Méthodes numériques pour la simulation d'équations aux dérivées partielles stochastiques non-linéaires en condensation de Bose-Einstein

Romain Poncet

NNT : 2017SACLX069

Thèse de doctorat de l'Université Paris-Saclay préparée à l'École Polytechnique

Ecole doctorale n°574
Ecole Doctorale de Mathématique Hadamard
Spécialité de doctorat : Mathématiques Appliquées

par

M. Romain Poncet

Méthodes numériques pour la simulation d'équations aux dérivées partielles stochastiques non-linéaires en condensation de Bose-Einstein

Thèse présentée et soutenue à École Polytechnique, le 02 octobre 2017.

Composition du Jury :

M.	Norbert J. Mauser	Professeur Universität Wien	(Rapporteur)
M.	Gabriel Stoltz	Professeur École Nationale des Ponts et Ch.	(Rapporteur)
M.	Christophe Besse	Professeur Université Paul Sabatier Toulouse	(Président du jury)
M.	Charles-Edouard Brehier	Chargé de recherche au CNRS Université Lyon 1	(Examinateur)
M.	Emmanuel Gobet	Professeur École Polytechnique	(Examinateur)
M.	Ludovic Goudenège	Chargé de recherche au CNRS CentraleSupélec	(Examinateur)
M.	Tony Lelièvre	Professeur École Nationale des Ponts et Ch.	(Examinateur)
Mme	Reika Fukuizumi	Directrice de recherche Université de Tohoku	(Membre invité)
Mme	Anne de Bouard	Directrice de recherche au CNRS École Polytechnique	(Directrice de thèse)

Table des matières

Table des figures

Chapitre 1

Introduction

1.1 La condensation de Bose-Einstein

La condensation de Bose-Einstein désigne un nouvel état de la matière qui peut être atteint sous des conditions de température extrêmement basse, et de faible densité. Le centre d'un condensat est typiquement 10^5 fois moins dense que l'air qui nous entoure. Il s'agit d'un état de la matière constitué de bosons (que sont les particules élémentaires, ou composites, de spin entier) occupant un unique état quantique fondamental. Cet état de la matière est intrinsèquement lié au caractère indiscernable des bosons. Si l'on cherche à comprendre comment un ensemble de bosons à énergie fixée se répartit sur tous les niveaux d'énergie (cinétique), alors une bonne manière de procéder consiste à considérer parmi toutes les répartitions possibles, celle(s) de plus grande probabilité. Le résultat de ce calcul est évidemment très dépendant du fait que les particules sont indiscernables. Il décrit le nombre de bosons qu'un système réparti dans chaque état d'énergie, et conduit à la *statistique de Bose-Einstein*. Ce calcul est conduit dans le Chapitre 2 au Paragraphe 2.1.2. Le phénomène de condensation apparaît dès lors que le nombre de particules qui peuvent être accommodées dans tous les états, à l'exception de l'état fondamental, est borné. Cela est le cas quand les niveaux d'énergie croissent suffisamment vite. Ainsi tout système dont le nombre de particules est très supérieur à ce nombre maximal, possédera une accumulation de particules dans l'état fondamental (car ce dernier, et contrairement aux autres états, peut accueillir une infinité de particules). C'est cet effet d'accumulation que l'on désigne par *condensation de Bose-Einstein*. Celui-ci est décrit plus précisément au Paragraphe 2.1.3, où des références sont proposées. Si en pratique de tels système ne peuvent exister qu'à des températures de l'ordre d'une fraction de microKelvins, c'est parce que cette borne sur le nombre de particules est une fonction croissante de la température. Cet état de la matière est en fait semblable à un état gazeux, mais avec cette propriété quantique très particulière. Les techniques de fabrication de tels systèmes sont aujourd'hui très bien connues, mais il aura fallu attendre 1995 pour la première mise en évidence expérimentale du phénomène de condensation de Bose-Einstein. Ces techniques consistent à confiner des bosons dans des pièges optiques ou magnétiques, et à refroidir le système

à l'aide d'une succession de manipulations. Celles-ci sont présentées au Paragraphe 2.3. En pratique, nous sommes capables de construire des condensats constitués d'environ un million d'atomes. Ces systèmes forment donc des objets quantiques macroscopiques, de l'ordre du micromètre. Bien que très petits, il est toutefois possible de les observer facilement grâce à des méthodes (destructrices) d'expansion spatiale.

Une particularité de ces systèmes, à laquelle nous nous sommes intéressés est la réponse du condensat à une mise en rotation. En effet, celle-ci révèle bien leur caractère quantique car la mise en rotation s'accompagne de l'apparition de vortex dont le nombre croît principalement avec la vitesse de rotation du condensat. Ces vortex constituent des points de singularité du champ de phase de la fonction d'onde, mais pas de la fonction d'onde. Ce sont donc des points d'annulation de la fonction d'onde. Le carré du module de cette dernière correspondant à une densité d'atomes dans le condensat, les vortex apparaissent alors sous la forme de trous lors de l'observation directe.

Le Chapitre 2 est consacré à une présentation plus détaillée du phénomène de condensation de Bose-Einstein, ainsi qu'à une description des procédés expérimentaux permettant leur construction. Nous présentons également dans ce chapitre une modélisation mathématique déterministe d'un condensat de Bose-Einstein à température nulle. Celle-ci porte le nom d'équation de Gross-Pitaevskii, et prend la forme d'une équation aux dérivées partielles de type Schrödinger non-linéaire, que nous présentons également brièvement dans le paragraphe suivant.

1.1.1 L'équation de Gross-Pitaevskii

Une exposition détaillée de ce modèle est fournie dans le Chapitre 2, au Paragraphe 2.2.1. Cette équation décrit l'évolution temporelle de la fonction d'onde ϕ d'un condensat de Bose-Einstein, à température nulle, et sous l'hypothèse qu'il est peuplé d'un grand nombre d'atomes. Nous notons d la dimension de l'espace ambiant. On supposera classiquement $d = 3$, ou $d = 2$ dans le cas où l'on procède à une réduction de la dimension (comme montré dans le Paragraphe 2.4.2). C'est une équation de type Schrödinger non-linéaire donnée par,

$$i\hbar\frac{\partial}{\partial t}\phi(t,\mathbf{x}) = \left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}) + \frac{4\pi\hbar^2 a}{m}\left|\phi(t,\mathbf{x})\right|^2\right)\phi(t,\mathbf{x}), \tag{1.1}$$

où $\mathbf{x}$ est un élément de $\mathbb{R}^d$, et $t \geq 0$. Dans cette équation, le terme V correspond au potentiel de confinement des atomes, a désigne la longueur de diffusion de l'espèce chimique modélisée. Une définition de cette grandeur, ainsi qu'une justification formelle de son utilisation, sont données au Paragraphe 2.2.2. La constante m désigne la masse de l'atome en question, et $\hbar$ désigne classiquement la constante de Planck réduite.

L'énergie H du système, appelée *énergie de Gross-Pitaevskii*, est donnée par,

$$H(\phi) = \frac{1}{2}\int_{\mathbb{R}^d}\left(\frac{\hbar^2}{2m}\left|\nabla\phi(\mathbf{x})\right|^2 + V(\mathbf{x})\left|\phi(\mathbf{x})\right|^2 + \frac{4\pi\hbar^2 a}{2m}\left|\phi(\mathbf{x})\right|^4\right)d\mathbf{x}. \tag{1.2}$$

Celle-ci, tout comme la norme $L^2(\mathbb{R}^d)$ de ϕ, est conservée pour l'Équation (1.1).

1.1.2 Deux généralisations de l'équation de Gross-Pitaevskii

Nous nous sommes intéressés dans cette thèse à deux généralisations stochastiques de l'équation de Gross-Pitaevskii. La première constitue une modélisation des fluctuations temporelles du potentiel de confinement. Elle prend la forme d'une équation aux dérivées partielles stochastique où le bruit multiplicatif est unidimensionnel que nous présentons au Paragraphe 1.2. Nous nous sommes intéressés d'une part à l'élaboration d'une méthode numérique pour l'approximation, à horizon temporel fini, de la solution de cette équation. Celle-ci fait l'objet du Chapitre 3. D'autre part, nous nous sommes intéressés à l'analyse à la fois théorique et numérique de la dynamique aléatoire d'une solution stationnaire (dans le cas déterministe) de type vortex. Cette étude fait l'objet du Chapitre 4. La seconde généralisation stochastique de l'équation de Gross-Pitaevskii modélise les effets de la température sur la dynamique du condensat. Elle est présentée dans ce chapitre au Paragraphe 1.3. Il est en particulier connu que cette dynamique possède une unique mesure invariante, et que la loi de sa solution converge exponentiellement vite vers cette mesure invariante. Nous nous sommes intéressés dans le Chapitre 5 à l'élaboration d'une méthode de type *Markov Chain Monte Carlo* (MCMC) permettant d'échantillonner efficacement cette loi invariante. Elle peut être vue comme une méthode de réduction de variance pour l'algorithme *Metropolis-Adjusted Langevin Algorithm* (MALA) [152]. Cette méthode n'étant pas limitée à notre cas d'application, ce chapitre est présenté dans un cadre général. Finalement, la dynamique suivie par cette généralisation stochastique de l'équation de Gross-Pitaevskii peut exhiber, dans certains cas, un comportement métastable. Nous étudions numériquement dans le Chapitre 6 des méthodes de simulation d'événements métastables pour cette dynamique.

Bien que ce chapitre ne nécessite pas la lecture préliminaire du Chapitre 2 pour comprendre les modèles présentés ci-dessous, celle-ci peut néanmoins permettre de comprendre plus en détail le contexte physique sous-jacent.

1.2 Une modélisation des défauts du confinement optique

La construction d'un condensat de Bose-Einstein nécessite de refroidir le gaz de bosons à des températures de l'ordre de la fraction de microKelvins. Celui-ci doit en plus être confiné pour éviter qu'il n'interagisse avec les parois de la chambre à vide, et afin de contrôler sa densité. Plusieurs techniques, reposant sur des méthodes par confinement optique ou magnétique peuvent être envisagées (voir Paragraphe 2.3). Le confinement optique possède de nombreux avantages par rapport au confinement magnétique. En particulier, il n'est en théorie pas sujet aux pertes de particules par transitions Majorana, et est donc censé permettre de préserver des condensats pendant de longues périodes. Il a cependant été constaté à la fin des années 1990 que ceux-ci ne permettaient pas de conserver ces systèmes au delà de la dizaine de secondes [73]. De plus, une augmentation inexpliquée

de la température était observée. Une des hypothèses avancées était que ces phénomènes pouvaient être provoqués par des fluctuations dans l'intensité des lasers utilisés pour le confinement optique. La vérification de cette hypothèse a justifié les travaux [85, 155]. Une modélisation aléatoire des fluctuations a également été étudiée dans [82] où les auteurs étudient leurs effets sur la durée de vie d'un condensat piégé. Dans [1] les auteurs étudient les propriétés qualitatives d'une modélisation aléatoire de ces fluctuations dans le modèle de Gross-Pitaevskii en deux dimensions et dans le cas d'une symétrie radiale à l'aide d'une *méthode des moments*. Celle-ci consiste à obtenir un système fermé d'équations différentielles ordinaires sur certaines intégrales de la solution.

Comme décrit au Paragraphe 2.3.3, le potentiel de confinement est négativement proportionnel au carré de l'amplitude du champ électrique $E(t, \mathbf{x})$. En notant $-\alpha \leq 0$ ce coefficient de proportionnalité, le potentiel de confinement est alors donné par $V(t, \mathbf{x}) = -\alpha E(t, \mathbf{x})^2$. Dans le cas d'un laser gaussien ce champ est donné par,

$$|E(t, \mathbf{x})|^2 = E_0(t)^2 \mathrm{e}^{-\frac{|\mathbf{x}|^2}{l^2}},$$

où l désigne le rayon caractéristique du faisceau. Pour un condensat suffisamment petit, ce champ peut être approximé à l'ordre 1 au voisinage du centre du faisceau par,

$$|E(t, \mathbf{x})|^2 \approx E_0(t)^2 \left(1 - \frac{|\mathbf{x}|^2}{l^2}\right),$$

et donc un potentiel de confinement,

$$V(t, \mathbf{x}) = -\alpha E_0(t)^2 + \frac{\alpha}{l^2} E_0(t)^2 |\mathbf{x}|^2 .$$

La composante uniforme de ce potentiel ne participe qu'à un déphasage uniforme de la solution de l'équation de Gross-Pitaevskii, et peut donc être retirée à l'aide d'un changement de jauge. On obtient alors l'équation de Gross-Pitaevskii suivante :

$$i\hbar \frac{\partial}{\partial t}\phi(t, \mathbf{r}) = \left(-\frac{\hbar^2}{2m}\Delta + \frac{\alpha}{l^2} E_0(t)^2 |\mathbf{x}|^2\right) \phi(t, \mathbf{r}) + \frac{4\pi\hbar^2 a}{m} |\phi(t, \mathbf{r})|^2 \phi(t, \mathbf{r}). \tag{1.3}$$

On introduit le processus $(\dot{\xi}(t))_{t\geq 0}$ qui correspond aux fluctuations relatives de l'intensité des lasers autour de leur valeur moyenne E_0. Ce processus est défini par,

$$\dot{\xi}(t) = \frac{|E_0(t)|^2 - |E_0|^2}{|E_0|^2}.$$

L'Équation (1.3) se réécrit alors

$$i\hbar \frac{\partial}{\partial t}\phi(t, \mathbf{r}) = \left(-\frac{\hbar^2}{2m}\Delta + \frac{\alpha}{l^2} E_0^2(1 + \dot{\xi}(t)) |\mathbf{x}|^2\right) \phi(t, \mathbf{r}) + \frac{4\pi\hbar^2 a}{m} |\phi(t, \mathbf{r})|^2 \phi(t, \mathbf{r}). \tag{1.4}$$

Les travaux [54, 56] ont permis d'offrir un cadre mathématique rigoureux pour l'étude de cette équation. Dans ces articles, le processus $(\dot{\xi}(t))_{t\geq 0}$ est défini par un bruit blanc en temps de fonction de corrélation $\mathbb{E}\left[\dot{\xi}(t)\dot{\xi}(s)\right] = \sigma_0^2\delta_0(t-s)$, où δ_0 correspond à une mesure de Dirac centrée en 0. Le produit $\dot{\xi}(t)\left|\mathbf{x}\right|^2\phi(t)$ apparaissant dans (1.4) est modélisé dans un cadre stochastique par un produit de Stratonovich. Cette modélisation est motivée par le fait que ce bruit peut être interprété comme une limite de processus de longueur de corrélation non nulle, comme il est modélisé dans [155]. Ce passage à la limite fait l'objet de [56, Théorème 3]. Finalement, cette équation est modélisée par une équation aux dérivées partielles stochastiques. Elle est définie dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une filtration standard $(\mathcal{F}_t)_{t\geq 0}$ telle que $\mathcal{F}_0$ soit complète. Soit W_t un mouvement brownien standard réel associé à la filtration $(\mathcal{F}_t)_{t\geq 0}$. Le modèle est donné, après adimensionnement, par,

$$id\phi(t,x) = \left(\frac{1}{2}(-\Delta + |x|^2) + \lambda\left|\phi(t,x)\right|^2\right)\phi(t,x)\,\mathrm{d}t + \frac{\sigma_0}{2}\left|x\right|^2\phi(t,x)\circ \mathrm{d}W_t, \tag{1.5}$$

où $\circ$ représente un produit de Stratonovich.

Dans le cas défocalisant ($\lambda > 0$), on peut énoncer un résultat d'existence globale pour des solutions assez régulières. On introduit alors l'espace de Hilbert $\Sigma^j(\mathbb{R}^d)$ défini par,

$$\Sigma^j(\mathbb{R}^d, \mathbb{C}) = \left\{ v \in L^2(\mathbb{R}^d),\ \sum_{|\alpha|+|\beta|\leq j} \left\| x^\beta \partial^\alpha v \right\|_{L^2}^2 = \|v\|_{\Sigma^j}^2 < +\infty \right\}. \tag{1.6}$$

Proposition 1.1 (Proposition 3.1)**.** *Supposons $\lambda = 1$, et $d \leq 3$. Alors, pour tout $j \in \mathbb{N}^*$, si $\phi_0 \in \Sigma^j(\mathbb{R}^d)$ alors il existe une unique solution globale $(\phi(t))_{t\geq 0}$ de (1.5) $(\mathcal{F}_t)_{t\geq 0}$ adaptée, telle que $\phi(0) = \phi_0$, et presque sûrement dans $C(\mathbb{R}^+; \Sigma^j(\mathbb{R}^d))$.*

Dans la suite de ce paragraphe nous présentons les résultats des Chapitres 3 et 4, puis nous présentons quelques perspectives de recherche.

1.2.1 Résultats du Chapitre 3

Ce chapitre correspond au preprint [142] "Numerical analysis of the Gross-Pitaevskii Equation with a randomly varying potential in time".

L'objectif de ce chapitre est de proposer un schéma numérique pour l'Équation (1.5) basé sur une discrétisation de type Crank-Nicolson en temps, et une discrétisation spectrale en espace. Cette discrétisation temporelle s'est montrée très adaptée dans le cas déterministe car elle permet de conserver la norme L^2 ainsi qu'une énergie modifiée (voir [7, 13]). De plus, elle permet, dans notre cadre stochastique, de discrétiser de manière consistante l'intégrale de Stratonovich (similairement à [20, 51, 84]). De plus, des simulations numériques pour des équations (déterministes ou stochastiques) similaires ont montré de bons résultats de convergence pour ce type de discrétisation temporelle [7, 20, 83].

Nous démontrons que ce schéma converge à l'ordre au moins 1 en probabilité, dans des

espaces $\Sigma^j(\mathbb{R}^d)$ définis par (1.6). Nous présentons également des simulations numériques qui mettent en évidence l'optimalité de l'ordre 1.

L'analyse numérique de l'Équation (1.5) présente deux difficultés principales. La première vient de la non-linéarité, qui n'est pas globalement Lipschitzienne. C'est un problème classique de l'analyse numérique des équations de Schrödinger non-linéaires, à la fois dans le cas déterministe [7, 10, 11, 161] et dans le cas stochastique [20, 51]. Cette difficulté implique par exemple qu'il n'est pas démontré que la solution $(\phi(t))_{t\geq 0}$ donnée par la Proposition 1.1 appartienne à l'espace $L^\infty(0,T;L^2(\Omega,\Sigma^j(\mathbb{R}^d)))$ pour $j>1$. Il est donc illusoire d'espérer montrer une convergence dans ces espaces. C'est pourquoi la convergence de notre schéma numérique n'est montrée qu'en probabilité (voir [20, 42, 118] pour des résultats analogues). Cependant, nous montrons aussi qu'en tronquant la non-linéarité de l'Équation (1.5), et en tronquant la non-linéarité dans le schéma numérique de la même manière également, il devient possible de montrer une convergence en moindre carré (*i.e.* dans $L^2(\Omega)$). La deuxième difficulté provient du terme stochastique multiplicatif. Sa présence complique la démonstration de la stabilité d'une discrétisation temporelle de type Crank-Nicolson, et ce même pour la partie uniquement linéaire de l'Équation (1.5). C'est une difficulté nouvelle pour ce type d'équation de Schrödinger stochastique. Elle est résolue en remplaçant l'opérateur linéaire non-borné $|x|^2$ par une suite d'opérateurs linéaires bornés convergents vers $|x|^2$ et pour lesquels on contrôle la croissance de leur norme d'opérateur dans $L^2(\mathbb{R}^d)$. Cette méthode introduit alors une condition de type CFL qui limite la dimension de l'espace d'approximation en fonction du pas de temps.

Présentation du schéma numérique et résultats de stabilité

Nous précisons maintenant le schéma de Crank-Nicolson classique, uniquement semi-discrétisé en temps, afin de fixer les idées. Celui-ci sera par la suite modifié pour nous permettre d'établir nos résultats de stabilité et de convergence. Toutefois, certaines de ces modifications nous semblent essentiellement techniques, et ne semblent pas être nécessaires pour l'implémentation de ce schéma. Soit $(t_n)_{n\in\mathbb{N}}$ une subdivision de $\mathbb{R}^+$, de pas de temps $\delta t>0$; c'est-à-dire, pour tout $n\in\mathbb{N}$, $t_n=n\delta t$. On note aussi $\chi^{n+1}=\delta t^{-1/2}(W_{t_{n+1}}-W_{t_n})$ et on pose $A=-\Delta+|x|^2$. Nous définissons alors le processus discret $(\phi_n)_{n\in\mathbb{N}}$ qui représente une approximation, par un schéma de Crank-Nicolson classique, du processus $(\phi(t_n))_{n\in\mathbb{N}}$, où $(\phi(t))_{t\geq 0}$ est la solution de l'Équation (1.5) donnée par la Proposition 1.1, par,

$$\phi^{n+1}-\phi^n=-i\left(\delta t A+\sqrt{\delta t}\chi^{n+1}\,|x|^2\right)\left(\frac{\phi^{n+1}+\phi^n}{2}\right)-i\lambda\delta t g(\phi^{n+1},\phi^n),$$

où g est une approximation classique de la non-linéarité donnée par,

$$g(\phi^{n+1},\phi^n)=\frac{1}{2}(|\phi^n|^2+\left|\phi^{n+1}\right|^2)\left(\frac{\phi^{n+1}+\phi^n}{2}\right).$$

En définissant formellement les opérateurs $T_{\delta t,n+1}$ et $S_{\delta t,n+1}$ par,

$$T_{\delta t,n+1} = \left(\mathrm{Id} + i\frac{\delta t}{2}A + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}\,|x|^2\right),$$
$$S_{\delta t,n+1} = \left(\mathrm{Id} + i\frac{\delta t}{2}A + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}\,|x|^2\right)^{-1}\left(\mathrm{Id} - i\frac{\delta t}{2}A - i\frac{\sqrt{\delta t}}{2}\chi^{n+1}\,|x|^2\right),$$

le schéma de Crank-Nicolson s'écrit, sous réserve d'inversibilité de l'opérateur $T_{\delta t,n+1}$, et d'existence de l'opérateur $S_{\delta t,n+1}$,

$$\phi^{n+1} = S_{\delta t,n+1}\phi^n - i\lambda\delta t T^{-1}_{\delta t,n+1} g(\phi^{n+1},\phi^n). \tag{1.7}$$

Afin de démontrer un résultat de convergence en norme $\Sigma^k(\mathbb{R}^d)$ d'espace, et dans un sens trajectoriel et probabiliste que nous préciserons dans la suite, il est classiquement nécessaire de démontrer la stabilité dans des normes en espace plus fortes de type $\Sigma^j(\mathbb{R}^d)$ avec $j > k$ assez grand. Cependant, la stabilité, voire même simplement l'existence d'une solution régulière pour le schéma donné par l'Équation (1.7) pose des difficultés dans des espaces $\Sigma^j(\mathbb{R}^d)$ plus réguliers que $L^2(\mathbb{R}^d)$. Afin de mener l'analyse mathématique, nous proposons deux modifications des opérateurs $T_{\delta t,n+1}$ et $S_{\delta t,n+1}$. Nous proposons de modifier ces opérateurs d'une part en tronquant les incréments browniens, et d'autre part, en régularisant l'opérateur $|x|^2$ comme évoqué précédemment. Cette régularisation correspond à introduire une discrétisation spatiale, et à introduire une condition de type CFL. Nous utilisons alors une discrétisation spectrale en espace. Plus précisément, nous introduisons les sous-espaces vectoriels de dimension finie $\Sigma_K(\mathbb{R}^d)$ pour $K \in \mathbb{N}$. L'espace $\Sigma_K(\mathbb{R})$ est l'espace engendré par les K premières fonctions de Hermite, et les espaces $\Sigma_K(\mathbb{R}^d)$ pour $d > 1$ sont obtenus par tensorisation de l'espace $\Sigma_K(\mathbb{R})$. Nous définissons alors une famille d'opérateurs linéaires bornés $(B_K)_{K\in\mathbb{N}}$, à valeurs dans $\Sigma_K(\mathbb{R}^d)$, approximant l'opérateur linéaire $|x|^2$. Nous imposons à cette famille d'opérateurs de satisfaire les Hypothèses 3.6 et 3.7 Chapitre 3. La première hypothèse consiste à imposer aux opérateurs $(B_K)_{K\in\mathbb{N}}$ de reproduire certaines propriétés de régularité de l'opérateur $|x|^2$ liées principalement à son commutateur avec A. La seconde hypothèse caractérise la vitesse de convergence de la suite $(B_K)_{K\in\mathbb{N}^*}$ vers l'opérateur de multiplication par $|x|^2$. Nous notons $T_{\delta t,K,n+1,j}$ et $S_{\delta t,K,n+1,j}$ les opérateurs analogues à $T_{\delta t,n+1}$ et $S_{\delta t,n+1}$ pour lesquels le bruit est tronqué, et l'opérateur $|x|^2$ régularisé. L'indice K renvoie à l'espace d'approximation spatiale $\Sigma_K(\mathbb{R}^d)$, tandis que l'indice j paramètre le niveau de troncature des incréments browniens $(\chi^{n+1})_{n\in\mathbb{N}}$. Plus précisément, il est lié à l'espace $\Sigma^j(\mathbb{R}^d)$ dans lequel nous souhaitons prouver la stabilité du schéma. La définition précise de ces opérateurs est donnée dans le Chapitre 3, au Paragraphe 3.2.2, par les Équations (3.9) et (3.10).

Les deux idées présentées précédemment sont suffisantes pour assurer que le schéma numérique est bien posé dans le cas linéaire ($g = 0$), comme assuré par le Lemme 3.10. Comme nous l'avons énoncé plus haut, le fait que la non-linéarité ne soit pas globalement

Lipschitzienne pose des problèmes de définition. Pour assurer que le schéma est bien posé, la technique classique consiste à introduire une troncature de cette non-linéarité. Nous paramétrisons le niveau de cette troncature par $L > 0$. On note alors g_L^k la troncature de la non-linéarité g présentée ci dessus, dans l'espace $\Sigma^k(\mathbb{R}^d)$. La définition précise de cette troncature est donnée par l'Équation (3.6).

