (cit. on p. 182).

[349] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready, "Qboost: Large scale classifier training with adiabatic quantum optimization," Proceedings of Machine Learning Research, vol. 25, S. C. H. Hoi and W. Buntine, Eds., pp. 333–348, 2012-11. [Online]. Available: https://proceedings.mlr.press/v25/neven12.html (cit. on pp. 184, 186).

[350] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997. DOI: 10.1006/jcss.1997.1504 (cit. on p. 184).

[351] R. Wang, "Adaboost for feature selection, classification and its relation with svm, a review," *Physics Procedia*, vol. 25, pp. 800–807, 2012. DOI: 10.1016/j.phpro.2012.03.160 (cit. on p. 184).

[352] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready, "Training a large scale classifier with the quantum adiabatic algorithm," *ArXiv:0912.0779*, 2009. [Online]. Available: http://arxiv.org/abs/0912.0779 (cit. on p. 185).

[353] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren, "The online performance estimation framework: Heterogeneous ensemble learning for data streams," *Machine Learning*, vol. 107, no. 1, pp. 149–176, 2018. DOI: 10.1007/s10994-017-5686-9 (cit. on p. 185).

[354] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006, vol. 4. [Online]. Available: https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf (cit. on p. 185).

[355] F. Glover, G. Kochenberger, and Y. Du, "A tutorial on formulating and using qubo models," *ArXiv:1811.11538*, 2018. DOI: 10.48550/ARXIV.1811.11538 (cit. on p. 187).

[356] E. Crosson and A. W. Harrow, "Simulated quantum annealing can be exponentially faster than classical simulated annealing," in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, 2016, pp. 714–723. DOI: 10.1109/FOCS.2016.81 (cit. on p. 187).

[357] S. Zbinden, A. Bärtschi, H. Djidjev, and S. Eidenbenz, "Embedding algorithms for quantum annealers with chimera and pegasus connection topologies," in *Lecture Notes in Computer Science*, ser. Lecture notes in computer science, Cham: Springer International Publishing, 2020, pp. 187–206. DOI: 10.1007/978-3-030-50743-5_10 (cit. on p. 187).

[358] M. Cerezo *et al.*, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, no. 9, pp. 625–44, 2021. DOI: 10.1038/s42254-021-00348-9 (cit. on pp. 188, 198).

[359] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, and E. C. Sewell, "Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning," *Discrete Optimization*, vol. 19, pp. 79–102, 2016. DOI: 10.1016/j.disopt.2016.01.005 (cit. on p. 198).

[360] F. V. Fomin, D. Lokshtanov, S. Saurabh, and M. Zehavi, *Kernelization: theory of parameterized preprocessing*. Cambridge University Press, 2019-12, ISBN: 9781107057760. DOI: 10.1017/9781107415157 (cit. on p. 198).

[361] J. Wurtz, S. Sack, and S.-T. Wang, *Solving non-native combinatorial optimization problems using hybrid quantum-classical algorithms*, 2024. arXiv: 2403.03153 [quant-ph] (cit. on p. 198).

[362] M. Schuld, I. Sinayskiy, and F. Petruccione, "The quest for a quantum neural network," *Quantum Information Processing*, vol. 13, pp. 2567–2586, 2014. DOI: 10.1007/s11128-014-0809-8 (cit. on p. 198).

[363] J. Desrosiers, F. Soumis, and M. Desrochers, "Routing with time windows by column generation," *Networks*, vol. 14, no. 4, pp. 545–565, 1984. DOI: 10.1002/net.3230140406 (cit. on p. 199).

[364] W. da Silva Coelho, L. Henriet, and L.-P. Henry, "Quantum pricing-based column-generation framework for hard combinatorial problems," *Phys. Rev. A*, vol. 107, no. 3, p. 032 426, 2023. DOI: 10.1103/PhysRevA.107.032426 (cit. on p. 199).

[365] E. Malaguti, M. Monaci, and P. Toth, "An exact approach for the vertex coloring problem," *Discrete Optimization*, vol. 8, no. 2, pp. 174–190, 2011. DOI: 10.1016/j.disopt.2010.07.005 (cit. on p. 199).

[366] M. Y. Naghmouchi and W. da Silva Coelho, *Mixed integer linear programming solver using benders decomposition assisted by neutral atom quantum processor*, 2024. arXiv: 2402.05748 [quant-ph] (cit. on p. 199).

[367] Y. Tang and J. Yan, "Graphqntk: Quantum neural tangent kernel for graph data," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., vol. 35, Curran Associates, Inc., 2022, pp. 6104–6118. [Online]. Available: https://proceedings.neurips.cc/paper%5C%5Ffiles/paper/2022/file/285b06e0dd856f20591b0a5beb954151-Paper-Conference.pdf (cit. on p. 199).