Le schéma numérique est finalement donné par,

$$\phi^{n+1} = S_{\delta t,K,n+1,j}\phi^n - i\lambda\delta t T^{-1}_{\delta t,K,n+1,j} g_L^k(\phi^{n+1},\phi^n). \tag{1.8}$$

L'existence et la stabilité du schéma numérique (1.8) sont données par la proposition suivante.

Proposition 1.2 (Proposition 3.9)**.** *Supposons que l'Hypothèse 3.6 énoncée dans le Chapitre 3 est vérifiée. Alors pour tout $k > d/2$, $j \geq k$, $L > 0$, $T > 0$ et $K_0 > 0$, il existe un pas de temps maximal $\delta t_0(j, L, K_0)$ et une constante $C(j, L, K_0, T) > 0$ tels que pour tout $\phi_0 \in \Sigma^j(\mathbb{R}^d)$, pour tout $\delta t = T/N \leq \delta t_0(j, L, K_0)$ et $K \leq K_0\delta t^{-1/4}$ il existe une unique solution discrète adaptée (à la filtration engendrée par le mouvement Brownien $(W_t)_{t\geq 0}$) $\phi = (\phi^n)_{n=0,\dots,N}$ satisfaisant l'Équation (1.8), presque sûrement dans $L^2(\Omega; L^\infty([0,T];\Sigma^j(\mathbb{R}^d)))$, et telle que*

$$\mathbb{E}\left[\sup_{n\leq N} \|\phi^n\|^2_{\Sigma^j}\right] \leq C(j, L, K_0, T)\,\|\phi_0\|^2_{\Sigma^j},$$

La preuve de ce résultat découle d'un théorème de point fixe dans l'ensemble des processus mesurables qui appartiennent à $L^2(\Omega; L^\infty([0,T];\Sigma^j(\mathbb{R}^d)))$. La difficulté principale réside dans la preuve de la stabilité dans le cas linéaire.

Résultats de convergence du schéma numérique

La convergence du schéma survient lors des passages à la limite $L, K \nearrow +\infty$, et $\delta t \searrow 0$. L'ordre dans lequel ces passages à la limite s'effectuent est crucial. En pratique on considère K comme une fonction de δt, que l'on note $K(\delta t)$, et telle que $K(\delta t)$ croisse à la vitesse $\delta t^{-1/4}$. La croissance de L en fonction de δt n'est pas claire. C'est la raison pour laquelle nous ne pouvons démontrer qu'un ordre de convergence en probabilité. Ce résultat est donné par le théorème suivant,

Theorem 1.3 (Théorème 3.14)**.** *Supposons les hypothèses 3.6 et 3.7 vérifiées. Pour tout $T > 0$, $k \in \mathbb{N}^*$, $\phi_0 \in \Sigma^{k+12}(\mathbb{R}^d)$, $K_0 > K_1 > 0$, $C > 0$, et pour tout $\alpha < 1$, il existe un choix de $L(\delta t)$ tel que,*

$$\lim_{\delta t\to 0} \sup_{n\delta t\leq T} \mathbb{P}\left(\left\|\phi^n_{K(\delta t),L(\delta t),\delta t} - \phi(t_n)\right\|_{\Sigma^k} \geq C\delta t^\alpha\right) = 0.$$

où $K(\delta t)$ vérifie pour tout $\delta t > 0$, $K_1 \leq K(\delta t)\delta t^{-1/4} \leq K_0$.

Nous renvoyons à la remarque 3.15 pour un commentaire sur l'hypothèse de régularité de la condition initiale ϕ_0. La preuve de ce résultat repose sur une majoration judicieuse de l'erreur

$$\left\|\phi^n_{K(\delta t),L(\delta t),\delta t} - \phi(t_n)\right\|_{\Sigma^k}.$$

Pour expliciter cette majoration, nous introduisons trois processus stochastiques. Le premier consiste seulement en une troncature de la non-linéarité,

$$\begin{cases} id\phi_L - \lambda f^k_L(\phi_L)\,\mathrm{d}t - A\phi_L\,\mathrm{d}t = |x|^2\phi_L \circ \mathrm{d}W_t, \\ \phi_L(0) = \phi_0, \end{cases}$$

où $f^k_L(\phi_L) = g^k_L(\phi_L, \phi_L)$. Le second processus incorpore de plus l'approximation spatiale,

$$\begin{cases} id\phi_{K(\delta t),L} - \lambda P_{K(\delta t)} f^k_L(\phi_{K(\delta t),L})\,\mathrm{d}t - P_{K(\delta t)} A\phi_{K(\delta t),L}\,\mathrm{d}t = B_{K(\delta t)}\phi_{K(\delta t),L} \circ \mathrm{d}W_t, \\ \phi_{K(\delta t),L}(0) = P_{K(\delta t)}\phi_0. \end{cases}$$

Le dernier processus prend finalement en compte le temps d'arrêt $\tau(\delta t)$ qui caractérise la première troncature des incréments browniens,

$$\psi^n_{K(\delta t),L,\delta t} = \begin{cases} P_K\phi_0, & \text{si} \quad n = 0, \\ \phi_{K(\delta t),L}(t_n), & \text{si} \quad 0 < n < \tau(\delta t), \\ \psi^{n-1}_{K(\delta t),L,\delta t}, & \text{sinon.} \end{cases}$$

En notant la $(\phi^n_{K(\delta t),L,\delta t})_{n\in\mathbb{N}}$ la solution du schéma numérique (1.8), nous séparons l'erreur $\left\|\phi^n_{K(\delta t),L,\delta t} - \phi(t_n)\right\|_{\Sigma^k}$ en quatre termes :

$$\begin{aligned} \left\|\phi^n_{K(\delta t),L,\delta t} - \phi(t_n)\right\|_{\Sigma^k} \leq{}& \left\|\phi^n_{K(\delta t),L,\delta t} - \psi^n_{K(\delta t),L,\delta t}\right\|_{\Sigma^k} + \left\|\psi^n_{K(\delta t),L,\delta t} - \phi_{K(\delta t),L}(t_n)\right\|_{\Sigma^k} \\ &+ \left\|\phi_{K(\delta t),L}(t_n) - \phi_L(t_n)\right\|_{\Sigma^k} + \left\|\phi_L(t_n) - \phi(t_n)\right\|_{\Sigma^k}. \end{aligned}$$

Nous démontrons ensuite que le premier et le troisième terme d'erreur du membre de droite convergent vers 0 dans $L^2(\Omega)$ quand δt tend vers 0. Le deuxième et le quatrième terme sont rendus de probabilité arbitrairement petite en choisissant L assez grand.

Résultats numériques

Dans la dernière partie du Chapitre 3, nous présentons trois autres schémas numériques, classiques dans le cas déterministe, pour ce type d'équations de Schrödinger. Nous présentons notamment une discrétisation temporelle de type *splitting*, c'est-à-dire basée sur une décomposition d'opérateurs, ainsi qu'une discrétisation en espace de type différences finies et une autre de type spectrale-Fourier. Nous présentons quelques résultats numériques de convergence pour ces schémas. Nous observons de bons résultats de convergence dans le

cas où l'intensité des fluctuations est petite, ce qui est le cas pratique d'intérêt. Dans le cas où les fluctuations ont une forte intensité, on observe que notre condition de type CFL entre δt et K est utile pour assurer la convergence de la solution numérique.

1.2.2 Résultats du Chapitre 4

Ce chapitre correspond à l'article [57] "Vortex solutions in Bose-Einstein condensation under a trapping potential varying randomly in time", publié dans Discrete and Continuous Dynamical Systems - Series B.

Nous étudions théoriquement et numériquement au Chapitre 4 l'influence des perturbations aléatoires du potentiel confinant sur la dynamique de solutions stationnaires (en l'absence de perturbations aléatoires) de type vortex dans le cas bidimensionnel. L'ajout du bruit dans la dynamique de Gross-Pitaevskii implique que les solitons de type vortex ne devraient pas persister à tous temps. Il est alors naturel de se demander combien de temps ces solitons peuvent persister, en fonction de l'intensité du bruit. Nous introduisons un paramètre $\varepsilon > 0$ devant le terme stochastique dans l'Équation (1.5). En prenant $\sigma_0 = 1$, celle-ci est donnée (à une renormalisation près, et en généralisant la non-linéarité) par,

$$id\psi(t,x) = \left(-\Delta + |x|^2 + \lambda\,|\psi(t,x)|^{2\sigma}\right)\psi(t,x)\,\mathrm{d}t + \varepsilon\,|x|^2\,\psi(t,x)\circ\,\mathrm{d}W_t. \tag{1.9}$$

Cette étude de la dynamique des solitons de type vortex est basée sur la méthode des coordonnées collectives. Elle consiste à décomposer la solution de l'Équation (1.9) en la somme d'une modulation (aléatoire) de la solution vortex initiale et d'un reste d'ordre ε. Cette modulation correspond à un déphasage de la solution vortex, ou encore à une rotation de celle-ci. Cette méthode repose tout particulièrement sur une propriété de stabilité orbitale des solitons, que nous précisons dans la suite. Elle a été utilisée pour de nombreuses équations dispersives à la fois dans le cadre déterministe [79, 104, 168] et dans le cadre stochastique [52, 53, 55, 58]. Nous établissons deux résultats principaux. Le premier assure que ce reste demeure *petit* pour des temps d'ordre au moins ε^{-2}. Nous reviendrons sur ce point, mais nous pouvons déjà préciser formellement que c'est notre choix de modulation aléatoire qui permet d'approcher la solution comme un vortex modulé pendant un temps caractéristique aussi long. Le second résultat caractérise le comportement asymptotique de la modulation et du reste quand ε tend vers 0. Plus précisément, il définit deux processus stochastiques, qui sont des semi-martingales correspondant aux processus limites de modulation et de reste, et démontre également la convergence en probabilité de la modulation et du reste vers ces processus. Dans une dernière partie, nous mettons en évidence numériquement une certaine optimalité des résultats théoriques obtenus, notamment à l'aide d'une méthode de Monte Carlo adaptée à la simulation d'événements rares.

Préliminaires

Une solution stationnaire vortex de l'Équation (1.9) dans le cas $\varepsilon = 0$ est une solution pour laquelle on peut séparer les variables t, r et θ (en coordonnées polaires). Ces solutions

s'écrivent alors,

$$u(t,r,\theta) = \mathrm{e}^{-i\mu t}\phi_{\mu,m}(r,\theta) = \mathrm{e}^{-i\mu t}\mathrm{e}^{im\theta}\psi_{\mu,m}(r).$$

On appelle m le degré, μ le potentiel chimique, et ψ le profil du vortex. Nous précisons (voir [120]) que pour tout entier m et pour tout réel positif μ assez grand, il existe une unique fonction $\psi_{\mu,m}(r)$ tel que $\mathrm{e}^{-i\mu t}\mathrm{e}^{im\theta}\psi_{\mu,m}(r)$ soit une telle solution stationnaire. Nous notons $\phi_{\mu,r}(r,\theta) = \mathrm{e}^{im\theta}\psi_{\mu,m}(r)$ Nous définissons le sous-espace fermé X_m composé des éléments de $\Sigma^1(\mathbb{R}^2)$ qui possèdent cette séparation en les variables r et θ,

$$X_m := \left\{ v \in \Sigma^1(\mathbb{R}^2), \quad \mathrm{e}^{-im\theta}v(x)\text{ne dépend pas de } \theta \right\}.$$

Le Lemme 4.3 assure qu'une solution de l'Équation (1.9) initialisée par un élément de X_m appartient presque sûrement à $C(\mathbb{R}^+; X_m)$.

Nous présentons maintenant la décomposition d'une solution $(u_\varepsilon(t))_{t\geq 0}$ initialisée par une solution stationnaire $\phi_{\mu,m}$. Nous cherchons à définir deux processus $(\xi^\varepsilon(t))_{t\geq 0}$ et $(\eta^\varepsilon(t))_{t\geq 0}$ tels que,

$$u^\varepsilon(t,x) = \mathrm{e}^{-i\xi^\varepsilon(t)}\left(\phi_{\mu,m}(x) + \varepsilon\eta^\varepsilon(t,x)\right),$$

où la modulation $(\xi^\varepsilon(t))_{t\geq 0}$ est à valeurs réelles, et le reste $(\eta^\varepsilon(t))_{t\geq 0}$ est à valeurs dans X_m. Cette décomposition n'étant pas unique, nous imposons en plus la condition d'orthogonalité suivante,

$$\mathrm{Re}(\eta^\varepsilon, i\phi_{\mu,m})_{L^2(\mathbb{R}^2,dx)} = 0, \quad \text{a.s.}, \quad t \leq \tau^\varepsilon, \tag{1.10}$$

Nous remarquons que le processus $\xi^\varepsilon(t)$ qui minimise presque sûrement la norme $L^2(\mathbb{R}^2)$ du reste $\eta^\varepsilon(t)$ à tous temps satisfait la condition d'orthogonalité (1.10).

Résultats théoriques

Le premier résultat principal du Chapitre 4 assure qu'il existe de tels processus $(\xi^\varepsilon(t))_{t\geq 0}$ et $(\eta^\varepsilon(t))_{t\geq 0}$ qui sont plus précisément des semi-martingales, et tels que le reste $\varepsilon\eta^\varepsilon(t)$ demeure petit pour des temps de l'ordre de ε^{-2}.

Theorem 1.4 (Théorème 4.5). *Soient $1/2 \leq \sigma < \infty$, $\mu_0 > 0$ et $m \in \mathbb{N}$ avec μ_0 assez grand par rapport à m. Pour tout $\varepsilon > 0$, soit $u^\varepsilon(t,x)$, la solution de l'Équation (1.9) avec $u(0,x) = \phi_{\mu_0,m}(x)$. Alors il existe $\alpha_0 > 0$ tel que, pour tout $\alpha \in]0,\alpha_0]$, il existe un temps d'arrêt $\tau^\varepsilon_\alpha \in (0,\infty)$ p.s., et une semi-martingale $\xi^\varepsilon(t)$, définie presque sûrement pour tout $t \leq \tau^\varepsilon_\alpha$, et à valeurs dans $\mathbb{R}$, telle que si nous posons $\varepsilon\eta^\varepsilon(t,x) = \mathrm{e}^{i\xi^\varepsilon(t)}u^\varepsilon(t,x) - \phi_{\mu_0,m}(x)$, alors la condition d'orthogonalité (1.10) est vérifiée. De plus, p.s. pour tout $t \leq \tau^\varepsilon_\alpha$,*

$$|\varepsilon\eta^\varepsilon(t)|_\Sigma \leq \alpha. \tag{1.11}$$

De plus, il existe une constante $C = C_{\alpha,\mu_0} > 0$, tel que pour tout $T > 0$ et pour tout $\alpha \leq \alpha_0$,

il existe $\varepsilon_0 > 0$, *tel que pour tout* $\varepsilon < \varepsilon_0$,

$$\mathbb{P}(\tau_\alpha^\varepsilon \leq T) \leq \exp\Big(-\frac{C}{\varepsilon^2 T}\Big). \tag{1.12}$$

La preuve de l'existence des semi-martingales $(\xi^\varepsilon(t))_{t\geq 0}$ et $(\eta^\varepsilon(t))_{t\geq 0}$ satisfaisant la condition d'orthogonalité (1.10) découle de l'application du théorème des fonctions implicites. La preuve de la seconde partie du Théorème 1.4 découle d'une méthode par fonction de Lyapunov. On introduit pour ce faire l'énergie

$$H(u) = \frac{1}{2}|\nabla u|_{L^2}^2 + \frac{1}{2}|xu|_{L^2}^2 + \frac{\lambda}{2\sigma+2}|u|_{L^{2\sigma+2}}^{2\sigma+2}, \tag{1.13}$$

ainsi que la norme $L^2(\mathbb{R}^2)$,

$$Q(u) = \frac{1}{2}|u|_{L^2}^2.$$

La fonction de Lyapunov utilisée, notée S_μ, et appelée aussi *fonctionnelle d'action*, est donnée par,

$$S_\mu(u) = H(u) - \mu Q(u), \quad u \in \Sigma^1(\mathbb{R}^d).$$

On note également que $\phi_{\mu,m}$ est le minimum global de S_μ dans l'ensemble X_m. La preuve utilise le fait que la Hessienne de S_μ en $\phi_{\mu,m}$ est positive, et que sa restriction à l'orthogonal de $i\phi_{\mu,m}$ (pour la norme $L^2(\mathbb{R}^2)$) est définie positive, ainsi que des estimées sur la croissance de l'énergie au cours du temps.

Le deuxième résultat principal assure de la convergence en probabilité dans $C([0, \tau_\alpha^\varepsilon \wedge T], L^2(\mathbb{R}^2))$ du processus η^ε vers un certain processus η. Un résultat de convergence pour le processus ξ^ε est également donné. Nous notons par $|\cdot|_{L_r^2}$ la norme L^2 radiale.

Theorem 1.5 (Théorème 4.6). *Fixons* $1 \leq \sigma < \infty$, μ_0 *assez grand par rapport à* m *et soient* η^ε, ξ^ε, *pour* $\varepsilon > 0$ *donnés par le Théorème 1.4, avec* $\alpha \leq \alpha_0$ *fixé. Alors, pour tout* $T > 0$, *le processus* $(\eta^\varepsilon(t))_{t\in[0,T\wedge\tau_\alpha^\varepsilon]}$ *à valeurs dans* X_m *converge en probabilité, quand* ε *tend vers 0, vers un processus* $\eta = (\eta(t))_{t\in[0,T]}$ *satisfaisant* $\mathrm{e}^{-im\theta}\eta = \tilde{\eta}$ *et*

$$d\tilde{\eta} = J\begin{pmatrix} L_{\mu_0,m}^{1,r} & 0 \\ 0 & L_{\mu_0,m}^{2,r}\end{pmatrix}\tilde{\eta}\,\mathrm{d}t - y\begin{pmatrix}0\\ \psi_{\mu_0,m}\end{pmatrix}\mathrm{d}t - \begin{pmatrix}0\\ r^2\psi_{\mu_0,m}\end{pmatrix}\mathrm{d}W - z\begin{pmatrix}0\\ \psi_{\mu_0,m}\end{pmatrix}\mathrm{d}W, \tag{1.14}$$

où

$$J = -i : \begin{pmatrix}\mathrm{Re}\ u\\ \mathrm{Im}\ u\end{pmatrix} \mapsto \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\mathrm{Re}\ u\\ \mathrm{Im}\ u\end{pmatrix},$$

et où

$$\begin{aligned} L_{\mu,m}^{1,r} &= -\frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + r^2 + \frac{m^2}{r^2} - \mu + (2\sigma+1)\psi_{\mu,m}^{2\sigma},\\ L_{\mu,m}^{2,r} &= -\frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + r^2 + \frac{m^2}{r^2} - \mu + \psi_{\mu,m}^{2\sigma}. \end{aligned}$$

avec $\tilde{\eta}(0) = 0$, *et*

$$y(t) = \frac{2\sigma(\mathrm{Re}\,\tilde{\eta}, \psi_{\mu_0,m}^{2\sigma+1})_{L_r^2}}{|\psi_{\mu_0,m}|_{L_r^2}^2}, \quad z(t) = \frac{|r\psi_{\mu_0,m}|_{L_r^2}^2}{|\psi_{\mu_0,m}|_{L_r^2}^2}, \tag{1.15}$$

pour tout $t \geq 0$. *La convergence a lieu dans* $C([0, \tau_\alpha^\varepsilon \wedge T], L^2)$, *et il existe une constante* $C > 0$ *tel que le processus* η *défini ci-dessus pour tout* $T > 0$ *satisfait l'estimée*

$$\mathbb{E}\left(\sup_{t\leq T} |\eta(t)|_{\Sigma^1}^2\right) \leq CT.$$

De plus, le paramètre de modulation ξ^ε *satisfait, pour tout* $t \leq \tau_\alpha^\varepsilon$,

$$d\xi^\varepsilon = \mu_0 \,\mathrm{d}t + \varepsilon y^\varepsilon \,\mathrm{d}t + \varepsilon z^\varepsilon \,\mathrm{d}W,$$

où y^ε *et* z^ε *sont des processus adaptés à valeurs réelles, convergeant en probabilité dans* $C([0,T])$ *respectivement vers* y, z *donnés par* (1.15).

Le calcul des équations satisfaites par les processus $(\eta^\varepsilon(t))_{t\geq 0}$ et $(\xi^\varepsilon(t))_{t\geq 0}$ repose principalement sur l'application du Lemme d'Itô et sur l'utilisation de la condition d'orthogonalité (1.10). La preuve de la convergence nécessite dans un premier temps d'obtenir des estimées uniformes en ε sur le reste η^ε. La non-linéarité est à l'origine de quelques difficultés. Ces estimations nécessitent par exemple une borne de $\|\eta^\varepsilon(t)\|_{L^\infty}$. Pour cela, nous introduisons (comme cela avait déjà été fait dans [55]) le temps d'arrêt $\tilde{\tau}_N^\varepsilon$ défini pour tout $N > 0$ par,

$$\tilde{\tau}_N^\varepsilon = \inf\left\{t \leq \tau^\varepsilon \wedge T,\ \|\varepsilon\eta^\varepsilon(t)\|_{\Sigma^2} \geq N\right\}. \tag{1.16}$$

Il est montré dans un premier temps la convergence du processus $(\eta^\varepsilon(t))_{t\geq 0}$ vers le processus η défini par le Théorème 1.5 dans l'espace $L^2(\Omega; C([0, \tilde{\tau}_N^\varepsilon \wedge T], L^2))$. Dans un second temps, nous passons à la limite $N \to +\infty$. C'est lors de cette étape que nous perdons la convergence en moindre carré pour une convergence en probabilité. Cette étape repose en particulier sur le fait que la solution de l'Équation (1.9) appartient presque sûrement à $C(\mathbb{R}^+; \Sigma^2(\mathbb{R}^2))$. Notons que ce phénomène de dégradation de la convergence est similaire à celui du Chapitre 3 où l'on passe à la limite sur la troncature de la non-linéarité. Nous passons de la même manière d'une convergence en moindre carré à une convergence en probabilité.

Résultats numériques

Dans cette partie, nous proposons de mettre en évidence numériquement les résultats du Théorème 1.4. Nous présentons des résultats numériques qui tendent à montrer l'optimalité (dans une certaine mesure) de ces résultats. Pour cela, nous proposons un schéma basé sur une discrétisation de type Crank-Nicolson en temps, et sur une discrétisation

par différences finies en espace. C'est un schéma unidimensionnel qui décrit l'évolution temporelle du profil de la solution. Contrairement à ce qui est fait au Chapitre 3, la non-linéarité ne conduit pas à un schéma non-linéairement implicite. Pour éviter l'utilisation d'un point fixe, nous proposons d'utiliser une méthode de relaxation où la non-linéarité est approximée par une extrapolation sur les deux derniers pas de temps.

Le calcul de la condition initiale repose sur une méthode de tir, et quelques exemples de solutions stationnaires sont présentés au Chapitre 4. Nous présentons aussi quelques trajectoires d'évolution des solutions de l'Équation (1.9) pour différents degrés m. Finalement, nous estimons par une méthode de Monte Carlo les probabilités $\mathcal{P}(\tau_\alpha^\varepsilon \leq t)$ par rapport à t et par rapport à ε pour $\alpha = 2.5 \cdot 10^{-3}$ et $\sigma = 1$. Ces résultats sont donnés en Figures 4.8 et 4.9. Ces simulations tendent à vérifier que $-\log\left(\mathcal{P}(\tau_\alpha^\varepsilon \leq t)\right)$ est asymptotiquement proportionnel à ε^{-2} et à t^{-1}.

Nous proposons également d'utiliser une méthode de Monte Carlo adaptée à l'estimation des probabilités $\mathcal{P}(\tau_\alpha^\varepsilon \leq t)$ dans le cas où celles-ci sont très petites, ce qui est le cas quand ε est choisi très petit. Nous proposons, pour ce faire, d'utiliser les algorithmes POP et IPS [87]. Pour cela, nous exprimons les événements rares en terme d'ensembles de trajectoires browniennes (qui conduisent à ces événements rares), ce qui permet d'utiliser une *shaking transformation* sur les trajectoires browniennes proposée dans [87].

1.2.3 Perspectives

Le Chapitre 3 propose une preuve de la convergence d'un schéma de Crank-Nicolson pour l'Équation (1.5) à l'ordre 1. Cependant, pour assurer la stabilité du schéma en norme $\Sigma^k(\mathbb{R}^d)$, nous n'avons pu nous passer d'une condition de type CFL. Un schéma de *splitting*, c'est-à-dire de décomposition d'opérateurs, permettrait de se passer de cette difficulté, et l'étude de sa convergence pourrait-être envisagé. Ce type de discrétisation temporelle a été étudiée dans un cas similaire pour un bruit espace-temps [118].

La question de la discrétisation spatiale reste également ouverte pour ce type de discrétisation temporelle. Pour des équations de type Schrödinger, une discrétisation temporelle de type splitting permet généralement de diagonaliser chaque sous-équation, ce qui permet leur résolution exacte dans une base adaptée, au prix d'un changement de base à chaque passage d'une équation à l'autre. Ce changement de base s'accompagne naturellement d'une projection, ou d'une interpolation, sur l'espace de discrétisation entre chaque résolution (voir Paragraphe 3.5). Dans ce cadre, une question intéressante est la quantification de la différence entre ces deux opérateurs. Nous nous attendons, par exemple, à ce que l'opérateur d'interpolation possède un bien meilleur comportement dans certains cas, et en particulier pour l'Équation (1.5).

1.3 Modélisation des condensats de Bose-Einstein à température non nulle

Effets de la température

Bien que l'équation de Gross-Pitaevskii permette de modéliser précisément (sous certaines hypothèses statistiques) la dynamique d'un condensat de Bose Einstein, celle-ci ne permet pas de modéliser les phénomènes liés à la température non-nulle. Dans cette situation, le condensat coexiste avec une composante non condensée avec laquelle il interagit. Pour des températures de l'ordre de la température critique de transition de phase, ces interactions sont cruciales pour comprendre le processus de croissance lors de la formation expérimentale du condensat. En effet, cette phase peut être accompagnée de la formation de défauts topologiques qui prennent la forme de vortex dans notre cas [132, 167]. Ce processus est modélisé par le scénario de Kibble-Zurek [59, 107, 166, 172]. Il s'agit d'un scénario qui présente certaines classes d'universalité, et que l'on peut retrouver tant en matière condensée qu'en cosmologie [107]. Nous présenterons dans le paragraphe suivant, et aussi au Paragraphe 2.4, un modèle permettant de prendre en compte ces interactions avec la phase non-condensée. Celui-ci a notamment permis de reproduire numériquement la formation spontanée de vortex durant la transition de phase [167]. Cependant, les applications de ce modèle restent à ce jour peu nombreuses. Dans le Chapitre 6, nous proposons une nouvelle application de ce modèle en nous intéressant au comportement métastable de sa solution.

Modélisation du système ouvert en interaction avec une phase non-condensée

Nous présentons dans cette partie un modèle qui permet de décrire les interactions entre un système condensé et un système non-condensé supposé en équilibre thermodynamique. Ce modèle a été dérivé d'abord par Stoof [159] puis par Gardiner et co-auteurs [80, 81]. Ces trois articles présentent trois dérivations différentes, mais qui aboutissent, sous certaines hypothèses, au même modèle. Nous présenterons quelques aspects de la modélisation de Gardiner [81].

La première hypothèse de ce modèle consiste à supposer que les niveaux de plus haute énergie ne sont pas occupés. Cela revient à éliminer tous les modes d'énergie supérieure à une énergie E_{eff}. Cette hypothèse est valide si $k_B T \ll E_{\text{eff}}$. C'est une simplification essentiellement technique pour justifier la dérivation physique du modèle.

La seconde hypothèse consiste à séparer le système en une partie condensée, et une partie non-condensée. Pour cela on introduit un niveau d'énergie E_{cut} (inférieure à E_{eff}) tel que les modes d'énergie inférieure à ce niveau seront considérés comme condensés, et les autres comme non-condensés. On parle aussi respectivement de région cohérente et incohérente. Malgré cette dénomination, la partie condensée contient plus d'atomes que simplement le condensat. Le niveau E_{cut} est choisi de telle sorte que les modes de la bande condensée soient occupés par au moins une dizaine d'atomes, pour permettre les

approximations nécessaires à la dérivation du modèle. Cette division permet de traiter les modes condensés à l'aide d'un formalisme purement quantique, tandis que les modes non-condensés sont traités comme un bain d'atomes thermalisés. Le modèle *Stochastic Projected Gross-Pitaevskii Equation* (SPGPE) constitue une description quantique du système ouvert constitué uniquement des particules dans les modes dits condensés. Il régit l'évolution temporelle de la fonction d'onde des modes condensés. Il prend la forme d'une équation de Schrödinger dont le support de la solution est l'espace de dimension finie engendré par les modes d'énergie inférieure à E_{cut}. Plus précisément, cette équation prend la forme d'une équation aux dérivées partielles stochastiques. La dérivation de ce modèle emploie des méthodes et des formalismes issus de la physique qui dépassent largement le cadre de cette introduction. L'idée générale consiste à construire une équation de Fokker-Planck pour la fonction de Wigner du champ quantique correspondant aux modes de la région cohérente, puis d'en déduire une représentation stochastique.

En pratique, nous ne considérons qu'un modèle simplifié du modèle complet, poussés à la fois par des considérations physiques qui justifient de négliger certains termes, et par des considérations calculatoires, pour réduire les coûts de calcul et simplifier les simulations.

Nous proposons dans la suite une formulation adimensionnée et simplifiée en dimension $d \in \{2, 3\}$ pour faciliter le traitement mathématique de ce modèle. Une formulation adimensionnée plus générale est donnée par l'Équation (2.35). Le lien entre cette dernière équation et la formulation proposée dans [25, Paragraphe 5.3] est détaillé au Paragraphe 2.4 du Chapitre 2. Par souci de simplification, le potentiel confinant est choisi isotropique. La réduction à un modèle bidimensionnel peut être obtenue sous certaines hypothèses. Celles-ci sont typiquement valides dans le cas où le confinement est beaucoup plus intense dans une direction d'espace. La procédure de réduction de la dimension est donnée au Paragraphe 2.4.2 dans un cadre plus général.

Le modèle SPGPE est donné par l'équation,

$$d\phi_t = -(i + c_1)\nabla E^K(\phi_t)\,\mathrm{d}t + c_1 c_3 dB_t, \tag{1.17}$$

où les constantes positives c_1 et c_3 proviennent de l'adimensionnement conduit au paragraphe 2.4.3. Le processus $(\phi_t)_{t\geq 0}$ est défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et est à valeurs dans un $\mathbb{C}$-sous-espace vectoriel de $L^2(\mathbb{R}^d, \mathbb{C})$ de dimension finie, que l'on note K et que l'on précise dans la suite. La dimension d est 2 ou 3 et les constantes c_1 et c_3 sont strictement positives. Le gradient ∇E^K est une application de K dans K, également précisée dans la suite. Le processus $(B_t)_{t\geq 0}$ est un mouvement brownien à valeurs dans K. C'est à dire que pour toute base orthogonale $(e_k)_{0\leq k<\dim(K)}$ pour le produit scalaire hermitien $L^2(\mathbb{R}^d, \mathbb{C})$, il existe une famille $((\beta_k(t))_{t\geq 0})_{0\leq k<\dim(K)}$ de mouvements browniens à valeurs dans $\mathbb{C}$ tels que,

$$B_t = \sum_{0\leq k<\dim(K)} \beta_k(t) e_k.$$

Ce mouvement brownien peut également s'interpréter comme la projection orthogonale, pour le produit scalaire hermitien de $L^2(\mathbb{R}^d, \mathbb{C})$, sur le $\mathbb{C}$-sous-espace vectoriel K d'un processus de Wiener de $L^2(\mathbb{R}^d, \mathbb{C})$.

Nous précisons maintenant la définition de ∇E^K avant de revenir à celle de K. Tout d'abord définissons l'énergie de Gross-Pitaevskii E pour un système en rotation. Celle-ci est définie pour tout élément $\phi \in L^2(\mathbb{R}^d, \mathbb{C})$ par,

$$E(\phi) = \frac{1}{2}\int_{\mathbb{R}^d} \left[\left(\frac{1}{2}(-\Delta + |\mathbf{x}|^2) - \rho L_z - \mu \mathrm{Id}\right)\phi(\mathbf{x})\right]\overline{\phi}(\mathbf{x})\mathrm{d}\mathbf{x} + \frac{1}{4}\int_{\mathbb{R}^d} g\,|\phi(\mathbf{x})|^4\,\mathrm{d}\mathbf{x}. \quad (1.18)$$

Cette équation est un enrichissement de l'énergie définie par l'Équation (1.2). Le coefficient $\rho \in (0, 1)$ caractérise la vitesse de rotation du condensat, et le coefficient $\mu > 0$ représente le potentiel chimique, et enfin $g > 0$ caractérise l'intensité des interactions. Comme pour l'équation de Gross-Pitaevskii, le terme $|\mathbf{x}|^2$ modélise le potentiel de confinement externe du condensat, et l'opérateur L_z, défini par $L_z = i(x\partial_y - y\partial_x)$, est l'opérateur de rotation. Nous définissons ensuite E^K la restriction de E au sous-espace vectoriel K, ainsi que ∇E^K comme le gradient de cette application. Cependant, comme cette dernière est une application du $\mathbb{C}$-espace vectoriel K, à valeurs dans $\mathbb{R}$, le choix du corps des scalaires ($\mathbb{R}$ ou $\mathbb{C}$) pour les espaces vectoriels de départ et d'arrivée de l'application E^K n'est pas clair, et donc la définition de ∇E^K non plus. Dans notre cas, nous choisissons le corps $\mathbb{R}$ et nous considérons K comme un $\mathbb{R}$-espace vectoriel en identifiant $\mathbb{C}$ à $\mathbb{R}^2$. Notons que l'alternative pour pouvoir définir la différentielle de l'application E^K serait d'étendre son espace d'arrivée à $\mathbb{C}$ tout entier, et de considérer $\mathbb{C}$ comme corps des scalaires. Cependant, dans ce cas, l'application ne serait pas différentiable. Le calcul donne alors pour tout $\phi \in K$,

$$\nabla E^K(\phi) = P_K\left(\left(\frac{1}{2}(-\Delta + |\mathbf{x}|^2) - \rho L_z - \mu \mathrm{Id}\right)\phi + g\,|\phi|^2\,\phi\right),$$

où l'application P_K désigne la projection orthogonale pour la norme $L^2(\mathbb{R}^d)$ sur l'espace vectoriel K.

Comme nous l'avons expliqué plus haut, la phase condensée est constituée par l'ensemble des bosons répartis sur les modes d'énergie plus petite que le niveau E_{cut}. Cette restriction de la phase condensée aux modes les plus bas correspond mathématiquement à restreindre la fonction d'onde de la phase condensée à un sous-espace vectoriel de dimension finie, engendré par ces modes les plus bas. Pour $\rho \in [0, 1)$ l'opérateur linéaire A, de domaine $\Sigma^2(\mathbb{R}^d)$ (défini par l'Équation (1.6)) à valeurs dans le dual de Σ, et défini par,

$$A = \frac{1}{2}(-\Delta + |x|^2) - \rho L_z, \quad (1.19)$$

est un opérateur auto-adjoint à résolvante compacte (car l'injection $\Sigma(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d)$ est compacte), son spectre est donc purement ponctuel. Les modes d'énergie évoqués précédemment sont définis comme les vecteurs propres de cet opérateur. L'espace vectoriel

K est alors l'espace engendré par ces premiers modes. Sa dimension dépend de E_{cut}, et d'autres paramètres physiques. La dérivation physique du modèle donne un calcul de cette dimension (dont le détail est donné au Paragraphe 2.4.1).

En pratique, le coefficient c_1 est très petit devant 1 (de l'ordre de $10^{-3} \sim 10^{-5}$) Les termes prépondérants dans l'Équation (1.17) correspondent à la dynamique donnée par l'équation de Gross-Pitaevskii (à la restriction à l'espace K près). C'est pourquoi ce modèle peut être vu comme une correction de l'équation de Gross-Pitaevskii déterministe.

Nous pouvons conclure sur un résultat d'existence et d'unicité forte pour l'Équation (1.17),

Theorem 1.6. *Pour tout $\phi_0 \in K$, il existe une unique solution globale $(\phi_t)_{t\geq 0}$ de l'Équation (1.17), adaptée par rapport à la filtration engendrée par le mouvement brownien $(B_t)_{t\geq 0}$, et dont les trajectoires sont continues en temps et en espace.*

L'existence et l'unicité locale en temps découlent du caractère localement Lipschitzien du drift. Le caractère global n'est pas immédiat car le terme de dérive n'est pas globalement Lipschitzien. Cependant, le terme de dérive est dissipatif pour la norme $L^2(\mathbb{R}^2)$. Pour le montrer, il suffit d'appliquer la formule d'Itô à la fonction $\phi \mapsto \frac{1}{2}\|\phi\|_{L^2}^2$ pour une solution locale de l'Équation (1.17), puis d'utiliser le Lemme de Gronwall pour borner la croissance du processus en norme $L^2(\Omega; L^2(\mathbb{R}^2))$. Pour justifier précisément ce calcul, il est possible de tronquer la non-linéarité pour la rendre globalement Lipschitz, puis de faire tendre le niveau de troncature vers l'infini.

1.3.1 Résultats du Chapitre 5

Ce chapitre correspond au preprint [141] "Generalized and hybrid Metropolis-Hastings overdamped Langevin algorithms".

Objectifs du chapitre

Le Chapitre 5 a pour but de proposer deux méthodes d'échantillonnage non biaisées d'une mesure de probabilité π de dimension finie, absolument continue par rapport à la mesure de Lebesgue. Cet objectif est motivé par le fait que l'Équation (1.17) possède une unique mesure invariante. Nous la supposons définie sur $\mathbb{R}^d$, et nous notons également π sa densité par abus de langage. Nous définissons aussi U le potentiel associé à la densité π par $U(x) = -\log \pi(x)$ pour tout $x \in \mathbb{R}^d$. Nous pouvons aussi supposer que le potentiel U n'est connu qu'à une constante additive près. Les méthodes d'échantillonnage proposées au Chapitre 5 permettent de calculer des espérances sous π de certaines observables $f \in L^1(\pi)$, c'est-à-dire de calculer des intégrales,

$$\pi(f) := \mathbb{E}_\pi(f) = \int_{\mathbb{R}^d} f(x)\pi(dx),$$

à l'aide d'une méthode de Monte-Carlo. Dans notre cas d'application, la mesure cible d'intérêt est la mesure invariante de l'Équation (1.17), donnée par l'Équation (5.3). De

telles quantités peuvent fournir des informations quantitatives sur le comportement en temps long des solutions de l'Équation (1.17). Cependant, les algorithmes proposés dans ce chapitre dépassent ce cas d'application et sont donc présentés dans un cadre plus général. Le problème de l'échantillonage de mesure en grande dimension est par exemple un problème central en statistique bayésienne [37] ainsi qu'un physique statistique numérique [36]. Le lien précis entre notre cas d'application et ces méthodes est discuté au Paragraphe 5.1.

Nous proposons deux méthodes d'échantillonnage appartenant à la classe des méthodes de type Markov-Chain Monte-Carlo (MCMC). Leur principe général est de construire une chaîne de Markov ergodique par rapport à la mesure cible à échantillonner π. Un exemple classique d'une telle méthode est donné par l'algorithme *Metropolis-Adjusted Langevin Algorithm* (MALA) ([152, 154]). Celui-ci exploite les bonnes propriétés de la dynamique de Langevin sur-amortie donnée par,

$$dX_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \quad \text{avec} \quad U = -\log \pi, \tag{1.20}$$

où $(W_t)_{t\geq 0}$ est un mouvement Brownien standard dans $\mathbb{R}^d$. En effet, sous de bonnes propriétés du potentiel U, ce processus de Markov est ergodique par rapport à la mesure cible π. L'algorithme MALA consiste à construire une chaîne de Markov définie comme une discrétisation par un schéma d'Euler-Maruyama, de cette équation. Un telle chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ est construite par récurrence pour $n \in \mathbb{N}$ par :

$$X_{n+1} = X_n - h\nabla(X) + \sqrt{2h}\chi^{n+1},$$

où $(\chi_n)_{n\in\mathbb{N}}$ est une suite i.i.d de variables aléatoires gaussiennes de dimension d centrées et réduites, et où $h > 0$ est le pas de temps. On note Q le noyau de transition de cette chaîne de Markov, c'est à dire pour tout $x \in \mathbb{R}^d$, $Q(x,\cdot) \sim \mathcal{N}(x - h\nabla U(x), 2h)$. Cependant, une telle chaîne de Markov n'est *a priori* pas ergodique par rapport à la mesure π. Elle peut être au mieux ergodique par rapport à une mesure perturbée, et dans certains cas être transiente [123, 152] Pour remédier à cette difficulté, cette chaîne de Markov est enrichie d'une étape d'acceptation/rejet de type Metropolis-Hastings, qui permet finalement de la rendre ergodique par rapport à la mesure π. Pour ce faire, cette étape impose la réversibilité de cette chaîne de Markov par rapport à la mesure π, en imposant une relation de balance détaillée, ce qui est une condition suffisante pour assurer cette propriété d'ergodicité. En résumé, cet algorithme est donné par,

Algorithm 1.7 (MALA). *Soit $h > 0$ et $x^0 \in \mathbb{R}^d$ un point initial.*
Itérer sur $n \geq 0$.

1. *Echantillonner y^{n+1} selon la loi $Q(x^n, dx)$.*
2. *Définir la probabilité d'acceptation $A(x^n, y^{n+1})$ par*

$$A(x^n, y^{n+1}) = 1 \wedge \frac{\pi(y^{n+1})Q(y^{n+1}, x^n)}{\pi(x^n)Q(x^n, y^{n+1})}.$$

3. Avec probabilité $A(x^n, y^{n+1})$, poser $x^{n+1} = y^{n+1}$; sinon poser $x^{n+1} = x^n$.

Dans ce chapitre, nous proposons deux méthodes qui tentent de tirer profit des bonnes propriétés de la dynamique de Langevin sur-amortie donnée par,

$$dX_t = -\nabla U(X_t)\,\mathrm{d}t + \xi\gamma(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{1.21}$$

où $\xi \in \mathbb{R}$ est un paramétrage de l'intensité et du sens de la non-réversibilité, et où le champ de vecteur γ est choisi tel que $\nabla \cdot (\gamma\pi) = 0$. Cette dynamique est également ergodique par rapport à la mesure π, et n'est réversible que si et seulement si $\xi\gamma = 0$. De tels champs peuvent par exemple être construits pour toute matrice antisymétrique J en posant $\gamma(x) = J\nabla \log \pi(x)$. Il est connu que de telles dynamiques non-réversibles convergent plus vite, à la fois en terme de trou spectral et en terme de variance asymptotique, que leur version réversible donnée par l'Équation (1.20) [72, 96, 97, 113, 150, 151, 170].

Le Chapitre 5 propose deux algorithmes qui permettent de construire deux chaînes de Markov non-réversibles, basées sur une discrétisation de l'Équation (1.21). Nous mentionnons [23, 24, 131] pour des travaux différents dans cette direction, ainsi que [43, 66, 94, 165] pour des travaux dans le cas d'un espace d'état discret. Dans ce chapitre nous construisons des chaînes de Markov à l'aide d'une étape d'acceptation/rejet permettant d'assurer l'ergodicité par rapport à la mesure cible π. Bien sûr, cette étape ne peut pas correspondre à un algorithme de Metropolis-Hastings car cela forcerait la réversibilité par rapport à une mesure cible. Toutefois une telle approche a été analysée dans [131].

L'algorithme *Generalized MALA*

Nous présentons maintenant le premier algorithme. Nous désignons par Q^ξ un noyau de densité de probabilité qui correspond à une proposition de déplacement sur un pas de temps. Il est censé approcher le noyau de la densité de probabilité de la chaîne de Markov construite comme un h-squelette de l'Équation (1.21), à l'aide d'une méthode de discrétisation. Le réel ξ correspond au paramétrage de l'intensité et au sens de la non-réversibilité dans l'Équation (1.21). Plusieurs noyaux, correspondant à plusieurs discrétisations, sont d'ailleurs proposés dans ce chapitre, et leurs performances, en terme de taux d'acceptation, sont comparés théoriquement et numériquement. Cet algorithme, appelé *Generalized MALA* est le suivant.

Algorithm 1.8 (**Generalized MALA**, Algorithm 5.6)**.** *Soit $h > 0$ et $(x^0, \xi^0) \in \mathbb{R}^d \times \mathbb{R}$ un point initial et une direction initiale.*
Itérer sur $n \geq 0$.

1. *Echantillonner y^{n+1} selon la loi $Q^{\xi^n}(x^n, dy)$.*
2. *Définir la probabilité d'acceptation $A^{\xi^n}(x^n, y^{n+1})$ par*

$$A^{\xi^n}(x^n, y^{n+1}) = 1 \wedge \frac{\pi(y^{n+1})Q^{-\xi^n}(y^{n+1}, x^n)}{\pi(x^n)Q^{\xi^n}(x^n, y^{n+1})}. \tag{1.22}$$

3. *Avec probabilité* $A^{\xi^n}(x^n, y^{n+1})$, *poser* $(x^{n+1}, \xi^{n+1}) = (y^{n+1}, \xi^n)$; *sinon poser* $(x^{n+1}, \xi^{n+1}) = (x^n, -\xi^n)$.

C'est un algorithme non-biaisé, dans le sens où π est une mesure invariante du processus $(x_n)_{n\in\mathbb{N}}$, et que ce dernier est ergodique par rapport à π. Cette propriété vient du fait qu'il est construit comme un algorithme de Metropolis-Hastings généralisé (voir [114, paragraphe 2.1.4]) sur l'espace d'état augmenté $E = \mathbb{R}^d \times \{-\xi^0, \xi^0\}$. Cette remarque justifie d'ailleurs le nom de l'algorithme. Le Lemme 5.7 montre que la probabilité d'acceptation définie par (1.22) permet de construire une chaîne de Markov $(x_n)_{n\in\mathbb{N}}$ qui conserve une propriété de l'équation continue (1.21). Il s'agit d'une propriété de *skew-detailed balance*, qui est une condition suffisante pour que la chaîne de Markov $(x_n)_{n\in\mathbb{N}}$ soit ergodique par rapport à π, et rende ainsi l'Algorithme 1.8 non-biaisé.

Le choix du noyau de proposition s'est révélé être une question très importante. Il est apparu qu'un noyau basé sur une discrétisation de l'Équation (1.21) par un schéma d'Euler-Maruyama (explicite), que l'on notera Q_1^ξ où ξ paramètre toujours l'intensité et le sens de la non-réversibilité, ne permet pas de bénéficier d'un taux de rejet convenable pour l'étape d'acceptation/rejet de l'Algorithme 1.8. En effet, le taux de rejet moyen pour l'algorithme MALA, initialisé avec la mesure invariante cible, est de l'ordre de $h^{3/2}$, où h désigne le pas de temps, alors qu'il n'est que de l'ordre de h pour l'Algorithme 1.8 basé sur le noyau de proposition Q_1^ξ. Nous proposons de construire un noyau de proposition Q_2^ξ basé sur une discrétisation de type point-milieu pour la partie non-réversible. Cette solution permet de retrouver un taux moyen de rejet d'ordre $h^{3/2}$, au prix d'un schéma implicite. Ces résultats sont donnés par la Proposition 1.9 où l'on note par $\alpha^1_{h,\xi}$ et $\alpha^2_{h,\xi}$ les probabilités d'acceptation des méthodes basées respectivement sur les noyaux Q_1^ξ et Q_2^ξ.

Proposition 1.9 (Proposition 5.12)**.** *Pour tout $l \geq 1$ il existe, sous certaines hypothèses de régularité sur le potentiel U, $C(l) > 0$ et $h_0 > 0$ tels que pour tout pas de temps $h \leq h_0$ strictement positif et pour tout $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left[(1 - \alpha^1_{h,\xi}(x, Y^1_{h,\xi}))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{2l},$$
$$\mathbb{E}\left[(1 - \alpha^2_{h,\xi}(x, Y^2_{h,\xi}))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{3l},$$

où $Y^1_{h,\xi} \sim Q_1^\xi(x, \cdot)$ et $Y^2_{h,\xi} \sim Q_2^\xi(x, \cdot)$.

Nous proposons également un troisième noyau de proposition qui permet de s'affranchir de certaines hypothèses de régularité sur U. Il est à noter que le calcul de la probabilité d'acceptation reste aisé avec une proposition point-milieu pour la non-réversibilité, ce qui n'aurait par exemple pas été le cas avec une discrétisation de type Crank-Nicolson.

Il est également démontré par la Proposition 5.15, sous de fortes conditions de régularité sur le potentiel U, une condition de Lyapunov qui permet d'assurer la convergence géométrique en variation totale de la solution vers la mesure cible π. La particularité de cette

condition de drift vient du fait qu'elle est basée sur le noyau de probabilité $P^2(x, dy)$ de transition de la chaîne de Markov à deux pas, défini par

$$P^2(x, dy) = \int_E P(x, dz)P(z, dy), \quad \forall x \in E.$$

Cette particularité vient du fait qu'il peut être nécessaire d'attendre une étape lors de laquelle la direction de la non-réversibilité est échangée, avant d'accepter une proposition qui contribue à la décroissance de la fonction de Lyapunov que nous envisageons.

L'algorithme *Generalized Hybrid MALA*

Nous proposons dans la suite du chapitre un deuxième algorithme qui peut être vu comme une version sur-amortie des méthodes *Hybrid Monte-Carlo*. L'idée consiste à résoudre alternativement la dynamique de Langevin réversible donnée par,

$$dX_t = -\nabla U(x_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{1.23}$$

et la dynamique purement non-réversible donnée par,

$$dX_t = -\xi J \nabla U(x_t)\,\mathrm{d}t. \tag{1.24}$$

La première équation est résolue par l'algorithme MALA, tandis que la deuxième est résolue par une méthode hybride. Plus précisément, nous introduisons Φ_h^ξ le flot numérique d'un intégrateur de l'Équation (1.24) sur un pas de temps h. Pour assurer que l'algorithme Generalized Hybrid MALA est non biaisé, nous supposons que l'intégrateur numérique vérifie les deux propriétés suivantes, classiques pour les méthodes de Monte-Carlo hybrides :

$$\Phi_h^\xi = (\Phi_h^{-\xi})^{-1}, \quad \det\left(\mathrm{Jac}\,\Phi_h^\xi\right) = 1.$$

Ces hypothèses sont raisonnables et sont par exemple satisfaites par le schéma point-milieu. L'algorithme est donné ci-après et le Lemme 5.20 assure que la deuxième étape laisse la mesure π invariante.