[368] S. Y.-C. Chen, T.-C. Wei, C. Zhang, H. Yu, and S. Yoo, "Hybrid quantum-classical graph convolutional network," 2021. arXiv: 2101.06189 [cs.LG] (cit. on p. 199).

[369] J. K. Gamble, M. Friesen, D. Zhou, R. Joynt, and S. N. Coppersmith, "Two-particle quantum walks applied to the graph isomorphism problem," *Phys. Rev. A*, vol. 81, no. 5, 2010-5. DOI: 10.1103/physreva.81.052313 (cit. on p. 199).

[370] S. Thabet *et al.*, *Enhancing graph neural networks with quantum computed encodings*, 2023. arXiv: 2310.20519 [quant-ph] (cit. on p. 199).

[371] K. Kitai *et al.*, "Designing metamaterials with quantum annealing and factorization machines," *Phys. Rev. Res.*, vol. 2, p. 013 319, 1 2020-3. DOI: 10.1103/PhysRevResearch.2.013319 (cit. on p. 200).

[372] T. Matsumori, M. Taki, and T. Kadowaki, "Application of qubo solver using black-box optimization to structural design for resonance avoidance," *Scientific Reports*, vol. 12, no. 1, p. 12 143, 2022-7, ISSN: 2045-2322. DOI: 10.1038/s41598-022-16149-8 (cit. on p. 200).

[373] S. Varsamopoulos *et al.*, *Quantum extremal learning*, 2022. arXiv: 2205.02807 [quant-ph] (cit. on p. 200).

[374] H. J. Manetsch, G. Nomura, E. Bataille, K. H. Leung, X. Lv, and M. Endres, *A tweezer array with 6100 highly coherent atomic qubits*, 2024. arXiv: 2403.12021 [quant-ph] (cit. on p. 203).

[375] T. Ayral, T. Louvet, Y. Zhou, C. Lambert, E. M. Stoudenmire, and X. Waintal, "Density-matrix renormalization group algorithm for simulating quantum circuits with a finite fidelity," *PRX Quantum*, vol. 4, no. 2, 2023-4, ISSN: 2691-3399. DOI: 10.1103/prxquantum.4.020304 (cit. on p. 203).

[376] C. Li, J. Shen, R. Shaydulin, and M. Pistoia, *Quantum counterdiabatic driving with local control*, 2024. arXiv: 2403.01854 [quant-ph] (cit. on p. 203).

[377] J. R. Finžgar, M. J. A. Schuetz, J. K. Brubaker, H. Nishimori, and H. G. Katzgraber, "Designing quantum annealing schedules using bayesian optimization," *Physical Review Research*, vol. 6, no. 2, 2024-4, ISSN: 2643-1564. DOI: 10.1103/physrevresearch.6.023063 (cit. on p. 203).

[378] F. Sauvage and F. Mintert, "Optimal control of families of quantum gates," *Phys. Rev. Lett.*, vol. 129, no. 5, 2022-7, ISSN: 1079-7114. DOI: 10.1103/physrevlett.129.050507 (cit. on p. 203).

[379] J. P. Covey, H. Weinfurter, and H. Bernien, "Quantum networks with neutral atom processing nodes," *npj Quantum Information*, vol. 9, no. 1, 2023-9, ISSN: 2056-6387. DOI: 10.1038/s41534-023-00759-9 (cit. on p. 203).

[380] M. Cao, F. Hoffet, S. Qiu, A. S. Sheremet, and J. Laurat, "Efficient reversible entanglement transfer between light and quantum memories," *Optica*, vol. 7, no. 10, p. 1440, 2020-10, ISSN: 2334-2536. DOI: 10.1364/optica.400695 (cit. on p. 203).

[381] P. W. Shor, *Fault-tolerant quantum computation*, 1997. arXiv: quant-ph/9605011 [quant-ph] (cit. on p. 203).

[382] L. Lovasz, "On the shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979. DOI: 10.1109/TIT.1979.1055985 (cit. on p. 205).

[383] C. A. Micchelli, Y. Xu, and H. Zhang, "Universal kernels," *Journal of Machine Learning Research*, vol. 7, no. 95, pp. 2651–2667, 2006. [Online]. Available: http://jmlr.org/papers/v7/micchelli06a.html (cit. on p. 206).