Algorithm 1.10 (**Generalized Hybrid MALA**, Algorithm 5.19)**.** *Soit x_0 un point initial. On choisit $\xi_0 \in \{-1, 1\}$ et $h > 0$. Itérer sur $n \geq 0$,*

1. *Intégration de la dynamique de Langevin réversible (1.23) :*
   utiliser MALA pour échantilloner $x_{n+1/2}$ à partir de x_n, avec un pas de temps h.
2. *Intégration de la dynamique purement non-réversible (1.24) :*
   (a) *Calculer $\widetilde{x}_{n+1} = \Phi_h^{\xi_n}(x_{n+1/2})$.*

(b) Poser $x_{n+1} = \widetilde{x}_{n+1}$ *et* $\xi_{n+1} = \xi_n$ *avec probabilité*

$$\begin{aligned}\beta_{h,\xi_n}(x_{n+1/2}) &= \min\left(1, \exp(U(x_{n+1/2}) - U(\widetilde{x}_{n+1}))\right) \\ &= \min\left(1, \exp(U(x_{n+1/2}) - U(\Phi_h^{\xi_n}(x_{n+1/2})))\right).\end{aligned}$$

Sinon poser $x_{n+1} = x_{n+1/2}$ *et* $\xi_{n+1} = -\xi_n$.

L'intérêt de cet algorithme réside dans le fait qu'il permet de découpler les taux de rejection venant de la résolution des dynamiques (1.23) et (1.24). En pratique, on souhaite obtenir un taux de rejection pour la deuxième étape de l'Algorithme 5.19 très faible car chaque rejection est accompagnée d'une inversion du sens de la dynamique purement non-réversible, qui est susceptible de ralentir la convergence de l'algorithme (dans le sens d'une augmentation de la variance asymptotique des estimateurs ergodiques). Cependant, le taux de rejection optimal pour l'algorithme MALA peut prendre des valeurs beaucoup plus grandes. L'avantage de cet algorithme sur GMALA est qu'il permet donc d'accélérer la dynamique de Langevin réversible, sans modifier le taux de changement de direction de la non-réversibilité. Par ailleurs, le Lemme 5.22 assure que le taux moyen de rejet de la deuxième étape de l'algorithme GHMALA est d'ordre h^3, au lieu de $h^{3/2}$ pour MALA et GMALA, sous des hypothèses de régularité sur U.

Une autre manière de voir l'intérêt du splitting entre les dynamiques (1.23) et (1.24) est de remarquer que le terme d'ordre principal en la variable h dans la probabilité de rejet $(1 - \alpha_{h,\xi}^2)$ provient de la partie réversible de la dynamique de Langevin. Cela signifie qu'essentiellement l'algorithme GMALA force à changer la direction de la non-réversibilité pour corriger une erreur commise sur l'intégration de la partie réversible, ce qui n'est pas naturel. C'est ce comportement que permet de corriger GHMALA.

Simulations numériques

Ces deux algorithmes sont testés sur trois exemples jouets de petite dimension au Paragraphe 5.5. D'abord, nous illustrons les taux moyens de rejet théoriques, et mettons en évidence les limitations d'une proposition explicite pour l'algorithme GMALA. Ensuite, nous nous intéressons particulièrement à la réduction de variance asymptotique offerte par GMALA et GHMALA. Nous observons des réductions de variance de plusieurs ordres de grandeur par rapport à l'algorithme MALA.

1.3.2 Résultats du Chapitre 6

Objectifs du chapitre

Le but du Chapitre 6 est de proposer et d'étudier des méthodes numériques permettant d'analyser un éventuel comportement métastable d'un condensat de Bose-Einstein en rotation, à des températures non-nulles. Cette dynamique métastable est modélisée par l'équation de Gross-Pitaevskii projetée stochastique (1.17). Sous l'effet de la rotation,

plusieurs minima locaux de l'énergie du système (donnée par (1.18)) peuvent exister, à invariance par rotation et changement de phase près. Ces minima sont charactérisés par différents arrangements de vortex dont le nombre peut varier. La Figure 1.1 [Figure 6.3] représente un exemple de trois minima locaux de l'énergie pour l'ensemble de paramètres numériques considéré dans le Chapitre 6. Dans ce cas, la dynamique de la solution de

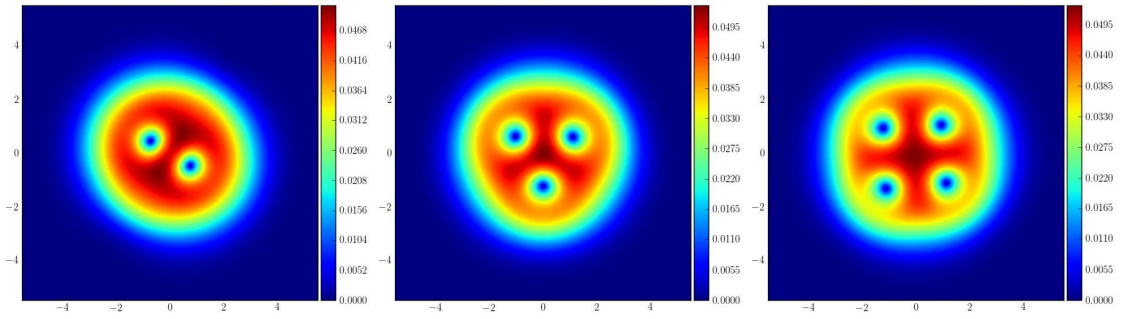

FIGURE 1.1 – Carré du module des minima locaux de l'énergie de Gross-Pitaevskii (1.18), pour l'exemple considéré dans le Chapitre 6.

l'Équation (1.17) contient deux échelles temporelles distinctes : une échelle de fluctuations rapides autour des minima de l'énergie, et une échelle beaucoup plus lente où le système "saute" d'un minimum local à un autre. C'est cette échelle de temps qui caractérise le comportement métastable de la dynamique.

Dans ce chapitre, nous nous sommes intéressés à la construction d'une méthode numérique dont l'objectif est de calculer le temps moyen passé dans un minimum local, avant de "sauter" dans un autre minimum. Pour cela, il est nécessaire d'une part d'être capable de résoudre la dynamique (1.17) précisément, et d'autre part de construire une méthode de Monte-Carlo adaptée à la simulation de ces événements rares de changement de configurations de vortex. La valeur ajoutée de ce chapitre réside principalement dans l'expérimentation numérique de méthodes de Monte-Carlo pour la simulation de dynamiques métastables (initialement développées dans le cadre de la dynamique moléculaire) dans le cadre de la condensation de Bose-Einstein. Ce dernier chapitre soulève de nombreuses questions théoriques sur le comportement asymptotique de la dynamique dans la limite où la température tend vers 0, ainsi que sur les justifications rigoureuses des algorithmes et des approximations numériques utilisées.

Construction d'un schéma numérique

Nous proposons dans le Chapitre 6 un schéma numérique pour l'approximation de l'Équation (1.17). Tout d'abord, il est à noter que, de part la restriction de $L^2(\mathbb{R}^d)$ au sous-espace vectoriel K, cette équation prend la forme d'une équation différentielle stochastique (EDS) classique, et non plus d'une équation aux dérivées partielles stochastique (EDPS). Il n'y a donc pas de problématique liée à une discrétisation en espace, mais seulement en temps. La difficulté essentielle vient du fait que le coefficient c_1 apparaissant dans

l'Équation (1.17) est très petit devant 1. Il est typiquement de l'ordre de $10^{-3} \sim 10^{-5}$ pour les expériences menées en laboratoire. De ce fait, deux échelles temporelles se distinguent. D'une part la partie purement Hamiltonienne de la dynamique est donnée par,

$$d\phi_t = -i\nabla E^K(\phi_t))\,\mathrm{d}t, \tag{1.25}$$

alors que la partie correspondant à une dynamique de Langevin réversible est donnée par,

$$d\phi_t = -c_1\nabla E^K(\phi_t))\,\mathrm{d}t + c_1c_3dB_t. \tag{1.26}$$

La partie Hamiltonienne est plus rapide (d'un facteur c_1^{-1}) que la partie correspondant à la dynamique de Langevin réversible. En pratique, cela implique qu'il sera nécessaire de résoudre la dynamique complète (1.17) sur des temps longs pour que l'effet de la dynamique de Langevin (qui agit en quelque sorte comme une correction de l'équation Hamiltonienne) soit significatif. Par ailleurs, nous savons construire des schémas adaptés aux dynamiques Hamiltoniennes et d'autres adaptés aux dynamiques de Langevin. Ces deux remarques justifient l'utilisation d'une méthode de splitting qui consiste à résoudre alternativement, et avec des schémas adaptés, ces deux dynamiques.

La discrétisation de la dynamique de Langevin réversible (1.26) est effectuée à l'aide d'un schéma explicite exponentiel, qui correspond à une discrétisation explicite d'une formulation mild de l'Équation (1.26). Cette méthode, classique pour des EDPS paraboliques, a deux avantages sur le schéma d'Euler-Maruyama. D'une part, il est en pratique plus stable que ce dernier (car le terme de dérive est sur-linéaire), et d'autre part, comme la dimension de l'espace K peut devenir arbitrairement grande, il est naturel de se tourner vers une méthode adaptée à la dimension infinie. Notons toutefois, qu'il ne serait pas possible de donner un sens classique à l'Équation (1.17) sans la restriction à l'espace K (autrement dit dans le cas $K = L^2(\mathbb{R}^d)$).

La discrétisation de la dynamique Hamiltonienne (1.25) repose sur une méthode de Gauss-Lawson symplectique d'ordre élevé. Elle est tirée de [22], et elle est adaptée au cas de la dimension infinie pour la même raison que précédemment. L'idée consiste à effectuer un changement de variable tel que l'équation décrivant l'évolution temporelle de cette nouvelle variable perde la raideur liée à la dimension. C'est cette nouvelle équation qui est discrétisée à l'aide d'une méthode de Runge-Kutta implicite (et symmétrique) d'ordre élevé, qui rende le flot numérique symplectique. Plus précisément, si l'on introduit S le semi-groupe associé à l'opérateur linéaire iA, où A est donné par (1.19), et que l'on introduit la variable $u(t, \mathbf{x})$ définie par,

$$u(t, \mathbf{x}) = S(t, 0)^{-1}\phi(t, \mathbf{x}),$$

où ϕ est la solution de l'Équation (1.25). Alors, u est solution de l'équation,

$$du(t) = S(t, 0)^{-1}\mathcal{N}(S(t, 0)u(t))\,\mathrm{d}t,$$

où $\mathcal{N}$ correspond à la partie non-linéaire de l'Équation (1.25). On vérifie donc bien que l'opérateur linéaire qui conférait sa raideur à l'Équation (1.25) n'apparaît plus explicitement. Un résultat d'existence et de convergence de ce schéma implicite est donné par le Théorème 6.2 (qui n'est autre que l'analogue du [22, Théorème 14]).

Méthode de Monte-Carlo pour les dynamiques métastables

La seconde partie du Chapitre 6 est dédiée à la mise en oeuvre de méthodes numériques pour l'analyse des dynamiques métastables de la solution de l'Équation (1.17). Nous sommes en particulier intéressés par l'estimation du temps moyen de sortie d'un état métastable. Une méthode naïve serait de simuler une trajectoire en temps long de ce processus stochastique, puis d'estimer les temps de sortie moyens par le biais d'une moyenne empirique des temps de sortie réalisés. Quand la température du système s'approche de zéro, cette méthode devient prohibitive en terme de temps de calcul.

Nous utilisons dans la suite une méthode de Monte-Carlo adaptée à cette question dont le principe a été introduit dans [40]. Celle-ci repose essentiellement sur l'algorithme *Adaptive Multilevel Splitting* (AMS, [33, 39]) qui permet de construire un estimateur pour une probabilité que nous décrivons formellement dans la suite. On s'intéresse à un processus stochastique $X = (X_t)_{t\geq 0}$ à valeur dans un ensemble E, supposé posséder une dynamique métastable. Pour mettre en évidence le point initial de cette dynamique, nous notons $X^x = (X_t^x)_{t\geq 0}$ où l'exposant x correspond au point initial ($X_0^x = x$). On notera également $X^\nu = (X_t^\nu)_{t\geq 0}$ dans le cas où la condition initiale est une variable aléatoire distribuée suivant une mesure de probabilité ν. Nous introduisons deux ensembles disjoints A et B, qui sont censés correspondre à deux sous-ensembles métastables inclus dans deux bassins d'attraction distincts. Nous considérons aussi x_A un point n'appartenant pas à $A \cup B$. Il est censé être choisi dans le même bassin d'attraction que A et proche de cet ensemble. Dans ce contexte, il est alors plus probable que la dynamique X^{x_A} atteigne A avant B, et l'événement correspondant au fait que l'ensemble B soit atteint avant A est un événement rare. En notant $\tau_A(X^{x_A})$ et $\tau_B(X^{x_A})$ respectivement le premiers temps d'atteinte de l'ensemble A et de l'ensemble B pour le processus X^{x_A}, on a donc $\mathbb{P}(\tau_B(X^{x_A}) < \tau_A(X^{x_A})) \ll 1$. C'est justement cette faible probabilité que permet d'estimer efficacement, et sans biais, l'algorithme AMS. Dans notre cas, la dynamique étudiée est bien entendue donnée par l'équation (1.17), et les ensembles A et B correspondent chacun à deux ensembles (métastables) contenant chacun un minimum local différent de l'énergie. Plus précisément, ils sont choisis comme des composantes connexes d'ensembles de niveaux de l'énergie qui ne contiennent qu'un minimum local.

Le but de l'algorithme AMS est de construire un ensemble pondéré de trajectoires dont la mesure empirique pondérée approche la loi des trajectoires initialisées en x, et arrêtées en $\tau_A \wedge \tau_B$. Il peut en ce sens être vu comme une méthode d'échantillonnage d'importance car il échantillonne surtout des trajectoires peu probables (qui seront donc faiblement pondérées) qui sont arrêtées en B avant d'être arrêtées en A, et qui sont donc nos trajectoires d'intérêt. Ces dernières sont appelées *trajectoires réactives*. C'est cette

description fine de la "queue de la distribution" des trajectoires initialisées en x et arrêtées en $\tau_A \wedge \tau_B$, qui permet de pouvoir construire des estimateurs de $\mathbb{P}(\tau_B(X^{x_A}) < \tau_A(X^{x_A}))$ de plus faible variance qu'un estimateur naïf.

Cet algorithme est détaillé dans le Chapitre 6. Il construit un ensemble de répliques par un processus de type sélection-mutation. L'idée consiste à itérer un processus où les trajectoires les plus proches de l'ensemble des trajectoires réactives sont dupliquées et mutées. Ainsi, les trajectoires mutent au fur et à mesure en des trajectoires de plus en plus proches des trajectoires réactives, jusqu'à le devenir. L'algorithme s'arrête quand un nombre suffisant de répliques sont devenues réactives. La difficulté principale de mise en oeuvre de cet algorithme vient de la manière dont on caractérise les trajectoires les plus proches de l'ensemble des trajectoires réactives. La sélection des "meilleures" trajectoires se fait à l'aide d'une fonction appelée *coordonnée de réaction* et notée ξ, associant à tout élément de l'espace d'état E un réel. C'est elle qui encode dans quelle mesure un point y de l'espace d'état a des chances d'évoluer, par la dynamique X^y, en une trajectoire réactive. Un exemple idéal est de définir $\xi(x) = \mathbb{P}(\tau_B(X^x) < \tau_A(X^x))$. Cependant, cette fonction n'est pas connue, et c'est même ce que l'on cherche à calculer. On utilise ainsi en pratique une fonction arbitraire, mais censée décrire assez bien l'évolution de la transition de A vers B. Le choix de cette fonction peut beaucoup influencer la variance de l'estimateur construit par l'algorithme AMS, et donc la performance de l'algorithme (voir les illustrations numériques de [33]). Dans notre cas d'application, nous sommes intéressés par la sortie d'un vortex du centre du condensat. En particulier, nous nous sommes intéressés au passage de 3 à 2 vortex dans le condensat. Nous proposons d'utiliser comme coordonnée de réaction la distance entre le troisième vortex le plus proche du centre du condensat et le centre du condensat. L'intérêt premier de cette coordonnée de réaction est qu'elle permet de séparer les deux états métastables en question. C'est-à-dire que $\sup_{x\in A} \xi(x) < \inf_{x\in B} \xi(x)$. Un deuxième intérêt de cette coordonnée est qu'elle est inchangée par un changement de phase uniforme, et par une rotation du condensat, qui sont deux transformations invariantes pour l'énergie. Elles ne devraient donc pas influer sur l'avancement de la réaction.

Le calcul du temps moyen de transition du bassin d'attraction de l'ensemble A à celui de B, noté $T_{A\to B}$ est basé sur le fait que l'on peut définir une constante α et une mesure ν telles que,

$$T_{A\to B} = \alpha \mathbb{P}(\tau_B(X^\nu) < \tau_A(X^\nu))^{-1}$$

où la constante α peut être estimée par une méthode de Monte-Carlo classique. En pratique la mesure ν est approchée par une mesure simulable ν', et nous estimons alors $\mathbb{P}(\tau_B(X^{\nu'}) < \tau_A(X^{\nu'}))$ à l'aide d'une moyenne empirique d'estimateurs de type AMS initialisés suivant la mesure ν'. La définition de cette mesure est donnée au Paragraphe 6.4.3.

Nous présentons dans le Chapitre 6 les estimations de $T_{A\to B}$ pour différentes valeurs des paramètres c_1 et c_3 apparaissant dans l'Équation (1.17). Cela correspond à différentes

températures et différentes intensités de la dissipation. Ce calcul a nécessité 30 jours de calcul sur 128 processeurs, et les résultats sont présentés en Figure 6.5.

Nous précisons que les paramètres numériques de cette expérience ont été choisis pour minimiser le temps de calcul qui demeure néanmoins très important. Notamment, l'intensité de la dissipation est plus grande que les valeurs typiques dérivées théoriquement. Une des limitations de cette méthode est la variance importante des estimateurs AMS utilisés. Cependant nous avons observé que celle-ci provient principalement de l'aléa sur la condition initiale de ces estimateurs qui est échantillonnée suivant la loi ν (suivant ν' en pratique). Autrement dit, la variable aléatoire $\mathbb{E}\left[\mathbb{P}(\tau_B(X^\nu) < \tau_A(X^\nu))|X_0^\nu\right]$ possède une grande variance relativement à son espérance. De plus, nous observons que la variance de cette variable aléatoire dépend fortement de l'intensité de la non-réversibilité. Ce résultat est présenté en Figure 6.6. On constate notamment dans notre cas qu'un peu de non-réversibilité permet de réduire la variance de l'estimateur de $T_{A\to B}$, mais que des valeurs de c_1 trop petites mènent à un accroissement de cette variance.

1.3.3 Perspectives

Dans le cadre du Chapitre 5, la question la plus naturelle qui se pose est de savoir s'il serait possible de se passer des schémas implicites utilisés par les algorithmes qui y sont présentés. Une piste éventuelle serait de s'intéresser à une formulation similaire à [75, Équation (24)].

Une ouverture pertinente sur ce chapitre serait d'étudier l'influence de la dimension sur les algorithmes décrits dans ce chapitre. Par exemple, calculer la décroissance des pas de temps associés aux deux étapes de l'algorithme GHMALA avec la croissance de la dimension, à probabilités moyennes d'acceptations fixées pourraient permettre de comprendre comment cet algorithme se comporte en grande dimension. Par exemple si le pas de temps associé à l'étape MALA décroît plus vite que le pas de temps associé à l'étape hybride, alors on pourrait espérer profiter d'autant plus de la réduction de variance asymptotique offerte dans le cas continu par la non-réversibilité.

Dans le cadre du Chapitre 6, nous aimerions justifier, dans de futurs travaux, l'approximation temporelle réalisée en séparant les dynamiques (1.25) et (1.26). Nous aimerions démontrer un résultat de convergence, uniforme en la constante c_1, pour cette approximation en temps uniquement (c'est-à-dire en mettant de côté l'erreur d'approximation numérique des dynamiques (1.25) et (1.26)).

Par ailleurs, la question du biais introduit par l'approximation de la loi ν par une loi ν' reste ouverte. Il pourrait aussi être intéressant d'étudier la possibilité d'échantillonner de façon non-biaisée la véritable loi ν en échantillonnant à la fois les trajectoires réactives qui vont de l'ensemble A vers l'ensemble B, et celles qui partent de B pour arriver dans A.

Chapitre 2

Physique de la condensation de Bose-Einstein

L'objectif de ce chapitre est de présenter une introduction simplifée au phénomène de la condensation de Bose-Einstein (en Paragraphe 2.1 et 2.2) ainsi qu'aux procédés expérimentaux de réalisation de ces systèmes (en Paragraphe 2.3). Toutefois, le contenu mathématique des chapitres suivants ne requiert pas la lecture de ce chapitre.

2.1 La physique de la condensation de Bose-Einstein

2.1.1 Un peu d'histoire

La prédiction du phénomène de condensation de Bose-Einstein remonte au début du vingtième siècle. En 1924, le physicien Satyendra Nath Bose s'intéressa aux statistiques quantiques des photons et proposa à Albert Einstein une dérivation de la loi de Planck, sans avoir recours à la mécanique classique. Einstein traduisit son travail en Allemand, et le publia [26]. Ce dernier étendit par la suite les résultats de Bose au cas d'un gaz parfait de particules identiques. C'est ainsi qu'il dériva la statistique que l'on nomme aujourd'hui statistique de Bose-Einstein.

Cette statistique prédit qu'un gaz de bosons subit une transition de phase sous certaines conditions de température (extrêmement basses) et de pression. C'est cet état de la matière que l'on appelle condensat de Bose-Einstein. C'est seulement en 1995 que les premiers condensats de Rubidium et de Sodium ont été réalisés [6, 50]. Cette mise en évidence expérimentale a marqué un tournant de la physique des atomes froids. Les recherches sur ce sujet ont commencé une quinzaine d'années auparavant, en particulier sur des atomes d'hydrogène. C'est à la fin des années 1980 que les techniques de refroidissement laser et de piégeage optico-magnétique, que nous présenterons dans la suite, ont été développées [44]. Ces méthodes se sont avérées particulièrement adaptées aux atomes alcalins. C'est finalement en combinant ces techniques de refroidissement laser avec des techniques de refroidissement par évaporation, dont nous parlerons aussi par la suite, que

ces expériences ont permis d'atteindre la condensation de Bose-Einstein.

2.1.2 La statistique de Bose-Einstein

On propose de dériver la statistique de Bose-Einstein dans cette partie. Celle-ci décrit pour un gaz parfait de bosons, le nombre moyen d'atomes $n_{T,\mu}(j)$ par état j d'énergie ε_j, pour le potentiel chimique μ et une température T fixés. On suppose que les niveaux d'énergie sont ordonnés pour que la suite $(\varepsilon_j)_{j\in N}$ soit croissante. On rappelle d'ailleurs qu'il suffit que le potentiel de confinement soit infini à l'infini pour que ce spectre d'énergie soit discret. Comme il est possible qu'un niveau d'énergie soit composé de plusieurs états, on introduit le nombre moyen d'atomes $n_{T,\mu}(\varepsilon_j)$ par niveau d'énergie donné par,

$$n_{T,\mu}(\varepsilon_j) = g_j n_{T,\mu}(j),$$

où g_j est la dégénérescence du niveau d'énergie ε_j. La statistique est alors donnée par,

$$n_{T,\mu}(j) = \frac{n_{T,\mu}(\varepsilon_j)}{g_j} = \frac{1}{\mathrm{e}^{(\varepsilon_j-\mu)/k_BT}-1}. \tag{2.1}$$

Tout d'abord, on peut noter que nécessairement $\varepsilon_0 > \mu$ pour que $n_{T,\mu}(0)$ soit bien défini et positif. Dans ce cas on a alors pour tout $j \in \mathbb{N}$, $n_{T,\mu}(j) > 0$. Avant d'expliquer comment cette statistique décrit le phénomène de condensation de Bose-Einstein, nous présentons maintenant deux dérivations de celle-ci.

Cette statistique est obtenue classiquement dans le formalisme de la physique statistique. Deux dérivations simples peuvent être obtenues dans l'ensemble grand-canonique, et dans l'ensemble microcanonique. Pour rappel, les ensembles statistiques sont des abstractions centrales en physique statistique. Ce sont des collections de copies d'un système physique qui décrivent l'ensemble des états possibles du système, sous des contraintes extérieures fixées, telles que la température, le nombre de particule, le volume ou encore l'énergie. L'ensemble grand-canonique correspond au cas d'un système en contact avec un *réservoir* (un ensemble de taille beaucoup plus grande) avec lequel il peut échanger de l'énergie et des particules. On suppose dans ce cas que le volume, la température et le potentiel chimique sont fixés. L'ensemble microcanonique correspond au cas d'un système isolé. Dans ce cas, le nombre de particules, le volume et l'énergie sont fixés.

Dérivation dans l'ensemble grand-canonique

On suit la dérivation de [67]. Puisque la statistique donnée par (2.1) fait intervenir le potentiel chimique et la température, le cadre naturel de sa dérivation semble donc être l'ensemble grand canonique. En effet, le raisonnement est particulièrement abordable dans ce cadre. Pour obtenir cette statistique, nous considérons un gaz parfait de bosons de potentiel chimique μ et de température T. Les particules n'étant pas en interaction, chaque sous-ensemble de particules d'énergie ε forme un système thermodynamique en interaction

avec le réservoir. L'ensemble grand-canonique est donc d'une certaine manière une tensorisation d'ensembles grand-canoniques pour chaque niveau d'énergie. Comme chaque niveau d'énergie peut contenir tout nombre de bosons, tous les états du système possibles sont les configurations avec n particules (indiscernables) dans l'état énergie ε, donc d'énergie $n\varepsilon$. Ainsi, la probabilité que le système soit composé de n particules est $\mathrm{e}^{n(\mu-\varepsilon)/k_BT}$, et la fonction de partition de l'ensemble grand-canonique pour le niveau d'énergie ε est donné par,

$$\mathcal{Z}(\varepsilon) = \sum_{n\in\mathbb{N}} \mathrm{e}^{n(\mu-\varepsilon)/k_BT} = \frac{1}{1-\mathrm{e}^{(\mu-\varepsilon)/k_BT}}.$$

On peut alors calculer le nombre moyen de particules dans l'état d'énergie ε,

$$n_{T,\mu}(\varepsilon_j) = k_BT\frac{1}{\mathcal{Z}}\left(\frac{\partial\mathcal{Z}}{\partial\mu}\right)_T = \frac{1}{\mathrm{e}^{(\varepsilon_j-\mu)/k_BT}-1}.$$

Dérivation dans l'ensemble microcanonique

On détaille une seconde dérivation approchée dans l'ensemble canonique. Ce calcul fut entrepris, à la suite des travaux de Bose, par Einstein en 1924-1925. Dans ce cadre ce sont la température et le nombre d'atomes qui sont fixés. Bien que ce cas soit moins naturel, que les calculs soient plus pénibles et que le résultat ne soit valide que dans la limite d'un fort taux de dégénérescence, le calcul est toutefois très intéressant car il fait intervenir le dénombrement des états d'énergie et permet de mieux comprendre la différence avec la statistique de Maxwell-Boltzmann. Comme on se place dans l'ensemble canonique, c'est cette fois-ci le nombre de particules N et l'énergie du système E qui sont fixés. Supposons que le niveau d'énergie ε_j comporte n_j atomes. On peut compter le nombre de façons dont ces atomes se répartissent sur les états d'énergie ε_j. On le note $w(n_j, g_j)$, et le calcul (mené dans [45]) donne,

$$w(n_j, g_j) = \frac{(n_j+g_j-1)!}{n_j!(g_j-1)!}.$$

Ainsi, le nombre de configurations pour le système entier s'écrit,

$$W((n_j),(g_j)) = \prod_{j\in\mathbb{N}} \frac{(n_j+g_j-1)!}{n_j!(g_j-1)!}.$$

De part le choix de l'ensemble statistique, les contraintes du système s'écrivent

$$\sum_{j\in[\![0,n]\!]} n_j = N \quad \text{et} \quad \sum_{j\in[\![0,n]\!]} n_j\varepsilon_j = E.$$

Il s'agit alors de maximiser (sous ces contraintes) le nombre de façons de répartir ces N particules pour obtenir la statistique de Bose-Einstein. En pratique, au lieu de maximi-

ser $W((n_j),(g_j))$, on maximise son logarithme à l'aide d'une approche Lagrangienne. On introduit α et β les multiplicateurs de Lagrange, et on maximise

$$\mathcal{L}((n_j),\alpha,\beta) = \log W((n_j),(g_j)) + \alpha\left(N - \sum_j n_j\right) + \beta\left(E - \sum_j n_j\varepsilon_j\right).$$

En utilisant la formule de Stirling, on obtient

$$\log w(n_j,g_j) = (n_j+g_j)\log(n_j+g_j) - n_j\log(n_j) + O(\ln(n_j+g_j)).$$

En négligeant le reste, $O(\ln(n_j+g_j))$, et en résolvant le problème d'optimisation simplifié, on obtient,

$$n_j = \frac{g_j}{\mathrm{e}^{\alpha+\beta\varepsilon_j}-1}.$$

On définit alors le potentiel chimique μ et la température T tels que $\alpha = -\dfrac{\mu}{k_BT}$ et $\beta = \dfrac{1}{k_BT}$.

Le lecteur pourra se convaincre que l'approximation faite par l'utilisation de la formule de Stirling ainsi que par la simplification du Lagrangien est justifiée dans la limite où g_j et n_j sont tous deux grands devant 1. Dans la limite où seulement g_j est grand, un calcul plus fin permettrait toujours de démontrer une décroissance exponentielle de n_j/g_j en le niveau d'énergie ε_j, ce qui est suffisant pour justifier le phénomène de condensation de Bose-Einstein.

L'aspect grégaire des bosons indiscernables

Considérons N particules qui peuvent se trouver dans deux états A ou B. Si l'on suppose les particules indiscernables, et que toutes les éventualités sont équiprobables, alors la probabilité que toutes les particules se trouvent dans l'état A est $p_1 = 1/(N+1)$. Si l'on considère maintenant les particules discernables et qu'à nouveau toutes les éventualités sont équiprobables, alors la probabilité que toutes les particules se trouvent dans l'état A est $p_2 = 1/2^N$. Ainsi pour un grand nombre de particules $p_1 \gg p_2$.

Ce petit exemple illustre bien que la statistique de Bose-Einstein (qui suppose les bosons indiscernables) entraîne une sorte de perte d'indépendance entre les particules par rapport à la statistique de Boltzmann. Cette dépendance est en fait l'analogue du fait qu'il faille restreindre l'espace de Hilbert aux fonctions d'onde symétriques par échange de deux particules.

En effet, un système à N bosons indiscernables peut être décrit par une fonction d'onde

$$\psi(t;x_1,\ldots,x_N),$$

où $x_k \in \mathbb{R}^3$ correspond à la variable d'espace pour la k-ème particule. De part l'hypothèse

d'indiscernabilité, une telle fonction d'onde ψ doit être inchangée par échange entre deux particules. Ainsi pour toute permutation σ de $[\![1, N]\!]$,

$$\psi(t; x_1, \ldots, x_n) = \psi(t; x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

De telles fonctions d'onde appartiennent donc à l'espace

$$L^2(\mathbb{R}^3, \mathbb{C}) \otimes_S \ldots \otimes_S L^2(\mathbb{R}^3, \mathbb{C}),$$

constitué des fonctions symétriques de l'espace (2.2) (voir [35]), et non pas à l'espace

$$L^2(\mathbb{R}^3, \mathbb{C}) \otimes \ldots \otimes L^2(\mathbb{R}^3, \mathbb{C}) \tag{2.2}$$

tout entier.

Il est possible de déduire de cette restriction aux fonctions d'ondes complètement symétriques le phénomène d'émission stimulée. Celui-ci consiste en le fait que la probabilité pour qu'une particule dans un état ϕ_k subisse une transition vers un état ϕ_l est multipliée par $N+1$ si cet état est déjà occupé par N particules. Il est alors possible de retrouver la statistique de Bose-Einstein à partir de bilan d'énergies lors de collisions entre particules. Voir [49].

2.1.3 Le phénomène de condensation de Bose-Einstein pour un gaz parfait

Le phénomène de condensation apparaît quand un grand nombre de particules est distribué suivant la statistique (2.1). Ce nombre total N de particules est donné, en fonction de T et μ par,

$$N = \sum_{j \in \mathbb{N}} n_{T,\mu}(j) = \sum_{j \in \mathbb{N}} \frac{1}{\mathrm{e}^{(\varepsilon_j - \mu)/k_B T} - 1}.$$

Comme précisé plus haut, il est nécessaire que le potentiel chimique soit plus petit que le plus petit niveau d'énergie ($\mu < \varepsilon_0$) pour que la densité de particules soit bien définie et positive pour chaque niveau d'énergie. Pour distinguer les particules réparties sur le niveau d'énergie fondamental (supposé non-dégénéré), et celles sur les autre niveaux, on note N_0 le nombre de particules sur ce premier niveau, et N_{exc} le nombre total de particules excitées sur les autres niveaux. On a alors,

$$N = N_0 + N_{\text{exc}}, \quad N_0 = \frac{1}{\mathrm{e}^{(\varepsilon_0 - \mu)/k_B T} - 1}, \quad N_{\text{exc}} = \sum_{j \neq 0} \frac{1}{\mathrm{e}^{(\varepsilon_j - \mu)/k_B T} - 1}.$$

On peut alors majorer N_{exc} en majorant le potentiel chimique μ par ε_0,

$$N_{\text{exc}} \leq N_{\text{exc}}^{\max}(T) = \sum_{j\neq 0} \frac{1}{\mathrm{e}^{(\varepsilon_j-\varepsilon_0)/k_BT}-1}. \tag{2.3}$$

Ainsi, lorsque le nombre de particules d'un système devient grand devant $N_{\text{exc}}^{\max}(T)$, il devient nécessaire qu'un nombre macroscopique d'entre elles s'accumulent dans l'état fondamental, qui lui n'est pas borné en terme de nombre de particules qu'il peut accueillir. C'est cette remarque qui justifie le phénomène de condensation de Bose-Einstein.

Notons que ce phénomène est bien spécifique à la statistique de Bose-Einstein. Dans le cadre de la statistique de Boltzmann ($n_{T,\mu}^{Boltz.}(j) = \mathrm{e}^{-(\varepsilon_j-\mu)/k_BT}$), une accumulation de particules dans le premier niveau peut également survenir, lorsque la température est suffisamment faible ($k_BT \ll \varepsilon_1 - \varepsilon_0$). Or, dans le cas de la statistique de Bose-Einstein, le phénomène est bien de nature différente car il peut intervenir pour toute température.

2.1.4 Le passage à la limite thermodynamique, sans interactions

Le but de cette partie est de vérifier si l'effet décrit dans la partie précédente est uniquement un effet de taille finie, ou s'il persiste à la limite thermodynamique. Celle-ci est définie comme la limite où le nombre de particules tend vers l'infini, mais à densité volumique constante. En effet, les niveaux d'énergies ε_j dépendent de la taille caractéristique L du potentiel de confinement. Plus cette grandeur sera grande, et plus les niveaux d'énergie seront proches, et donc plus le nombre d'atomes N_{exc} sera grand. La question est donc de savoir comment grandit $N_{\text{exc}}^{\max}/L^d$ dans le cas d-dimensionnel. Si ce nombre n'est pas borné dans la limite thermodynamique, alors le phénomène d'accumulation disparaîtra dans cette limite.

Nous suivons ici l'approche de [138]. Pour mener ces calculs, la technique consiste à approcher la sommation discrète apparaissant dans (2.3) par une sommation continue. Cette procédure donne une bonne approximation du nombre d'atomes que peuvent contenir les états excités, dans la limite thermodynamique. Notons alors $g(\varepsilon)$ la densité d'états excités d'énergie ε. Nous traiterons dans la suite le cas d'une particule libre confinée dans une boîte, et celui d'une particule confinée dans un un piège harmonique. Nous verrons (Équation (2.6) et (2.8)) que dans ces deux cas, cette densité g est donnée par

$$g(\varepsilon) = C_\alpha \varepsilon^{\alpha-1}, \tag{2.4}$$

où α dépend de la dimension, et C_α est une constante. Nous verrons aussi (Équation (2.7) et (2.9)) que dans ces deux cas le passage à la limite thermodynamique survient lorsque $C_\alpha^{-1} N_{\text{exc}}$ reste borné car cette quantité est proportionnelle à la densité de particules. Le nombre de particules $N_{\text{exc}}(T,\mu)$ dans les états excités est alors approché par,

$$N_{\text{exc}}(T,\mu) = \int_{\varepsilon_0-\mu}^{+\infty} g(\varepsilon) \frac{1}{\mathrm{e}^{\varepsilon/k_BT}-1} d\varepsilon,$$

d'où,

$$C_\alpha^{-1} N_{\text{exc}}(T,\mu) = \int_{\varepsilon_0-\mu}^{+\infty} \varepsilon^{\alpha-1} \frac{1}{\mathrm{e}^{\varepsilon/k_BT}-1} d\varepsilon.$$

On obtient alors que la condensation est préservée par le passage à la limite si et seulement si cette intégrale converge quand μ tend vers ε_0. Cela est le cas si et seulement si $\alpha > 1$. Dans ce cas, le nombre de particules maximale $N_{\text{exc}}^{\max}(T)$ que peuvent contenir les états excités à la température T est donné par,

$$\begin{aligned} C_\alpha^{-1} N_{\text{exc}}(T) &= (k_BT)^\alpha \int_0^{+\infty} \frac{x^{\alpha-1}}{\mathrm{e}^x-1} dx, \\ &= (k_BT)^\alpha \Gamma(\alpha)\zeta(\alpha), \end{aligned}$$

où $\Gamma(\alpha)$ désigne la fonction gamma, et ζ la fonction zeta de Riemann.

Pour un système à N particules, on peut alors introduire la notion de température critique T_c, qui correspond à la température minimale telle que toutes les particules puissent être contenues dans les états d'énergie excités. Elle est donc donnée par,

$$k_BT_c = \left(\frac{C_\alpha^{-1} N}{\Gamma(\alpha)\zeta(\alpha)} \right)^{1/\alpha}.$$

On peut également introduire pour un système à N particules à la température $T \leq T_c$, la notion de fraction condensée. Notons N_0 le nombre de particules dans l'état fondamental, et N_{exc} le nombre de particules dans les états excités. Par définition de la température critique, on a,

$$N_{\text{exc}} = N \left(\frac{T}{T_c} \right)^\alpha.$$

Ainsi, la fraction condensée est définie par N_0/N, et est donnée dans la limite thermodynamique pour $T \leq T_c$ par,

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c} \right)^\alpha. \tag{2.5}$$

Calcul de la densité d'énergie pour une particule libre dans une boîte cubique de taille L

Pour un tel système, les niveaux d'énergie $\varepsilon(n_1,n_2,n_3)$ sont donnés par,

$$\varepsilon(n_1,n_2,n_3) = \frac{\pi^2\hbar^2}{2mL^2}(n_1^2+n_2^2+n_3^2).$$

Ce sont les niveaux d'énergie ε tels qu'il existe une solution non-nulle au problème,

$$\begin{cases} -\dfrac{\hbar^2}{2m}\Delta\phi = \varepsilon\phi \quad \text{sur} \quad \Omega =]0, L[^3, \\ \phi = 0 \quad \text{sur} \quad \partial\Omega. \end{cases}$$

Le nombre de modes $G(\tilde{\varepsilon})$ d'énergie inférieure à $\tilde{\varepsilon}$ est donné par,

$$G(\tilde{\varepsilon}) = \text{Card}\{(n_1, n_2, n_3) \in (\mathbb{R}_+)^3; \varepsilon(n_1, n_2, n_3) \leq \tilde{\varepsilon}\}.$$

Ce cardinal peut être approché, dans la limite thermodynamique, par le volume d'un octant d'une boule de rayon $R = (\frac{2mL^2\varepsilon}{\pi^2\hbar^2})^{1/2}$,

$$G(\varepsilon) = \frac{1}{8} \cdot \frac{4}{3}\pi R^3 = \frac{L^3\sqrt{2}(m\varepsilon)^{3/2}}{3\pi^2\hbar^2}.$$

La densité d'énergie $g(\varepsilon)$ est alors donnée par,

$$g(\varepsilon) = \frac{dG(\varepsilon)}{d\varepsilon} = \frac{L^3 m^{3/2}}{\sqrt{2}\pi^2\hbar^3}\varepsilon^{1/2}. \tag{2.6}$$

Ainsi, en reprenant la notation donnée par l'équation (2.4), on obtient,

$$\alpha = \frac{3}{2}, \quad C_{3/2} = \frac{L^3 m^{3/2}}{\sqrt{2}\pi^2\hbar^3}. \tag{2.7}$$

On peut aussi vérifier que $C_{3/2}$ est bien proportionnel au volume L^3 du système, ce qui justifie que la limite thermodynamique doit être prise quand le nombre de particule tend vers l'infini, mais que $C_{3/2}^{-1}N$ converge vers une limite finie. Ce calcul peut être généralisé pour une dimension d quelconque, et on peut alors démontrer que $\alpha = \frac{d}{2}$, et que la constante C_α est à nouveau proportionnelle au volume du système. On peut donc conclure que le phénomène de condensation de Bose-Einstein ne survit à la limite thermodynamique, dans le cas d'une particule libre, que lorsque la dimension est plus grande ou égale à trois.

Calcul de la densité d'énergie pour une particule dans un piège harmonique

Par souci de simplification, on suppose que le potentiel quadratique (3D) est harmonique et donné par,

$$V(\mathbf{r}) = \frac{1}{2}m\omega(x^2 + y^2 + z^2).$$

Dans ce cas, les niveaux d'énergie $\varepsilon(n_1, n_2, n_3)$ sont donnés par,

$$\varepsilon(n_1, n_2, n_3) = \left(n_1 + n_2 + n_3 + \frac{3}{2}\right)\hbar\omega.$$

Comme précédemment, ce sont les niveaux d'énergie pour lesquels il existe une solution non nulle au problème,

$$-\frac{\hbar^2}{2m}\Delta\phi + V\phi = \varepsilon\phi, \quad \text{sur} \quad \Omega = \mathbb{R}^3.$$

La nombre de modes $G(\varepsilon)$ d'énergie inférieure à ε est approché dans la limite thermodynamique par,

$$G(\varepsilon) = \frac{1}{(\hbar\omega)^3}\int_0^{\varepsilon} d\varepsilon_x \int_0^{\varepsilon-\varepsilon_x} d\varepsilon_y \int_0^{\varepsilon-\varepsilon_x-\varepsilon_y} d\varepsilon_z = \frac{\varepsilon^3}{6(\hbar\omega)^3},$$

et la densité d'énergie $g(\varepsilon)$ est alors donnée par,

$$g(\varepsilon) = \frac{dG(\varepsilon)}{d\varepsilon} = \frac{\varepsilon^2}{2(\hbar\omega)^3}. \tag{2.8}$$

Ainsi, nous obtenons dans ce cas,

$$\alpha = 3, \quad C_3 = \frac{1}{2(\hbar\omega)^3}. \tag{2.9}$$

La distance caractéristique pour ce potentiel est donnée par $(\hbar/m\omega)^{3/2}$, et l'on vérifie ainsi que C_3 est proportionnel au volume caractéristique du système, ce qui justifie comme précédemment le passage à la limite thermodynamique pour $C_3^{-1}N$ constant. Ce résultat peut être généralisé au cas de de la dimension d quelconque, et dans ce cas,

$$g(\varepsilon) = \frac{\varepsilon^{d-1}}{(d-1)!(\hbar\omega)^d},$$

et $\alpha = d$. Ainsi, pour les dimensions plus grandes que deux, le phénomène de condensation de Bose-Einstein survit à la limite thermodynamique dans le cas d'un système confiné par un piège harmonique.

2.2 Les effets des interactions

Nous avons décrit précédemment le phénomène de condensation de Bose-Einstein pour un gaz parfait, c'est-à-dire en négligeant les effets des interactions entre particules. Cependant celles-ci jouent un rôle complexe et essentiel. Elles sont par exemple à l'origine du phénomène de superfluidité (que nous n'aborderons pas ici). Elles influent également sur la température critique de transition de phase (voir [5] dans le cas d'un gaz homogène, et [48] dans le cas d'un gaz confiné par un piège harmonique), et sur la répartition spatiale des atomes [138]. Cette question de la répartition peut être résolue à l'aide de l'équation de Gross-Pitaevskii. L'objectif de ce paragraphe est d'introduire cette équation, et de justifier sa dérivation physique.

Les distances inter-particules dans les gaz d'atomes froids (environ 10^2 nm) sont en général d'un ordre de grandeur plus grand que les distances typiques d'interaction entre deux atomes. Ainsi, il est souvent raisonnable de négliger les interactions à trois corps, ou même plus, ce qui facilite le traitement analytique des interactions. Par exemple, la modélisation par l'équation de Gross-Pitaevskii repose sur cette hypothèse. Cependant, les interactions entre atomes demeurent extrêmement complexes, et l'on ne saurait pas calculer théoriquement le potentiel d'interaction entre atomes avec une bonne précision. C'est la raison pour laquelle, les interactions peuvent être modélisées, d'une manière simplificatrice, par un potentiel d'interaction effectif prennant la forme d'un potentiel de contact. Nous détaillerons en Paragraphe 2.2.2 le principe de cette approximation. Comme nous le verrons dans la suite, ce potentiel effectif est paramétré par une distance, appelée *longueur de diffusion*. Celle-ci peut alors être estimée de manière implicite en observant certaines propriétés d'un condensat, et leur prédiction à l'aide de ce potentiel effectif. En particulier, l'utilisation des résonances de Feshbach a permis de mieux calculer ces longueurs de diffusion.

2.2.1 Une modélisation de la phase condensée : l'équation de Gross-Pitaevskii

Cette partie est dédiée à la dérivation de l'équation de Gross-Pitaevskii pour un gaz dilué de bosons. Ce modèle décrit les propriétés d'un gaz de bosons en interactions à température nulle. Nous en présentons une dérivation "physique" en suivant l'approche proposée dans [138]. Elle est basée sur deux hypothèses primordiales. D'une part, comme expliqué ci-dessus, nous nous plaçons dans le cas où les distances inter-particules sont grandes devant la longueur de diffusion, ce qui permet de modéliser les interactions entre particules par un potentiel effectif d'interaction de contact. Nous justifions cette approximation en Paragraphe 2.2.2. D'autre part, nous utilisons une approximation de champ moyen, appelée encore *approximation de Hartree*. Celle-ci consiste à supposer que la fonction d'onde Ψ d'un ensemble de N atomes est tensorisée comme produit de la fonctions d'onde ϕ à une particule,

$$\Psi(t, \mathbf{r}_1, \mathbf{r}_1, \ldots, \mathbf{r}_N) = \prod_{j=1}^{N} \phi(t, \mathbf{r}_j), \tag{2.10}$$

où ϕ est normalisée,

$$\int_{R^3} |\phi(\mathbf{r})|^2 \, d\mathbf{r} = 1.$$

Cette approche permet d'approximer l'énergie du système uniquement en fonction de la fonction d'onde ϕ. Il est alors convenable de renormaliser cette fonction d'onde pour pouvoir l'interpréter comme une densité de particule. On posera alors $\psi = \sqrt{N}\phi$. La justification rigoureuse de ce modèle repose principalement sur la justification de l'approximation

de Hartree. De tels résultats ont été établis dans la limite, appelée *limite champ moyen*, où le nombre de particules N tend vers l'infini et que l'intensité des interactions tend vers 0.

Dérivation physique

Le Hamiltonien d'un ensemble de N atomes est donné par

$$H = \sum_{j=1}^{N} -\frac{\hbar^2}{2m}\Delta_{r_j} + V(\mathbf{r}_j) + \sum_{1\leq j<k\leq N} U(\mathbf{r}_i, \mathbf{r}_j),$$

où V est le potentiel confinant externe et $U(\mathbf{r}_i, \mathbf{r}_j)$ le potentiel d'interaction entre deux particules de positions $\mathbf{r}_i$ et $\mathbf{r}_j$. Ainsi, l'énergie du système est donnée par,

$$\begin{aligned} E(\phi) &= \langle H\phi, \phi\rangle \\ &= N\int_{\mathbb{R}^3}\left(-\frac{\hbar^2}{2m}\Delta\phi(\mathbf{r}) + V(\mathbf{r})\right)\overline{\phi(\mathbf{r})}d\mathbf{r} \\ &+ \frac{N(N-1)}{2}\int_{(\mathbb{R}^3)^2} U(\mathbf{r},\mathbf{r}')\left|\phi(\mathbf{r})\right|^2\left|\phi(\mathbf{r}')\right|^2 d\mathbf{r}'d\mathbf{r}. \end{aligned} \tag{2.11}$$

En utilisant le changement de variable $\sqrt{N}\phi \to \phi$ et l'approximation $(N-1)/N \approx 1$ on peut écrire,

$$\begin{aligned} E(\phi) \approx &\int_{\mathbb{R}^3}\left(\frac{\hbar^2}{2m}\left|\nabla\phi(\mathbf{r})\right|^2 + V(\mathbf{r})\left|\phi(\mathbf{r})\right|^2\right)d\mathbf{r} \\ &+ \frac{1}{2}\int_{(\mathbb{R}^3)^2} U(\mathbf{r},\mathbf{r}')\left|\phi(\mathbf{r})\right|^2\left|\phi(\mathbf{r}')\right|^2 d\mathbf{r}'d\mathbf{r}. \end{aligned} \tag{2.12}$$

L'énergie de Gross-Pitaevskii est obtenue en approchant maintenant le potentiel U par un potentiel effectif U_{eff} de contact donné par, $U_{eff}(\mathbf{r},\mathbf{r}') = U_0\delta(\mathbf{r}-\mathbf{r}')$, avec $U_0 = 4\pi\hbar^2 a/m$. La longueur a est la longueur de diffusion, qui est définie dans le paragraphe suivant, dans lequel nous justifions cette approximation de U par un potentiel effectif. L'énergie de Gross-Pitaevskii est finalement donnée par,

$$E_{\mathrm{GP}}(\phi) = \int_{\mathbb{R}^3}\left(\frac{\hbar^2}{2m}\left|\nabla\phi(\mathbf{r})\right|^2 + V(\mathbf{r})\left|\phi(\mathbf{r})\right|^2 + \frac{U_0}{2}\left|\phi(\mathbf{r})\right|^4\right)d\mathbf{r}. \tag{2.13}$$

L'équation de type Schrödinger non-linéaire qui décrit l'évolution temporelle de la fonction d'onde ϕ est donnée par,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\phi(t,\mathbf{r}) &= \nabla E_{\mathrm{GP}}(\phi(t,\mathbf{r})) \\ &= \left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}) + \frac{4\pi\hbar^2 a}{m}\left|\phi(t,\mathbf{r})\right|^2\right)\phi(t,\mathbf{r}), \end{aligned} \tag{2.14}$$

où ∇E_{GP} est le gradient de E_{GP} pour la norme $L^2(\mathbb{R}^3)$. Ce modèle porte le nom des deux physiciens qui l'ont obtenu, Gross [89] et Pitaevskii [140].