[384] N. Shervashidze, S. Vishwanathan, T. Petri, K. Mehlhorn, and K. Borgwardt, "Efficient graphlet kernels for large graph comparison," in *JMLR Workshop and Conference Proceedings Volume 5: AISTATS 2009*, 2009-4, pp. 488–495. [Online]. Available: https://proceedings.mlr.press/v5/shervashidze09a/shervashidze09a.pdf (cit. on p. 206).

[385] T. Gärtner, P. Flach, and S. Wrobel, "On graph kernels: Hardness results and efficient alternatives," in *Learning Theory and Kernel Machines*, B. Schölkopf and M. K. Warmuth, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 129–143, ISBN: 978-3-540-45167-9. DOI: 10.1007/978-3-540-45167-9_11 (cit. on p. 207).

[386] S. Vishwanathan, N. N. Schraudolph, R. Kondor, and K. M. Borgwardt, "Graph kernels," *Journal of Machine Learning Research*, vol. 11, no. 40, pp. 1201–1242, 2010. [Online]. Available: http://jmlr.org/papers/v11/vishwanathan10a.html (cit. on p. 207).

[387] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959-12, ISSN: 0945-3245. DOI: 10.1007/BF01386390 (cit. on p. 208).

[388] R. W. Floyd, "Algorithm 97: Shortest path," *Commun. ACM*, vol. 5, no. 6, p. 345, 1962-6, ISSN: 0001-0782. DOI: 10.1145/367766.368168 (cit. on p. 208).

université PARIS-SACLAY | ÉCOLE DOCTORALE Ondes et matière (EDOM)

Titre: Calcul quantique avec des atomes de Rydberg: contrôle et modélisation pour simulation et algorithmes quantiques

Mots clés: Ordinateur quantique, Atome de Rydberg, Contrôle optimal, Jumeau numérique.

Résumé: Améliorer sa compréhension d'un système en le modélisant permet d'espérer le contrôler de manière plus optimale et ouvre la voie à une myriade d'applications potentielles, exploitant les effets jusqu'alors énigmatiques de ce système désormais familier. Cette thèse applique ce paradigme au calcul quantique analogique avec des atomes de Rydberg, montrant comment à l'aide d'une modélisation minutieuse du bruit, de protocoles de contrôle optimaux et de techniques d'apprentissage automatique, on peut espérer améliorer des expériences de simulation de magnétisme quantique ou la résolution de problèmes d'optimisation et de classification de graphes. Après avoir décrit la plateforme expérimentale permettant de contrôler les atomes de Rydberg, nous introduisons des outils classiques tels que les jumeaux numériques de systèmes enclins à des erreurs, la modélisation d'un grand nombre d'atomes par réseaux de tenseurs, le contrôle optimal robuste et l'optimisation bayésienne pour les algorithmes variationnels. Nous appliquons ces outils à plusieurs applications prometteuses. Nous améliorons la préparation d'états antiferromagnétiques dans le modèle d'Ising et réalisons une évaluation détaillée de l'influence d'erreurs sur l'étude de phases magnétiques du modèle dipolaire XY et lors de la tomographie d'états quantiques. En utilisant des techniques d'optimisation et des méthodes d'apprentissage automatique, nous abordons également des cas d'usage industriels tels que la résolution du problème de stable maximum sur des graphes représentant des tâches de planification de charge de batteries de voitures électriques, la classification de composés moléculaires toxiques ou inoffensifs, et des tâches de prédiction dans la gestion des risques financiers.

Title: Quantum computing with Rydberg atoms: control and modelling for quantum simulation and practical algorithms

Keywords: Quantum computer, Rydberg atom, Optimal control, Digital twin.

Abstract: Refining our understanding of an unknown system through modelling lays the groundwork for optimally controlling it and opens the door to a myriad of potential applications, exploiting the once enigmatic and unpredictable effects of this now-known system. This thesis applies this paradigm to analog quantum computing with Rydberg atoms, showcasing how careful noise modelling, optimal control and machine learning frameworks can support and enhance the simulation of quantum magnetism and the solving of graph-based optimisation and classification problems. After describing the experimental platform enabling the control of Rydberg atoms, we introduce classical tools such as digital twins of noisy systems, tensor network modelling, robust optimal control, and Bayesian optimisation for variational algorithms. We apply the latter to several applications. We improve the preparation of antiferromagnetic state in the Ising model and benchmark the noisy behaviour of a dipolar XY quantum simulator when probing continuous symmetry breaking and performing quantum state tomography. Using optimisation techniques and machine learning methods, we also tackle industrial use cases such as maximum independent set on graphs representing smart charging tasks, binary classification of toxic or harmless molecular compounds, and prediction of fallen angel companies in financial risk management.