Limite champ moyen

La limite champ-moyen survient lorsque l'intensité des interactions (supposées faibles) est renormalisée par un paramètre $\kappa(N)$, qui dépend du nombre N d'atomes. On considère alors le Hamiltonien

$$H_N = \sum_{j=1}^{N} \left(-\frac{\hbar^2}{2m} \Delta_{r_j} + V(\mathbf{r}_j) \right) + \kappa(n) \sum_{1 \leq j < k \leq N} U(\mathbf{r}_i, \mathbf{r}_j).$$

Il apparaît alors que l'énergie du système est d'ordre $\mathcal{O}(N) + \kappa(N)\mathcal{O}(N^2)$. En choisissant $\kappa(N) = \mathcal{O}(N^{-1})$, on obtient une énergie proportionnelle au nombre de particules. La limite champ moyen correspond plus précisément à la limite

$$N \to +\infty \quad \text{and} \quad \kappa \to 0, \quad \text{où} \quad \kappa N = \text{const.}$$

On pose $\kappa = (N-1)^{-1}$. L'énergie du système s'écrit alors, en remplaçant le potentiel par son potentiel effectif, par

$$E_N(\Psi) = \int_{\mathbb{R}^{3N}} \left(\frac{\hbar^2}{2m} |\nabla \Psi(\mathbf{r})|^2 + V(\mathbf{r}) |\Psi(\mathbf{r})|^2 + \frac{U_0}{2} |\Psi(\mathbf{r})|^4 \right) d\mathbf{r}.$$

Dans le cas de l'approximation de Hartree donnée par l'Équation (2.10), nous retrouvons bien

$$E_N(\Psi) = N E_{\mathrm{GP}}(\phi), \tag{2.15}$$

où E_{GP} est donnée par l'Équation (2.13). La justification de l'approximation de Hartree tient à pouvoir établir le passage à la limite suivant :

$$\lim_{N \to +\infty} \frac{E(N)}{N} = e_{\mathrm{GP}}, \tag{2.16}$$

où $E(N)$ est l'infimum pour la fonctionnelle d'énergie E_N du problème à N corps sur l'ensemble des fonctions d'onde à N corps, de norme $L^2(\mathbb{R}^{3N})$ unitaire et symmétriques, et où $e_{\mathrm{GP}} = \inf_{\|\phi\|_{L^2}=1} E_{\mathrm{GP}}(\phi)$. Notons que l'on a pour tout entier N, $\frac{E(N)}{N} \leq e_{\mathrm{GP}}$ d'après l'Équation (2.15). Ce type de passage à la limite a été démontré dans [17, 74, 117, 116]

Modélisation de la rotation

Cette équation peut-être étendue à des systèmes plus complexes. Par exemple, il est possible de modéliser une rotation du condensat. Pour ce faire, la dérivation ci-dessus doit être menée dans le référentiel tournant du condensat dans lequel l'impulsion sera modifiée. Une dérivation claire et détaillée pourra être trouvée dans le [146, Chapitre 1]. On suppose alors que la rotation est centrée en l'origine du repère, et on note $\mathbf{\Omega}$ son axe de rotation. Sa norme est notée Ω, et elle correspond à la vitesse de rotation du condensat.

On retrouve alors le résultat en ajoutant au membre de droite de l'équation (2.14) un générateur infinitésimal de la rotation donné par $\mathbf{\Omega} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{p}})$ où $\hat{\mathbf{r}}$ et $\hat{\mathbf{p}}$ sont respectivement les opérateurs positions et impulsions. On rappelle que l'opérateur $\hat{\mathbf{r}} \times \hat{\mathbf{p}}$ n'est autre que l'opérateur moment cinétique. En représentation de position, l'opérateur impulsion devient $\hat{\mathbf{p}} = -i\hbar\nabla$, et on obtient alors l'équation de Gross-Pitaevskii,

$$i\hbar\frac{\partial}{\partial t}\phi(t,\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\Delta + i\hbar\mathbf{\Omega}\cdot(\mathbf{r}\times\nabla) + V(\mathbf{r})\right)\phi(t,\mathbf{r}) + \frac{4\pi\hbar^2 a}{m}\left|\phi(t,\mathbf{r})\right|^2\phi(t,\mathbf{r}). \tag{2.17}$$

Ce terme de rotation sera pris en compte dans le Chapitre 6, consacré à l'analyse numérique de dynamiques métastables. C'est d'ailleurs ce terme de rotation qui sera responsable de cette métastabilité (voir le Paragraphe 6.1.1 du Chapitre 6).

2.2.2 Définition de la longueur de diffusion

L'objectif de cette partie est de justifier formellement l'utilisation d'un potentiel d'interaction effectif, et de définir la notion de *longueur de diffusion*. Dans un premier temps, nous reformulons le processus de collision entre deux particules comme un processus de diffusion d'une particule par un potentiel d'interaction U. Ce dernier processus suffit pour décrire le premier. Dans un second temps, nous expliquons que pour les échelles spatiales de ce problème de diffusion à un corps, celui-ci peut se reformuler comme un problème de diffusion d'une onde plane incidente sur le potentiel U. Nous faisons le calcul (formel) d'une approximation de l'onde diffusée, que l'on nomme approximation de Born, qui nous permet d'introduire la notion de longueur de diffusion. Ce calcul est en partie inspiré par le développement rigoureux de [47, Chapitre 8]. Nous remarquons alors que l'amplitude de l'onde diffusée lointaine ne dépend que de la masse (ou la valeur moyenne) du potentiel U, dans la limite où l'énergie de l'onde tend vers 0. Autrement dit, deux potentiels différents, mais de même masse, produisent des ondes diffusées équivalentes dans la limite lointaine. C'est cette remarque qui justifie l'introduction d'un potentiel effectif de contact, et qui permet de simplifier le processus de collision en l'approchant par une interaction de contact, c'est à dire un potentiel de type fonction delta de Dirac de même masse que le potentiel initial.

Considérons dans un premier temps le cas d'une interaction à deux corps. Notons $(\hat{\mathbf{r}}_1, \hat{\mathbf{p}}_1)$ et $(\hat{\mathbf{r}}_2, \hat{\mathbf{p}}_2)$ les opérateurs positions et impulsions des deux particules de masses respectives m_1 et m_2. Notons U le potentiel d'interaction entre les deux particules de telle sorte que le Hamiltonien du problème est donné par :

$$H(\hat{\mathbf{r}}_1, \hat{\mathbf{p}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{p}}_2) = \frac{\hat{\mathbf{p}}_1^2}{2m_1} + \frac{\hat{\mathbf{p}}_2^2}{2m_2} + U(\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2).$$

Classiquement, en séparant le mouvement d'ensemble du système du mouvement relatif entre les deux particules, cette dynamique peut être réduite à la dynamique d'un seul corps fictif d'opérateurs position $\hat{\mathbf{r}} = \hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2$ et impulsion $\hat{\mathbf{p}} = \frac{m_2\hat{\mathbf{p}}_1 - m_1\hat{\mathbf{p}}_2}{m_1+m_2}$. Le Hamiltonien

de cette particule (fictive) est donné par :

$$H(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \frac{\hat{\mathbf{p}}^2}{2m} + U(\hat{\mathbf{r}}), \tag{2.18}$$

où $m = \frac{m_1 m_2}{m_1 + m_2}$. Cette procédure est décrite dans [19, Chapitre 11].

Formellement nous sommes intéressés par le calcul des états propres d'énergie E pour l'Hamiltonien H défini par (2.18), c'est à dire les solutions de l'équation aux valeurs propres suivantes,

$$\left(-\frac{\hbar^2}{2m}\Delta + U(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{2.19}$$

On suppose que le potentiel U décroit plus vite que $|\mathbf{r}|^{-1}$ à l'infini, de telle sorte que le spectre de l'opérateur $\left(-\frac{\hbar^2}{2m}\Delta + U(\mathbf{r})\right)$ soit continu de 0 à l'infini. On s'intéresse aux états de diffusions qui sont les états propres du spectre continu. On suppose alors $E > 0$, et on pose $E = \frac{\hbar^2 k^2}{2m}$. Ces états propres étant non normés dans $L^2(\mathbb{R}^3)$, nous changeons de point de vue pour corriger cette difficulté. On suppose que l'on se place dans le cas asymptotique où une particule s'approche du potentiel depuis l'infini, avec une impulsion bien déterminée, et on s'intéresse à l'état de la particule diffusée dans un futur lointain. Cette hypothèse est justifiée, dans le cadre de gaz dilués, par le fait que le potentiel d'interaction est concentré dans une région de l'espace de taille très inférieure aux distances inter-particules typiques. Cette hypothèse revient à supposer que la particule est dans un *état asymptotique* modélisé par une onde plane incidente $\phi(\mathbf{r}) = \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}$ d'impulsion $\mathbf{k} \in \mathbb{R}^3$ sur le potentiel U, solution de l'équation de Helmholtz,

$$\left(\Delta + k^2\right)\phi(\mathbf{r}) = 0, \quad \text{où } |\mathbf{k}| = k.$$

Bien que cet état ne soit pas physique à proprement parler, cette modélisation correspond au cas limite d'un paquet d'onde dont la dispersion relative en impulsion tend vers 0 : $\Delta p/p \to 0$. Cependant nous présentons seulement le raisonnement à l'aide de cette modélisation asymptotique, par souci de simplicité. Le but est dans ce cas de sélectionner une solution de (2.19) de la forme,

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \psi_{\text{diff}}(\mathbf{r}).$$

Comme ϕ n'est pas intégrable, il est judicieux de formuler le problème (2.19) en fonction de ψ_{diff} plutôt que ψ. C'est ce qui permet d'obtenir un problème bien posé, à onde incidente ϕ fixée. Ainsi, on obtient formellement le problème suivant,

$$\frac{2m}{\hbar^2}U(\mathbf{r})(\phi(\mathbf{r}) + \psi_{\text{diff}}(\mathbf{r})) = (\Delta + k^2)\psi_{\text{diff}}(\mathbf{r}). \tag{2.20}$$

Pour obtenir un problème bien posé, il est classique d'ajouter une condition sur le comportement à l'infini de $\psi_{\text{diff}}(\mathbf{r})$. C'est la condition de radiation de Sommerfeld, donnée

par,

$$\lim_{r\to+\infty} r\left(\frac{\partial\psi_{\text{diff}}}{\partial r} - ik\psi_{\text{diff}}\right) = 0. \tag{2.21}$$

Celle-ci assure l'unicité d'une solution au problème (2.20)-(2.21), et donc son caractère bien posé. Formellement, elle permet de sélectionner la solution physique du problème. On peut dériver du problème (2.20)-(2.21) une équation intégrale, appelée équation de Lippmann-Schwinger, qui lui est équivalente (sous réserve de régularité). Afin d'introduire cette équation, introduisons le noyau de Green *sortant* de l'équation de Helmholtz, noté G_k^+ et défini par,

$$G_k^+(\mathbf{r}) = \frac{m}{2\pi\hbar^2}\frac{\mathrm{e}^{+ikr}}{r}, \forall \mathbf{r}\in\mathbb{R}^3. \tag{2.22}$$

Celui-ci est une solution au sens des distributions de l'équation de Helmholtz impulsionnelle

$$-\frac{\hbar^2}{2m}(\Delta + k^2)G_k^+(\mathbf{r}) = \delta(\mathbf{r}).$$

L'équation de Lippmann-Schwinger est finalement donnée par,

$$\psi_{\text{diff}}(\mathbf{r}) = -\int_{\mathbb{R}^3} G_k^+(\mathbf{r}-\mathbf{r}')U(\mathbf{r}')(\psi_{\text{diff}}(\mathbf{r}') + \phi(\mathbf{r}'))d^3\mathbf{r}'. \tag{2.23}$$

A partir de l'équation (2.23), on peut calculer une approximation du champ diffusé ψ_{diff} lointain. C'est ce que l'on nomme *l'approximation de Born*. Celle-ci repose sur l'observation que le champ diffusé lointain décroit comme $1/r$ (comme le noyau de Green). Ainsi pour de larges valeurs r on peut approcher $(\psi_{\text{diff}}(\mathbf{r})+\phi(\mathbf{r}))$ par $\phi(\mathbf{r})$, et l'équation (2.23) devient alors,

$$\psi_{\text{diff}}(\mathbf{r}) \underset{r\to+\infty}{\approx} -\int_{\mathbb{R}^3} G_k^+(\mathbf{r}-\mathbf{r}')U(\mathbf{r}')\phi(\mathbf{r}')d^3\mathbf{r}'. \tag{2.24}$$

On rappelle que l'on est intéressé par une approximation du champs diffusé $\psi_{\text{diff}}(\mathbf{r})$ pour des valeurs de $|\mathbf{r}|$ plus grandes que la distance typique d'interaction entre particules. Notons a cette dernière. Cela revient à supposer que $|\mathbf{r}| \gg a$, et que l'intégrale ci-dessus ne porte que sur des valeurs $\mathbf{r}'$ telles que $|\mathbf{r}| \gg |\mathbf{r}'|$. Cette remarque permet de faire l'approximation que $1/|\mathbf{r}-\mathbf{r}'| \approx 1/r$, et que $|\mathbf{r}-\mathbf{r}'| \approx r - \mathbf{u}'\cdot\mathbf{r}'$, avec $\mathbf{u}' = \mathbf{r}/r$. On peut alors approcher (2.24), à très basse énergie, par,

$$\psi_{\text{diff}}(\mathbf{r}) \approx f(\mathbf{k})\frac{\mathrm{e}^{ikr}}{r}, \quad \text{avec} \quad f(\mathbf{k}) = -\frac{m}{2\pi\hbar^2}\int_{\mathbb{R}^3} \mathrm{e}^{i(\mathbf{k}-\mathbf{u}'k)\cdot\mathbf{r}'}U(\mathbf{r}')d^3\mathbf{r}', \tag{2.25}$$

où l'on a posé $\mathbf{u}' = \mathbf{r}/r$. On définit la *longueur de diffusion* a_s par, $\lim_{k\to 0} f(\mathbf{k}) = -a_s$,

c'est-à-dire,

$$a_s = \frac{m}{2\pi\hbar^2} \int_{\mathbb{R}^3} U(\mathbf{r}')d^3\mathbf{r}'. \tag{2.26}$$

Par ailleurs, comme nous l'avons énoncé en introduction, l'approximation donnée par (2.25)-(2.26) pour une onde incidente d'énergie suffisamment basse ne dépend que de l'intégrale du potentiel $U(\mathbf{r})$. Ainsi, approcher ce potentiel par un autre potentiel de même masse, conduit au même comportement diffusif asymptotique que le potentiel original. C'est la raison pour laquelle, on introduit un potentiel effectif $U_{eff}(\mathbf{r})$ donné par

$$U_{eff}(\mathbf{r}) = \frac{2\pi\hbar^2 a_s}{m}\delta(\mathbf{r}),$$

où δ représente la fonction delta de Dirac. Ce potentiel représente une interaction de contact, et permet de simplifier de nombreuses dérivations physiques analytiques.

2.3 Le protocole expérimental

La réalisation d'un condensat de Bose-Einstein, nécessite de pouvoir refroidir un gaz de bosons à quelques fractions de microKelvins. A titre de comparaison, la température du fond diffus cosmologique est d'environ 2,728 K. La réalisation expérimentale de condensats de Bose-Einstein nécessite, à l'heure actuelle, l'utilisation de plusieurs méthodes de refroidissement telles que le refroidissement par laser, ainsi que le refroidissement par évaporation. Nous présentons dans cette partie le mécanisme général de ces méthodes de refroidissement, qui permet de mieux comprendre comment les condensats sont conçus en pratique. Cette compréhension du protocole expérimental est particulièrement intéressante pour comprendre la mise en évidence numérique (à l'aide du modèle *Stochastic Gross-Pitaevskii Equation* que nous présentons en Paragraphe 2.4) de l'apparition spontanée de vortex lors de la transition de phase. Voir par exemple [30, 167]. De plus nous introduisons à la fin de ce paragraphe une modélisation des défauts du protocole expérimental dont l'analyse numérique fait l'objet du Chapitre 3. Nous décrivons aussi succinctement les mécanismes de piégeage des atomes, qui sont liés aux mécanismes de refroidissement.

Les procédés expérimentaux de refroidissement à des températures de l'ordre de la dizaine de nanoKelvins sont multiples. Deux méthodes principales sont utilisées successivement. D'abord les atomes sont *pré-refroidis* par des méthodes de refroidissement par laser. Cette étape permet de réduire la température jusqu'à quelques dizaines de microKelvins. Ce procédé est limité par les émissions spontanées des photons absorbés qui ne permettent pas d'atteindre des températures plus basses. Cette méthode de refroidissement est en général couplée à une méthode de piégeage optico-magnétique (MOT pour Magneto-Optical Trap). Les températures obtenues avec cette méthode sont suffisamment basses pour permettre ensuite l'utilisation des méthodes de refroidissement par évaporation qui

permettent d'atteindre des températures inférieures aux températures critiques de transition de phase. Cette méthode de refroidissement est quant à elle couplée à une méthode de confinement purement magnétique. Cependant, ces pièges possèdent de nombreux inconvénients. Par exemple, ils ne permettent pas de confiner tous les atomes du gaz (mais seulement ceux dans certains états hyperfins) et conduisent à de fortes pertes d'atomes lors des expériences. Pour ces raisons, les physiciens ont souvent recours à l'utilisation de pièges purement optiques une fois que le condensat est formé (voir [18] pour la première expérience de ce type). Dans la suite de ce paragraphe, nous tentons de décrire brièvement le mécanisme général de ces méthodes qui conduisent à la formation de condensats.

2.3.1 Le refroidissement par laser, et les pièges optico-magnétiques (MOT)

Cette partie s'inspire de [106, 124]. Pour bien comprendre cette méthode, il faut d'abord comprendre les interactions entre atomes et photons. Quand un atome est éclairé par un laser, deux phénomènes distincts apparaissent, qui conduisent à deux types de forces que sont la pression de radiation et la force dipolaire. Laissons pour le moment cette dernière de côté (nous y reviendrons en Paragraphe 2.3.3) et expliquons brièvement le phénomène de pression de radiation. Un photon possède une énergie élémentaire donnée par $\hbar\nu$ où $\hbar$ est la constante de Planck et ν sa fréquence. Si cette dernière est telle que l'énergie élémentaire $\hbar\nu$ est égale à la différence d'énergie entre deux niveaux d'énergie de l'atome, alors on parle de fréquence de résonance, et l'atome peut absorber les photons possédant ces fréquences résonantes. Lors d'une absorption, la quantité de mouvement de l'atome est modifiée par conservation de la quantité de mouvement du système atome/photon. On rappelle que la quantité de mouvement d'un photon est donnée par $\hbar\nu/c$, où c est la célérité. Après cette phase d'absorption, l'atome peut se désexciter par émission spontanée d'un photon, ce qui modifie de nouveau la quantité de mouvement de l'atome. Dans ce cas, la direction du photon est aléatoire. Ainsi, la quantité de mouvement moyenne des photons émis est nulle grâce à l'isotropie de l'émission. On peut donc modifier la quantité de mouvement d'un atome en le bombardant d'un grand nombre de photons dans une certaine direction. C'est ce phénomène que l'on nomme pression de radiation.

Nous expliquons maintenant comment ce phénomène peut permettre de réduire la température d'un gaz d'atomes en réduisant la quantité de mouvement de chaque particule. Cette méthode repose sur l'effet Doppler. Ce phénomène conduit un objet en mouvement à observer une onde incidente à une fréquence différente de sa fréquence émise. Plus précisément, pour un atome se déplaçant à la vitesse $\mathbf{v}$ et une onde plane de vecteur d'onde $\mathbf{k}$ et de fréquence ν, l'atome percevra une onde de fréquence modifiée $\nu' = \nu - \mathbf{k}\cdot\mathbf{v}$. On peut alors s'apercevoir qu'un atome qui se déplace dans la direction d'une source laser observera une fréquence de l'onde incidente légèrement supérieure à la fréquence du laser. Ainsi pour ralentir un atome qui se déplace vers une source laser, il est nécessaire de choisir une fréquence ν inférieure à la fréquence de résonance de l'atome, de telle sorte que ν' soit égale à cette fréquence. Il est donc possible de ne ralentir que les atomes se déplaçant dans la direction du laser. En éclairant le gaz de six lasers (pour les six directions orientées de

l'espace) il est alors possible de ralentir tous les atomes contenus dans ce gaz.

Cependant, bien que la vitesse des photons émis soit en moyenne nulle (par isotropie) cela ne suffit pas pour refroidir un atome à des températures arbitrairement basses. Pour un atome se déplaçant vers la source d'un laser, chaque cycle absorption/émission mène presque sûrement à une augmentation de la quantité de mouvement de l'atome dans le plan normal à son impulsion. Ainsi, il existe une température minimale, appelée *température Doppler*, qu'il n'est pas possible de franchir avec un refroidissement par laser. C'est ce qui justifie l'utilisation du refroidissement par évaporation.

On précise maintenant la méthode de piégeage utilisée dans les pièges MOT. Supposons qu'un atome s'éloigne, dans une certaine direction, du centre d'un piège. Alors, en favorisant l'absorption des photons incidents, on peut inverser la quantité de mouvement de l'atome, et ainsi rapprocher l'atome du centre du piège. C'est exactement ce qu'il se passe en pratique. Pour créer une variation spatiale de la propension des atomes à absorber un photon, on crée une dépendance spatiale des écarts d'énergies avec les états excités. Cette dépendance repose sur l'effet Zeeman qui implique qu'un atome soumis à l'effet d'un champ magnétique externe voit ses niveaux atomiques d'énergie modifiés. D'ailleurs, ce phénomène peut mener à une levée de la dégénérescence des états d'énergie. Ainsi, la méthode consiste à plonger le piège dans un champ magnétique variant spatialement, pour que les différences d'énergie entre deux niveaux varient également spatialement, et ainsi obtenir une hétérogénéité de la pression de radiation qui permette de confiner les atomes.

2.3.2 Le refroidissement par évaporation et les pièges magnétiques

Tout d'abord, commençons par expliquer le principe de confinement magnétique. Nous donnerons ensuite quelques explications des mécanismes utilisés lors des refroidissements par évaporation. La compréhension précise de ces méthodes nécessite des connaissances plus approfondies sur les structures atomiques élémentaires et nous renvoyons à [88, 106, 129, 138, 136, 169] pour une compréhension plus approfondie des phénomènes décrits ci-dessous. La présentation demeurera néanmoins accessible au lecteur non-spécialiste. Le mécanisme de ce piège repose sur l'existence d'un moment magnétique μ des atomes à confiner. Celui-ci est associé au moment cinétique orbital des électrons et à leur spin. Comme précédemment, en présence d'un champ magnétique $\mathbf{B}$, les niveaux d'énergie se séparent par effet Zeeman, et l'énergie d'un sous niveau est alors donné par $E(m_F) \propto \mu_B g_F m_F B$, où m_F désigne le nombre quantique magnétique. La grandeur μ_B désigne le magnéton de Bore, qui est une grandeur positive. On peut alors distinguer deux comportements différents suivant le signe de $g_F m_F$. On dit des états tels que $g_F m_F < 0$ que ce sont des *strong field seeking states* et les autres sont appelés *weak field seeking states*. Cette dénomination vient simplement du fait que les atomes tels que $g_F m_F < 0$ minimisent leur énergie dans les zones de forte intensité du champ magnétique, et inversement pour les autres. Ainsi, pour confiner des atomes de type *weak field seeking states* il suffirait de construire un champ magnétique possédant un minimum local d'intensité, et les atomes de type *strong field seeking states* pourraient être confinés par des champs magnétiques

possédant un maximum local d'intensité. Cependant, ce dernier type de champ magnétique n'est pas constructible en vertu du théorème d'Earnshaw. Ainsi, seuls les atomes de type *weak field seeking states* peuvent être confinés par ce type de piège (dans des minima locaux du champ magnétique).

Une limitation de ces pièges est la perte d'atomes qui survient lorsque un état *weak field seeking state* devient *strong field seeking states*. Le piège devient alors repoussant après ces transformations. Celles-ci peuvent survenir pour deux raisons. Le premier phénomène provient des *transitions Majorana*. Celles-ci interviennent lorsque le moment magnétique de l'atome ne parvient pas à suivre (adiabatiquement) la direction du champ magnétique lors de son déplacement dans le piège. Ce phénomène apparaît lorsque le champ magnétique devient suffisamment faible, ce qui peut être le cas au centre du piège. En effet, comme nous l'avons dit précédemment, les atomes sont attirés par les zones de faible intensité du champ magnétique, et donc le centre du piège correspond au minimum du champ magnétique. Plus précisément, cela peut se produire quand la fréquence de Larmor (qui est la fréquence de précession du moment magnétique des atomes autour d'un champ externe) devient du même ordre de grandeur, ou plus petite, que la variation du champ magnétique (dans le référentiel de l'atome). Le second phénomène est dû aux collisions inélastiques. Ce sont celles qui ne préservent pas l'énergie cinétique, et qui peuvent donc entraîner un changement d'état. Il se trouve que celles-ci sont beaucoup moins probables que les collisions élastiques, et ne sont donc pas limitantes dans la durée de vie du piège (voir [129, 138]).

La méthode de refroidissement par évaporation consiste à éliminer du piège les atomes de plus haute énergie cinétique, tout en imposant la thermalisation du système, ce qui permet d'amener le système dans un état d'équilibre d'énergie plus faible, au prix d'une perte significative du nombre d'atomes piégés. On peut aussi noter que le temps de thermalisation doit être très inférieur au temps caractéristique de perte des atomes pour que cette méthode soit possible. Pour éliminer ces atomes de plus haute énergie, on peut avoir recours à une diminution de la profondeur du piège, ce qui permet aux atomes les plus rapides de franchir cette barrière. Une méthode plus efficace, appelée *couteau radio-fréquence*, permet d'éliminer les atomes de plus haute énergie en induisant une transition vers un *strong field seeking state*. On ne détaille pas cette méthode ici, et l'on renvoie à [76, Chapitre 4].

2.3.3 Le confinement optique

Comme nous l'avons énoncé plus haut, cette étape de confinement utilise la force dipolaire. Commençons par expliciter ce mécanisme avant de présenter cette méthode de confinement. Quand un atome est soumis à un champ électrique extérieur $\mathbf{E}$, il se polarise. On peut comprendre ce phénomène en imaginant les électrons et les protons tirés dans les deux sens opposés de la direction du champ électrique. Son moment dipolaire $\mathbf{p}$ est proportionnel à ce champ électrique, et à une constante α (dépendant de l'atome) appelée *polarisabilité de l'atome*. On a alors $\mathbf{p} = \varepsilon_0 \alpha \mathbf{E}$, où ε_0 est la permittivité diélectrique

du vide. Ce champ électrique induit une énergie potentielle dipolaire donnée par $U = -\mathbf{p} \cdot \mathbf{E} = -\varepsilon_0 \alpha E^2$, qui implique une force dipolaire $\mathbf{F}_{\text{dipolaire}} = -\nabla U$, qui pousse donc le dipôle vers les zones de maximum du champ E. Ainsi, si l'on éclaire un atome à l'aide d'un laser, celui-ci sera attiré vers les zones de maximum du champ électrique, situées généralement au centre du faisceau. Il suffit alors d'éclairer un atome par deux faisceaux lasers concourants pour le piéger, si son énergie cinétique est assez faible. C'est pour cette raison que ce piège est en général utilisé après un refroidissement par évaporation. Bien sûr, pour éviter l'absorption de photons, la fréquence des lasers doit être choisie loin des fréquences de résonance des atomes à confiner. C'est pour cela que ces pièges sont appelés dans la littérature physique *far-off-resonance laser traps*. Il est même possible à l'aide de ce principe de déplacer le condensat, c'est pourquoi cette méthode est aussi appelée *pince optique*.

2.4 Quelques précisions sur SPGPE

Le modèle SPGPE a été introduit au Paragraphe 1.3. Nous faisons ici le lien entre la formulation proposée dans [25, Paragraphe 5.3], et la formulation adimensionnée proposée en introduction.

Rappelons qu'à température non-nulle, la condensation n'est pas complète, et qu'alors une fraction des bosons demeure non-condensée. On distingue alors deux phases au sein du gaz de bosons dites condensée et non-condensée. La première est composée des bosons dans les états d'énergie les plus bas, tandis que la deuxième est composée des bosons dans les états d'énergie les plus hauts. Cette phase non-condensée se comporte comme un bain d'atomes proches de l'équilibre thermique. La dynamique de la phase condensée se trouve alors perturbée par cette phase non-condensée par interactions. En pratique, nous sommes intéressés par la dynamique de la fonction d'onde décrivant la phase condensée. Le modèle SPGPE modélise la dynamique du système ouvert que constitue la phase condensée. Son intérêt réside dans la description de l'interaction entre ces deux phases. Cette modélisation partitionne l'ensemble des niveaux d'énergie en deux ensembles. Pour cela elle introduit un niveau d'énergie E_{cut} tel que les modes d'énergie inférieure à ce niveau sont considérés comme condensés, et les autres comme non-condensés.

2.4.1 Formulation physique de SPGPE

On note dans la suite par $\mathbf{x}$ un élément $(x, y, z) \in \mathbb{R}^3$, et par $\mathbf{r}$ un élément $(x, y) \in \mathbb{R}^2$. L'évolution temporelle de la fonction d'onde $\alpha_t(\mathbf{x})$ de la bande condensée est décrite par l'équation,

$$d\alpha_t = \mathcal{P}\left\{-\frac{i}{\hbar} L_{\text{GP}} \alpha_t(\mathbf{x}) dt + \frac{\gamma(\mathbf{x})}{k_B T}(\mu - L_{\text{GP}})\alpha_t(\mathbf{x}) dt + \sqrt{2\gamma(\mathbf{x})} dW(\mathbf{x}, t)\right\}, \tag{2.27}$$

avec,

$$L_{\mathrm{GP}}\alpha_t(\mathbf{x}) = \left(-\frac{\hbar^2\Delta}{2m} + V(\mathbf{x}) - \Omega L_z + u\left|\alpha_t(\mathbf{x})\right|^2\right)\alpha_t(\mathbf{x}),$$

où le processus $(\alpha_t)_{t\geq 0}$ est à valeurs dans $\mathbb{C}$ et $(W_t)_{t\geq 0}$ est un processus de Wiener cylindrique dont la définition est précisée plus loin. La constante de Planck réduite est notée $\hbar$, la constante de Boltzmann est notée k_B, la température T, le potentiel chimique μ et $\gamma(\mathbf{x})$ est une fonction que l'on précise plus loin. Comme expliqué en Paragraphe 2.2.2, l'intensité des interactions u est donnée par,

$$u = 4\pi\hbar^2 a/m,$$

où a est la longueur de diffusion et m la masse d'un atome. Les conditions expérimentales permettent de fixer la valeur de la température T et du potentiel chimique μ, qui est défini comme la grandeur conjuguée par la transformée de Legendre du nombre d'atomes dans le condensat. De plus, l'opérateur de rotation L_z est donné par,

$$L_z = i\hbar(x\partial_y - y\partial_x),$$

et le potentiel extérieur de confinement V par,

$$V(\mathbf{x}) = \frac{1}{2}m\left(\omega_x^2x^2 + \omega_y^2y^2 + \omega_z^2z^2\right), \quad \text{et en 2 dimensions} \quad V(\mathbf{r}) = \frac{1}{2}m\left(\omega_x^2x^2 + \omega_y^2y^2\right).$$

Pour définir précisément le projecteur $\mathcal{P}$ et le processus de Wiener (W_t), nous introduisons $(e_k(\mathbf{x}))_{k\in\mathbb{N}}$ une base orthonormale de $L^2(\mathbb{R}^3)$ constituée de vecteurs propres de l'opérateur

$$-\frac{\hbar^2\Delta}{2m} + V(\mathbf{x}) - \Omega L_z. \tag{2.28}$$

Notons $(\lambda_k)_{k\in\mathbb{N}}$ la suite de valeurs propres associée. Posons alors K le sous-espace vectoriel de $L^2(\mathbb{R}^3)$ de dimension finie défini par,

$$K = \mathrm{Span}\left\{e_k; \lambda_k \leq E_{\mathrm{cut}}\right\}.$$

Ainsi, le projecteur $\mathcal{P}$ correspond à la projection orthogonale sur le sous-espace K qui correspond à l'espace naturel de la fonction d'onde de la bande condensée. De la même manière, le processus de Wiener cylindrique $(W_t)_{t\geq 0}$ est défini par,

$$W(\mathbf{x}, t) = \sum_{k\in\mathbb{N}} \beta_k(t)e_k(\mathbf{x}),$$

où $(\beta_k(t))_k$ est une famille indépendante de mouvements browniens à valeurs complexes.

Le champ scalaire γ est donné par [25, Équation 172]. En pratique, ce champ est

choisi constant et proportionnel à la constante $\gamma_0 = 4m(ak_BT)^2/\pi\hbar^3$. La constante de proportionnalité ρ (telle que $\gamma = \rho\gamma_0$) est choisie en accord avec les expériences, et est généralement comprise dans l'intervalle $[1, 10]$. Dans ce cas, par définition de la projection $\mathcal{P}$, le processus de Wiener W peut être défini par,

$$W^{\mathcal{P}}(\mathbf{x}, t) = \sum_{\{k; e_k \in K\}} \beta_k(t) e_k(\mathbf{x}).$$

Ainsi l'équation (2.27) s'écrit

$$d\alpha_t = -i\omega_r L_{\mathrm{GP}}^{3D}\alpha_t(\mathbf{x})dt + \frac{\gamma\hbar\omega_r}{k_BT}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}^{3D}\right)\alpha_t(\mathbf{x})dt + \sqrt{2\gamma}dW^{\mathcal{P}}(\mathbf{x}, t), \tag{2.29}$$

avec $L_{\mathrm{GP}}^{3D} = \frac{1}{\hbar\omega_r}\mathcal{P}L_{\mathrm{GP}}$ et $\omega_r = (\omega_x\omega_y)^{1/2}$.

Dans les cas où le confinement dans la direction z est plus important que dans les directions x et y, il est possible que seul l'état fondamental dans la direction z fasse partie de la phase condensée. Dans ce cas, le modèle tridimensionnel peut se ramener (de manière exacte) à un modèle bidimensionnel. Cette réduction de la dimension est détaillée dans la partie 2.4.2. Nous décrivons en Paragraphe 2.4.3 la procédure d'adimensionnement de l'équation (2.29) ou de sa formulation bidimensionnelle donnée en Paragraphe 2.4.2.

2.4.2 Réduction de la dimension pour SPGPE

Cette partie est dédiée à la description de la procédure de réduction de la dimension pour l'équation SPGPE, donnée par l'équation (2.29). Généralement, la valeur de E_{cut} est choisie de l'ordre de grandeur de 2μ-3μ, de sorte que les modes au dessus de ce niveau sont composés de très peu d'atomes. Lorsque la valeur de ω_z est assez grande, uniquement l'état fondamental dans la direction z fait partie de la région cohérente. Dans ce cas, il est possible de procéder à une réduction de la dimension. Cette affirmation vient du fait que les vecteurs propres de l'opérateur (2.28) sont tensorisés en $\mathbf{r} = (x, y)$ et z, c'est-à-dire que pour tout $k \in \mathbb{N}$, et pour tout $\mathbf{x} \in \mathbb{R}^3$, il existe deux entiers k_1 et k_2 tels que $k_1 + k_2 = k$ et $e_k(\mathbf{x}) = \mathrm{e}_{k_1}^r(\mathbf{r})\mathrm{e}_{k_2}^z(z)$, où l'on a défini par $(\mathrm{e}_k^r)_{k\in\mathbb{N}}$ et $(\mathrm{e}_k^z)_{k\in\mathbb{N}}$ respectivement les vecteurs propres des opérateurs,

$$-\frac{\hbar^2\Delta_r}{2m} + V(\mathbf{r}) - \Omega L_z, \quad \text{avec} \quad \Delta_r = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{2.30}$$

et

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{1}{2}m\omega_z^2 z^2, \tag{2.31}$$

ordonnés par croissance de leurs valeurs propres. Celles-ci sont notées respectivement $(\lambda_k^r)_{k\in\mathbb{N}}$ et $(\lambda_k^z)_{k\in\mathbb{N}}$. Avec ces notations, dire que seul l'état fondamental dans la direction z fait partie de la région cohérente signifie que pour tout entier k tel que $e_k \in K$,

le vecteur propre e_k se décompose sous la forme $e_k(\mathbf{x}) = \mathrm{e}_k^r(\mathbf{r})\mathrm{e}_0^z(z)$. Cette condition est équivalente à supposer que E_{cut} est tel que,

$$\lambda_0^r + \lambda_1^z > E_{\mathrm{cut}}. \tag{2.32}$$

Par ailleurs une condition suffisante pour que l'ensemble K soit non vide est donnée par,

$$\lambda_0 = \lambda_0^r + \lambda_0^z \leq E_{\mathrm{cut}}.$$

On notera également que les vecteurs propres de l'opérateur (2.31) sont les fonctions de Hermite, et que pour tout entier k, $\lambda_k^z = \frac{\hbar\omega_z}{2}(2k+1)$.

Dans le cas particulier où la condition (2.32) est satisfaite, les solutions $(\alpha_t(\mathbf{x}))_{t\geq 0}$ de l'équation (2.29) s'écrivent sous la forme tensorisée

$$\alpha_t(\mathbf{x}) = \alpha_t^r(\mathbf{r})\mathrm{e}_0^z(z),$$

ce qui permet de reformuler cette équation en une équation ne portant que sur la composante tensorielle en $\alpha_t^r(\mathbf{r})$. On on définit alors K_{2D} l'espace des solutions défini par,

$$K_{2D} = \mathrm{Span}\left\{\mathrm{e}_k^r; \lambda_k^r \leq E_{\mathrm{cut}} - \lambda_0^z\right\} = \mathrm{Span}\left\{\mathrm{e}_k^r; \lambda_k \leq E_{\mathrm{cut}}\right\}.$$

Pour ce faire, il suffit de considérer le produit scalaire en z seulement de l'équation (2.29) contre e_0^z. On obtient,

$$\begin{aligned}
\alpha_t^r(\mathbf{r}) &= \int_{\mathbb{R}} \mathrm{e}_0^z(z)\alpha_t(\mathbf{r},z)dz \\
&= \int_{\mathbb{R}} \mathrm{e}_0^z(z)\alpha_0(\mathbf{r},z)dz \\
&\quad + \int_0^t \int_{\mathbb{R}} \mathrm{e}_0^z(z)\left[-i\omega_r L_{\mathrm{GP}}^{3D}\alpha_t(\mathbf{r},z) + \frac{\gamma\hbar\omega_r}{k_B T}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}^{3D}\right)\alpha_t(\mathbf{r},z)\right]dzdt \\
&\quad + \sqrt{2\gamma}\int_{\mathbb{R}} \mathrm{e}_0^z(z)W((\mathbf{r},z),t)dz.
\end{aligned}$$

On définit l'opérateur L_{GP}^{2D} pour tout $\alpha^r \in K_{2D}$ par,

$$(L_{\mathrm{GP}}^{2D}\alpha^r)\otimes \mathrm{e}_0^z = \mathcal{P}L_{\mathrm{GP}}^{3D}(\alpha^r \otimes \mathrm{e}_0^z),$$

ou autrement dit,

$$L_{\mathrm{GP}}^{2D}\alpha^r(\mathbf{r}) = \int_{z\in\mathbb{R}} \mathrm{e}_0^z(z)\mathcal{P}\left\{\frac{1}{\hbar\omega_r}\left(-\frac{\hbar^2\Delta_r}{2m} + V(\mathbf{r}) + \frac{\hbar\omega_z}{2} - \Omega L_z + u\left|\alpha^r(\mathbf{r})\right|^2\left|\mathrm{e}_0^z(z)\right|^2\right)\alpha(\mathbf{r})\mathrm{e}_0^z(z)\right\},$$

où $\Delta_r = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. En observant que

$$\|\alpha_0^z\|_{L^2}^2 = 1, \quad \text{et,} \quad \|\alpha_0^z\|_{L^4}^4 = \frac{1}{\sqrt{2\pi}\sigma_z},$$

avec $\sigma_z = (\hbar/m\omega_z)^{1/2}$, on obtient,

$$L_{\mathrm{GP}}^{2D}\alpha^r = \mathcal{P}_{2D}\left\{\frac{1}{\hbar\omega_r}\left(-\frac{\hbar^2\Delta_r}{2m} + V(\mathbf{r}) + \frac{\hbar\omega_z}{2} - \Omega L_z + \frac{1}{\sqrt{2\pi}\sigma_z}u\left|\alpha^r(\mathbf{r})\right|^2\right)\alpha(\mathbf{r})\right\},$$

où $\mathcal{P}_{2D}$ est le projecteur orthogonal sur K_{2D}. L'équation réduite au cas 2 dimensionnel est alors donnée par,

$$d\alpha_t^r = -i\omega_r L_{\mathrm{GP}}^{2D}\alpha_t^r(\mathbf{r})dt + \frac{\gamma\hbar\omega_r}{k_BT}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}^{2D}\right)\alpha_t^r(\mathbf{r})dt + \sqrt{2\gamma}dW(\mathbf{r},t), \tag{2.33}$$

où le processus de Wiener $W(\mathbf{r},t)$ est donné par,

$$W(\mathbf{r},t) = \sum_{\{k;\mathrm{e}_k^r \in K_{2D}\}} \beta_k(t)\mathrm{e}_k^r(\mathbf{r}).$$

2.4.3 Adimensionnement pour SPGPE

Dans cette partie, L_{GP} désigne indifféremment les opérateurs L_{GP}^{2D} ou L_{GP}^{3D}. Pour adimensionner les équations (2.29) et (2.33), on commence par noter que la constante $\frac{k\gamma}{k_BT}$ est adimensionnelle. Ensuite nous procédons à l'adimensionnement des variables spatiales. Pour cela, nous définissons les longueurs caractéristiques dans les directions x, y et z respectivement par,

$$\sigma_x = \left(\frac{\hbar}{m\omega_x}\right)^{1/2}, \quad \sigma_y = \left(\frac{\hbar}{m\omega_y}\right)^{1/2}, \quad \sigma_z = \left(\frac{\hbar}{m\omega_z}\right)^{1/2}.$$

Nous définissons les variables spatiales adimensionnelles x, y, et z par,

$$\mathbf{x} = (\mathsf{x},\mathsf{y},\mathsf{z}) = \Phi(\mathbf{x}) = (x/\sigma_x, y/\sigma_y, z/\sigma_z).$$

Nous posons aussi $\omega_r = (\omega_x\omega_y)^{1/2}$ et $\boldsymbol{\varepsilon} = (\varepsilon_x,\varepsilon_y,\varepsilon_z) = \left(\frac{\omega_x}{\omega_r}, \frac{\omega_y}{\omega_r}, \frac{\omega_z}{\omega_r}\right)$.

On introduit la fonction d'onde $a_t(\mathbf{x})$ (et $a_t(\mathbf{r})$ dans le cas 2D) adimensionnée en espace,

$$a_t(\mathbf{x}) = \alpha_t \circ \Phi(\mathbf{x}), \quad a_t(\mathbf{r}) = \alpha_t^r \circ \Phi(\mathbf{r}),$$

Dans le cas 2D, nous prenons $\mathbf{r} = (\mathsf{x},\mathsf{y}) = \Phi(\mathbf{r}) = (x/\sigma_x, y/\sigma_y)$. Nous obtenons alors,

$$\frac{1}{\hbar\omega_r}\left(-\frac{\hbar^2\Delta}{2m} + V(\mathbf{x}) - \Omega L_z\right)\alpha_t(\mathbf{x}) = \left(\frac{1}{2}\boldsymbol{\varepsilon}\cdot(-\Delta_{\mathbf{x}} + \mathbf{x}^2) - \frac{\Omega}{\omega_r}L_{\mathsf{z}}\right)a_t(\mathbf{x}),$$

où $L_{\mathsf{z}} = i(\mathsf{x}\partial_{\mathsf{y}} - \mathsf{y}\partial_{\mathsf{x}})$, $\Delta_{\mathbf{x}} = (\frac{\partial^2}{\partial\mathsf{x}^2}, \frac{\partial^2}{\partial\mathsf{y}^2}, \frac{\partial^2}{\partial\mathsf{z}^2})$, et où $\boldsymbol{\varepsilon}\cdot(-\Delta_{\mathbf{x}} + \mathbf{x}^2)$ représente le produit scalaire

entre les vecteurs $\boldsymbol{\varepsilon}$ et $(-\Delta_{\mathbf{x}} + \mathbf{x}^2)$

Le nombre de particules N est donné par $\|\alpha_t\|_{L^2}^2$. Ainsi α_t est homogène à une distance à la puissance $-d/2$ en dimension d et est d'ordre $\sqrt{N}$. On définit $(\phi_t)_{t\geq 0}$ le processus adimensionnel d'ordre 1 par,

$$\phi_t = \delta_d^{d/2} N^{-1/2} a_t,$$

où δ_d est la distance caractéristique en dimension d. Pour $d = 2, 3$ on choisit,

$$\delta_3 = \sqrt{\frac{\hbar}{m(\omega_x\omega_y\omega_z)^{1/3}}}, \quad \text{et} \quad \delta_2 = \sqrt{\frac{\hbar}{m\omega_r}}.$$

On obtient,

$$L_{\mathrm{GP}}^{3D}\phi = \left(\frac{1}{2}\boldsymbol{\varepsilon}\cdot\left(-\Delta_{\mathbf{x}} + \mathbf{x}^2\right) - \frac{\Omega}{\omega_r}L_{\mathbf{z}} + 4\pi N a\sqrt{\frac{m\omega_z}{\hbar}}\,|\phi|^2\right)\phi,$$

et

$$L_{\mathrm{GP}}^{2D}\phi = \left(\frac{1}{2}\boldsymbol{\varepsilon}\cdot\left(-\Delta_{\mathbf{r}} + \mathbf{r}^2\right) + \frac{\varepsilon_z}{2} - \frac{\Omega}{\omega_r}L_{\mathbf{z}} + \sqrt{8\pi} N a\sqrt{\frac{m\omega_z}{\hbar}}\,|\phi|^2\right)\phi,$$

Les équations (2.29) et (2.33) deviennent,

$$d\phi_t = -i\omega_r L_{\mathrm{GP}}\phi_t dt + \frac{\gamma\hbar\omega_r}{k_B T}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}\right)\phi_t dt + \delta_d^{d/2} N^{-1/2}\sqrt{2\gamma}\,dW(t). \tag{2.34}$$

Le processus $\delta_d^{d/2} W(\mathbf{r}, t)$ constitue un processus de Wiener adimensionnel en espace, et l'on pose,

$$B(\mathbf{r}, t) = \delta_d^{d/2} W(\phi(\mathbf{r}), t).$$

L'équation (2.34) devient,

$$d\phi_t = -i\omega_r L_{\mathrm{GP}}\phi_t(\mathbf{r})dt + \frac{\gamma\hbar\omega_r}{k_B T}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}\right)\phi_t(\mathbf{r})dt + \sqrt{\frac{2\gamma}{N}}\,dB(t).$$

La dernière étape consiste à adimensionner la variable temporelle. On pose $\mathsf{t} = \omega_r t$, et on obtient,

$$d\phi_{\mathsf{t}} = -i\omega_r L_{\mathrm{GP}}\phi_{\mathsf{t}}(\mathbf{r})d\mathsf{t} + \frac{\gamma\hbar\omega_r}{k_B T}\left(\frac{\mu}{\hbar\omega_r} - L_{\mathrm{GP}}\right)\phi_{\mathsf{t}}(\mathbf{r})d\mathsf{t} + \sqrt{\frac{2\gamma}{N\omega_r}}\,dB(\mathsf{t}).$$

En notant que,

$$\sqrt{\frac{2\gamma}{N\omega_r}} = \frac{\gamma\hbar}{k_B T}\sqrt{\frac{\pi\hbar}{2N\rho\omega_r m a^2}},$$

nous obtenons finalement l'équation adimensionnelle suivante,

$$d\phi_{\mathtt{t}} = [-(i+c_1)(L_{\mathrm{GP}} - c_2 Id) - ic_2]\phi_{\mathtt{t}} d\mathtt{t} + c_1 c_3 dB_{\mathtt{t}},$$

avec,

$$c_1 = \frac{\gamma\hbar}{k_B T}, \quad c_2 = \frac{\mu}{\hbar\omega_r}, \quad c_3 = \sqrt{\frac{\pi\hbar}{2N\rho\omega_r m a^2}}.$$

Le terme $-ic_2 Id$ ne participe qu'à un déphasage global de la solution de ce modèle. Il peut donc être retiré de ce modèle à l'aide d'un changement de jauge, ce qui donne avec les mêmes constantes,

$$d\phi_{\mathtt{t}} = -(i+c_1)(L_{\mathrm{GP}} - c_2 Id)\phi_{\mathtt{t}} dt + c_1 c_3 dB_{\mathtt{t}}. \tag{2.35}$$

Des valeurs typiques expérimentales (voir par exemple [30]) correspondent à $T \sim 10^{-8}$K, et pour le rubidium 87 $m = 1.44 \cdot 10^{-25}$kg et $a = 5.77 \cdot 10^{-9}$m. En prenant de plus $\rho \sim 1$, nous obtenons alors $c_1 \sim 10^{-5}$. Il n'est pas étonnant d'obtenir une petite valeur pour ce coefficient car la dynamique de Langevin réversible est censée correspondre à une correction de l'équation de Gross-Pitaevskii dans le cas de petites températures non-nulles. En pratique cette très petite valeur pose des difficultés numériques qui seront discutées dans le Chapitre 6.

Première partie

Méthodes numériques pour la modélisation des fluctuations de l'intensité des lasers confinants

Chapitre 3

Numerical analysis of the Gross-Pitaevskii Equation with a randomly varying potential in time

The Gross-Pitaevskii equation with white noise perturbations in time of the harmonic potential is considered. In this chapter we define a Crank-Nicolson scheme based on a spectral discretisation and we show the convergence of this scheme in the case of focusing cubic nonlinearity with space dimension is equal to, or less than 3, and when the exact solution is uniquely defined and global in time. We prove that the strong order of convergence in probability is at least one.

This chapter corresponds to the preprint [142] "Numerical analysis of the Gross-Pitaevskii Equation with a randomly varying potential in time".

3.1 Introduction

The Gross-Pitaevskii equation (GPe) with cubic nonlinearity and quadratic potential is used to model the evolution of Bose-Einstein Condensate (BEC) macroscopic wave functions in an all-optical far-off-resonance laser trap. Fluctuations of laser intensity can be modelled by white noise in time perturbations of the harmonic potential ([1, 82]). More precisely, we will be interested in the following dimensionless equation:

$$\begin{cases} id\phi - \lambda \left|\phi\right|^{2\sigma} \phi dt - (-\Delta\phi + |x|^2\phi)dt = |x|^2\phi \circ dW_t, \\ \phi(0) = \phi_0, \end{cases} \tag{3.1}$$

where the unknown ϕ is a random field taking complex values, on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, depending on $t \geq 0$ and $x \in \mathbb{R}^d$. We take $\sigma > 0$ and $\lambda = \pm 1$. Here, $(W_t)_{t\geq 0}$ is a Brownian motion taking real values associated with the filtration $(\mathcal{F}_t)_{t\geq 0}$. The symbol $\circ$ denotes a Stratonovich product. Although the stochastic integral associated with this product does not verify the martingale property as Itô's product does, this choice of model

is natural since the noise is actually a *physical noise* which in the real physical case is not exactly white, but has small correlation length. A good way to understand this is to refer to [56] where the authors explain that Equation (3.1) can be seen in the subcritical case as a limit of equations where the Stratonovich product is replaced by a random process with nonzero correlation length.

Existence and uniqueness of a solution for (3.1) has been proven in [54] by the use of a compactness method in the case $d = 1, 2$ with restrictions on σ. The proof has then been generalized in [56] to the case $d \geq 3$ thanks to a dispersive estimate for the linear equation.

The use of a Crank-Nicolson scheme is natural since the mid-point discretisation of the noise is consistent with the Stratonovich product. For example such a discretisation has been used in [20, 51, 84]. Moreover, as this discretisation preserves the L^2-norm, it has also been used to discretize the deterministic equation (see [7, 13]) and nonlinear Schrödinger equations with additive noise (see [51]). Numerical simulations have provided good results of convergence for the Crank-Nicolson in similar contexts [7, 20, 83].

One of the main difficulties of this work is to prove stability for the naive Crank-Nicolson scheme in the linear case. Lack of stability is also encountered in [51], but actually comes from the non-linearity without truncation. Thus our case-study is more complicated. To estimate the growth of the solution between two time-steps, one can express the norm (in which we expect stability) of the solution at a given time-step with respect to the solution at the previous time-step. Such an equation is the discrete analogous to the Itô formula in the continuous case. However, this method yields some new terms that we do not know how to estimate, as explained in Remark 3.11. The solution we chose is to replace the unbounded operator $|x|^2$ by a bounded approximation such that we can control its operator norm with respect to $\delta t^{-1/4}$. This approximation is built using Hermite functions.

Another difficulty comes from the non-globally Lipschitz behaviour of the nonlinearity. We cannot ensure existence of a global solution for the discrete scheme by the standard use of a fixed point method. To circumvent this difficulty we use a classical argument (already used in [7, 10, 11, 20, 51, 161]) that consists in truncating the nonlinearity to make it globally Lipschitz, and then making the truncation level go to infinity. This argument leads to a weaker sense of convergence for the untruncated equation compared to the truncated equation. Indeed, we are only able to obtain a rate of convergence in probability instead of the mean-square sense. This kind of result is classical for globally non-Lipschitz non-linearities [20, 42, 118]. We also recall that even though the solution of Equation (3.1) belongs almost surely to some regular spaces, as stated by Proposition 3.1, we cannot prove that it has finite moments in these regular spaces. Thus it is delusive to hope to establish a strong order of convergence in these spaces (see Remark 3.2).

Besides, split-step methods have been used to approximate SPDEs and SDEs. For the deterministic Gross-Pitaevskii equation we refer to [14, 15]. One of the main benefit is that they offer an easy and consistent way to approximate the Stratonovich integral. In [118] the author proposes a splitting scheme to solve the deterministic nonlinear Schrödinger equation and the stochastic nonlinear Schrödinger equation with multiplicative noise of Stratonovich type. In both cases the nonlinearity is chosen to be non globally Lipschitz. For the deterministic scheme an order of convergence is given for small enough integration time intervals, and in the stochastic case an order of convergence in probability is given on a random time interval. The proof uses a truncation on the nonlinearity. Moreover, in [71], the authors propose a Lie time-splitting scheme for a nonlinear partial differential equation driven by a random time-dependent dispersion coefficient. In this case the nonlinearity is supposed to be Lipschitz, but the dispersion coefficient can approximate a fractional Brownian motion.

Let us introduce some notation. We define for $k \in \mathbb{N}$ and $p \in \mathbb{N}^*$,

$$\Sigma^{k,p}(\mathbb{R}^d) = \left\{ v \in L^p(\mathbb{R}^d),\ \sum_{|\alpha|+|\beta|\leq k} \left\| x^\beta \partial^\alpha v \right\|_{L^p}^p = \|v\|_{\Sigma^{k,p}}^p < +\infty \right\}. \tag{3.2}$$

We simply denote Σ^k instead of $\Sigma^{k,2}$, and Σ instead of Σ^1. We denote by $C(\cdot)$ the constants, specifying dependencies with $\cdot$. These constants can vary from line to line. We denote by $\langle \cdot, \cdot \rangle$ the usual Hermitian inner product in L^2, defined for all $u, v \in L^2(\mathbb{R}^d)$ by $\langle u, v \rangle = \int_{\mathbb{R}^d} u(x)\overline{v}(x)dx$. We denote by A the operator $-\Delta + |x|^2$. In dimension $d = 1$, we recall that this operator has purely discrete eigenvalues $\lambda_k = 2k + 1$, with $k \in \mathbb{N}$. The corresponding eigenfunctions are the Hermite functions, denoted by $\{e_k, \quad k \in \mathbb{N}\}$, that form a complete orthonormal system in $L^2(\mathbb{R})$. In higher dimensions, the tensorization of this basis also form a complete orthonormal system in $L^2(\mathbb{R}^d)$. For example with $d = 2$, we get the basis $\{e_{k,j} := (x, y) \mapsto e_k(x)e_j(y), \quad k, j \in \mathbb{N}\}$, and for $k, j \in \mathbb{N}$, the eigenvalue of $e_{k,j}$ is $2(k + j) + 2$. Moreover the sesquilinear form defined on $\Sigma^k(\mathbb{R}^d) \times \Sigma^k(\mathbb{R}^d)$ by $(u, v) \mapsto \langle A^k u, v \rangle$ defines a Hermitian inner product on Σ^k, and the associated norm is equivalent to the norm defined in (3.2). For $K \in \mathbb{N}$, we denote by $\Sigma_K(\mathbb{R})$ the finite dimensional vector space spanned by the K first Hermite functions. In the multidimensional case we denote by $\Sigma_K(\mathbb{R}^d)$ the finite dimensional subspace of $\Sigma(\mathbb{R}^d)$, of dimension K^d, obtained by tensorisation of the space $\Sigma_K(\mathbb{R})$ (see [147]). We denote by P_K the orthogonal projection of $L^2(\mathbb{R}^d)$ onto $\Sigma_K(\mathbb{R}^d)$, and we set $A_K = AP_K$.

We now give some notation about the numerical scheme. Let $T > 0$ be the time horizon, and $N \in \mathbb{N}^*$ be the number of time discretisation steps. We denote by $\delta t = T/N$ the time step. We also set $t_n := n\delta t$ for $n \leq N$, and for all $n \in \mathbb{N}$, $\chi^{n+1} = \delta t^{-1/2}(W_{(n+1)\delta t} - W_{n\delta t})$. Thus $(\chi^n)_{n\in\mathbb{N}^*}$ is a sequence of independent standard normal deviates. Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be the filtration defined for all $n \in \mathbb{N}$ by $\mathcal{F}_n = \sigma(W_s, s \leq t_n)$, so that for all $n \in \mathbb{N}^*$ χ^n is $\mathcal{F}_n$-measurable.

In Section 3.2 we define the numerical scheme. We introduce truncations of the Brownian motion increments (that will enable us to control the growth of the solution of the scheme), and the non-linearity. We also introduce an approximation of the term $|x|^2$, that will enable us to prove stability. In Section 3.3 we prove existence and uniqueness of a solution for the numerical scheme. We also prove stability. The stability heavily relies on the spectral discretisation. We then prove in Section 3.4 the convergence in probability of our scheme toward the GP equation, with a strong order at least one. We finally present in Section 3.5 simulation results for the numerical scheme we suggest, and compare its accuracy with various other schemes.

3.2 Definition of the numerical scheme and main result

This section is devoted to the construction of the numerical scheme and to the statement of the main convergence result. The technical precisions are quite cumbersome and interfere with the comprehension of the general idea. For the sake of clarity we first present a refined version of the main statements in section 3.2.1, introducing the main ideas and the technical issues. We then give in Section 3.2 a rigorous definition of the scheme that enables us to state the convergence result.

3.2.1 Objectives and formal result

As stated in Section 3.1, Equation (3.1) is known to be well-posed under some assumptions on λ, σ and d (see [54, 56, 57]). We state a slightly stronger result in the specific case $\lambda = 1$, $\sigma = 1$ and $d \leq 3$ for arbitrarily regular initial condition. All the results of this article hold only in this specific case since we require the following result.

Proposition 3.1. *Assume $\lambda = 1$, $\sigma = 1$ and $d \leq 3$. Then $\forall j \in \mathbb{N}^*$, if $\phi_0 \in \Sigma^j(\mathbb{R}^d)$ then there exists a unique global solution ϕ of (3.1) adapted to $(\mathcal{F}_t)_{t\geq 0}$, with $\phi(0) = \phi_0$, almost surely in $C(\mathbb{R}^+; \Sigma^j(\mathbb{R}^d))$.*

The case $j = 1$ and $\sigma < 2$ if $d = 3$ or $\sigma < +\infty$ if $d = 2$, has been initially partially treated in [54] and generalized in [56]. Then, the case $j = 2$, $d = 2$ and $\sigma \geq 1/2$ has been treated in [57]. The proof is given in Appendix and is done recursively using the idea introduced in [57]. This extra regularity is needed in order to show an order of convergence of the numerical scheme (that will be introduced thereafter) toward the solution of Equation (3.1).

Remark 3.2. *Even though Proposition 3.1 states that the solution of Equation 3.1 has regular trajectories, we were not able to prove that the solution has finite moments. For instance we could not prove that the solution ϕ given by Proposition 3.1 belongs to $L^\infty([0,T], L^2(\Omega, \Sigma^k(\mathbb{R}^d)))$ for $k > 1$ and for all $T > 0$. Thus, we cannot hope to show an order of convergence in the mean-square sense in the space $\Sigma^1(\mathbb{R}^d)$ since this kind of results requires more regularity.*

The naive Crank-Nicolson discretisation of Equation (3.1) would be given by,

$$\phi^{n+1} - \phi^n = -i\left(\delta t A + \sqrt{\delta t}\chi^{n+1}|x|^2\right)\left(\frac{\phi^{n+1}+\phi^n}{2}\right) - i\lambda\delta t g(\phi^{n+1},\phi^n), \tag{3.3}$$

where g is a classical approximation of the non-linearity [60] given by

$$g(\phi^{n+1},\phi^n) = \frac{1}{\sigma+1}\left(\frac{\left|\phi^{n+1}\right|^{2\sigma+2} - |\phi^n|^{2\sigma+2}}{|\phi^{n+1}|^2 - |\phi^n|^2}\right)\left(\frac{\phi^{n+1}+\phi^n}{2}\right), \tag{3.4}$$

which can be simplified (to get rid of the singularity) in our case $\sigma = 1$ by

$$g(\phi^{n+1},\phi^n) = \frac{1}{2}\left(|\phi^n|^2 + \left|\phi^{n+1}\right|^2\right)\left(\frac{\phi^{n+1}+\phi^n}{2}\right).$$

The scheme (3.3) can be formally written in the following form

$$\phi^{n+1} = S_{\delta t,n+1}\phi^n - i\lambda\delta t T^{-1}_{\delta t,n+1} g(\phi^{n+1},\phi^n), \tag{3.5}$$

where

$$T_{\delta t,n+1} = \left(Id + i\frac{\delta t}{2}A + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}|x|^2\right),$$
$$S_{\delta t,n+1} = \left(Id + i\frac{\delta t}{2}A + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}|x|^2\right)^{-1}\left(Id - i\frac{\delta t}{2}A - i\frac{\sqrt{\delta t}}{2}\chi^{n+1}|x|^2\right).$$

Let us suppose that this scheme is well-defined in the following sense.

Claim 3.3. *There exists a unique discrete solution $\mathcal{F}_n$-adapted $\phi = (\phi^n)_{n=0,\dots,N}$ satisfying (3.3) which belongs a.s. to $L^\infty(0,T;\Sigma^j)$.*

Then we want to establish a convergence result in probability for this discrete solution towards the solution of Equation (3.1):

Claim 3.4. *For all $k \in \mathbb{N}^*$, $\phi_0 \in \Sigma^{k+12}$ and $\alpha < 1$, there exists $C > 0$ such that*

$$\lim_{\delta t\to 0}\sup_{n\delta t\le T}\mathbb{P}\left(\|\phi^n - \phi(t_n)\|_{\Sigma^k} > C\delta t^\alpha\right) = 0,$$

where $(\phi(t))_{t\in[0,T]}$ denotes the solution of (3.1) given by Proposition 3.1.

The respective rigorous analogous of these claims are Proposition 3.9 and Theorem 3.14. We aim to show a convergence in probability, and not a stronger convergence, as mean-square for example, because the nonlinearity is not globally Lipschitz. To show this result, we actually prove an intermediate result of convergence in the mean-square sense when the nonlinearity is approximated by a Lipschitz truncation. This approximation has

been used in the deterministic setting to establish error bounds [10, 11, 161]. We lose the strong convergence by making the level of truncation go to infinity. Similar methods have been used in [20, 42, 51]. We refer to Remark 3.15 for comment about the high regularity assumed on the initial condition.

We did not manage to prove these two claims for the scheme given by (3.5). The main issues are linked to the invertibility of the operator $T_{\delta t,n+1}$, and to the stability of the linear equation. The goal of Section 3.2.2 is to slightly modify the operators $T_{\delta t,n+1}$ and $S_{\delta t,n+1}$ in order to be able to prove Claims 3.3 and 3.4. More precisely these modifications should take into account that for strictly positive values of j, the invertibility of the operators $T_{\delta t,n+1}$ when $\chi^{n+1}\delta t$ is below a certain negative constant is not clear. Indeed, for all $\phi \in \Sigma^1(\mathbb{R}^d)$, $\langle T_{\delta t,n+1}\phi, \phi\rangle_{L^2(\mathbb{R}^d)} = \|\phi\|_{L^2}$, which enables us to prove the injectivity of $T_{\delta t,n+1}$. Its surjectivity can be shown by the Lax-Milgram theorem when $\chi^{n+1}\delta t$ is greater than a negative constant. Thereafter we propose to circumvent this difficulty by truncating the increments of the Brownian motion. Nevertheless this modification is not sufficient to enable us to prove Claim 3.4. In order to be able to prove it, we introduce thereafter an approximating sequence $(B_K)_{K\in\mathbb{N}^*}$ of the operator $|x|^2$, composed by bounded operators in every Σ^l for $l \in \mathbb{N}$. By replacing $|x|^2$ by B_K, we are then able to prove the stability. The next subsection defines rigorously the modifications on the operators $T_{\delta t,n+1}$ and $S_{\delta t,n+1}$ described above in order to show a convergence in probability toward Equation (3.1).

3.2.2 Rigorous definition of the scheme

In all the following $k \in \mathbb{N}$ refers to the space $\Sigma^k(\mathbb{R}^d)$ in which convergence of the numerical scheme is expected, and $j \in \mathbb{N}$ refers to the space $\Sigma^j(\mathbb{R}^d)$ in which stability is expected. To be able to prove an order of convergence, we require j to be greater than k. More precisely, our convergence theorem will require $k > d/2$ and $j = k + 12$.

We first introduce some objects. For all $K \in \mathbb{N}^*$, we consider B_K, a self-adjoint operator in L^2 that we will specify later. It is essentially an approximation of the operator $|x|^2$ which is a bounded operator in every Σ^j for $j \in \mathbb{N}$. We furthermore suppose that for all $K \in \mathbb{N}^*$, B_K takes its values into $\Sigma_K(\mathbb{R}^d)$. This way, the scheme will take values into $\Sigma_K(\mathbb{R}^d)$, which makes it implementable. Yet, the fact that the numerical scheme takes values in the finite-dimensional space $\Sigma_K(\mathbb{R}^d)$ is not an essential assumption. We could have only looked at a semi-discrete scheme in time, and considered a truncation of the operator $|x|^2$, instead of relying on a space discretisation.

We now set θ and θ_j, regular truncations of the Heaviside function such that $\theta \in \mathcal{C}^\infty(\mathbb{R}^+), \theta \geq 0$, $\operatorname{Supp}\theta \subset [0,2]$ and $\theta \equiv 1$ on $[0,1]$. For all $L \in \mathbb{N}^*$ and all $x \in \mathbb{R}^+$ we also set $\theta_L(x) = \theta(x/L)$. We introduce f_L and g_L, two Lipschitz approximations of the

non-linearity defined by,

$$\forall \phi \in \Sigma^k, \quad g_L^k(\phi_1, \phi_2) = \theta_L(\|\phi_1\|^2_{\Sigma^k})\theta_L(\|\phi_2\|^2_{\Sigma^k})\frac{1}{2}\left(|\phi_1|^2 + |\phi_2|^2\right)\left(\frac{\phi_1+\phi_2}{2}\right), \tag{3.6}$$

$$\forall \phi \in \Sigma^k, \quad f_L^k(\phi) = g_L^k(\phi, \phi). \tag{3.7}$$

For all $C_0 > 0$, we introduce the $\mathcal{F}_n$-measurable stopping time $\tau_{\delta t, C_0}$ defined by,

$$\tau(\delta t, C_0) = \min\left(\left\{n \in \mathbb{N}, 1 \leq n \leq N, \left|W_{t_n} - W_{t_{n-1}}\right| \geq C_0\right\} \cup \{N+1\}\right). \tag{3.8}$$

We introduce this stopping time for technical purposes. It prevents the most shaken up trajectories of the Brownian motion to downgrade the regularity of the discrete solution. In the regime where δt is small, $\tau(\delta t, C_0) = N+1$ with high probability, as stated by the following Lemma. It shows that the case $\tau(\delta t, C_0) < N+1$ happens with small probability when δt vanishes and thus corresponds to pathological cases.

Lemma 3.5. *For all $C_0 > 0$, there exists $R_0 > 0$ such that*

$$\mathbb{P}\left(\tau(\delta t, C_0) < N+1\right) \leq C_0 N e^{-\frac{C_0^2 N}{2T^2}}.$$

Proof of Lemma 3.5. The domination comes from the fact that

$$\begin{aligned}\mathbb{P}\left(\tau(\delta t, C_0) < N+1\right) &\leq \sum_{1\leq n\leq N} \mathbb{P}\left(\left|W_{t_n} - W_{t_{n-1}}\right| \geq C_0\right) \\ &\leq N\mathbb{P}\left(|G| \geq \frac{C_0\sqrt{N}}{\sqrt{T}}\right),\end{aligned}$$

where G denotes a standard normal deviate. The result comes from the fact that for all $x > 0$, $\mathbb{P}\left(|G| \geq x\right) \leq 2e^{-x^2/2}$. □

We define the random operators $T^{-1}_{\delta t, K, n+1, C_0}$ and $S_{\delta t, K, n+1, C_0}$ on $\Sigma_K(\mathbb{R}^d)$ for $n \leq N$ by,

$$T^{-1}_{\delta t, K, n+1, C_0} = \begin{cases} \left(Id + i\frac{\delta t}{2}A_K + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K\right)^{-1}, & \text{if} \quad n+1 < \tau_{C_0, \delta t} \\ 0, \quad \text{otherwise} \end{cases} \tag{3.9}$$

$$S_{\delta t, K, n+1, C_0} = \begin{cases} T^{-1}_{\delta t, K, n+1, C_0}\left(Id - i\frac{\delta t}{2}A_K - i\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K\right), & \text{if} \quad n+1 < \tau_{C_0, \delta t}, \\ Id, \quad \text{otherwise.} \end{cases} \tag{3.10}$$

The invertibility of the operator $\left(Id + i\frac{\delta t}{2}A_K + i\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K\right)$ follows from the fact that B_K will be chosen symmetric (see Assumption 3.6), and thus this operator is a skew-

symmetric pertubation of the identity for the Hermitian inner product in L^2, which implies that it is injective, and thus bijective since $\Sigma_K(\mathbb{R}^d)$ is finite-dimensional.

We now present our Crank-Nicolson scheme, with a spectral discretisation in space such that it takes values in the space $\Sigma_K(\mathbb{R}^d)$. We define the family of stochastic processes $(\phi^n_{K,L,\delta t})$, indexed by $K \in \mathbb{N}^*$, $L \in \mathbb{N}^*$ and all possible values of δt, as,

$$\begin{cases} \phi^{n+1}_{K,L,\delta t} = S_{\delta t,K,n+1,C_0}\phi^n_{K,L,\delta t} - i\delta t\lambda T^{-1}_{\delta t,K,n+1,C_0} P_K g^k_L(\phi^n_{K,L,\delta t}, \phi^{n+1}_{K,L,\delta t}), \\ \phi^0_{K,L,\delta t} = P_K\phi_0. \end{cases} \tag{3.11}$$

Existence and uniqueness of a solution for (3.11) is not obvious since this equation is non-linearly implicit. We show in the next section that it is actually well-posed. The approximation of the non-linearity is classical, see [20, 51, 60], and comes from the fact that it is conservative for the energy in the deterministic case (without truncation). Moreover, we emphasize that the initial condition ϕ_0 will always be deterministic in this article, and so will be the initial condition $\phi^0_{K,L,\delta t}$ of the numerical scheme.

We now specify how to choose the family of operators $(B_K)_{K\in\mathbb{N}^*}$. We require it to satisfy the following assumptions where we denote by $|||\cdot|||_{\mathcal{L}(\Sigma^m,\Sigma^n)}$ the operator norm for the linear operators from Σ^m to Σ^n. We also denote by $[\cdot,\cdot]$ the commutator of two operators. It is defined for all operators U, V by $[U, V] = UV - VU$.

Assumption 3.6. *The family $(B_K)_{K\in\mathbb{N}^*}$ is such that for all $k \in \mathbb{N}$, there exists $C(k) > 0$ for which,*

$$B_K = B_K^*, \quad \forall K \in \mathbb{N}^*, \tag{3.12}$$

$$|||B_K|||_{\mathcal{L}(\Sigma^{k+2},\Sigma^k)} \leq C(k), \quad \forall K \in \mathbb{N}^*, \tag{3.13}$$

$$|||B_K|||_{\mathcal{L}(\Sigma^k,\Sigma^k)} \leq C(k)K, \quad \forall K \in \mathbb{N}^*, \tag{3.14}$$

$$|||[A^k, B_K]|||_{\mathcal{L}(\Sigma^{2k},L^2)} \leq C(k), \quad \forall K \in \mathbb{N}^*, \tag{3.15}$$

$$\forall \phi \in \Sigma^k, \quad \left|\Re\langle [A^k, B_K]\phi, B_K\phi\rangle\right| \leq C(k)\left\|\phi\right\|^2_{\Sigma^k}, \quad \forall K \in \mathbb{N}^*. \tag{3.16}$$

Assumption 3.7. *We suppose that there exists $C > 0$ such that for all $K, k, p \in \mathbb{N}$, and for all $\phi \in \Sigma^{k+p+2}(\mathbb{R}^d)$,*

$$\left\|(|x|^2 - B_K)\phi\right\|^2_{\Sigma^k} \leq CK^{-p}\left\|\phi\right\|^2_{\Sigma^{k+p+2}}. \tag{3.17}$$

Assumption 3.6 is required for proving the stability of the numerical scheme. Assumption 3.7 makes precise the meaning of the convergence of the sequence B_K toward the operator $|x|^2$. Properties (3.12), (3.13), (3.15) and (3.16) are satisfied by the operator $|x|^2$, and are natural to impose to the sequence (B_K), for technical purposes. The essential difference between (B_K) and $|x|^2$ stated in (3.14) is that the B_K are not unbounded operators from Σ^k to Σ^k, in contrast to $|x|^2$. This is why we expect $|||B_K|||_{\mathcal{L}(\Sigma^k,\Sigma^k)}$ to go to $+\infty$ since B_K converges to $|x|^2$ when K increases. Moreover (3.14) actually controls the

speed at which this operator norm diverges. This speed comes from the fact that there exists $C > 0$ such that for all $K, k, p \in \mathbb{N}$ and for all $\phi \in \Sigma^{k+p}(\mathbb{R}^d)$,

$$\|P_K \phi\|^2_{\Sigma^{k+p}} \leq C K^p \|\phi\|^2_{\Sigma^k} . \tag{3.18}$$

Assumption 3.7 is natural since there exists $C > 0$, such that for all $k, p \in \mathbb{N}$, for all $K \in \mathbb{N}$ and for all $\phi \in \Sigma^{k+p}(\mathbb{R}^d)$,

$$\|(Id - P_K)\phi\|^2_{\Sigma^k} \leq C K^{-p} \|\phi\|^2_{\Sigma^{k+p}} , \tag{3.19}$$

and for all $\phi \in \Sigma^{k+p+2}(\mathbb{R}^d)$,

$$\left\|(Id - P_K) |x|^2 \phi\right\|^2_{\Sigma^k} \leq C K^{-p} \|\phi\|^2_{\Sigma^{k+p+2}} . \tag{3.20}$$

Equations (3.19) and (3.20) follow from a direct computation based on the expansion of ϕ in eigenvectors of A.

To satisfy hypotheses (3.14) and (3.17), one could think of choosing $B_K = P_K |x|^2 P_K$, but then (3.16) would not hold anymore. To satisfy all hypotheses, we can define B_K as a smooth (actually Lipschitz) spectral cutoff of $|x|^2$, also taking values in $\Sigma_K(\mathbb{R}^d)$, instead of a cutoff such as $P_K |x|^2 P_K$. We now give an example of such a family of operators. We begin by constructing it in dimension one, and we then generalize to higher dimensions. We recall that in dimension one, the operator $|x|^2$ satisfies,

$$\forall m \in \mathbb{N}, \quad |x|^2 e_m = \frac{1}{2}\left(\sqrt{(m-1)m} e_{m-2} + (2m+1) e_m + \sqrt{(m+1)(m+2)} e_{m+2}\right).$$

This is why we want (B_K) to be of the following form.

$$\forall m \in \mathbb{N}, \quad B_K e_m = \alpha_m^K e_{m-2} + \beta_m^K e_m + \gamma_m^K e_{m+2}. \tag{3.21}$$

We fix $\theta \in (0,1)$, the parameter that specifies the smoothness of the cutoff, and set for all $m \in \mathbb{N}$, and for all $K \in \mathbb{N}^*$,

$$\alpha_m^K = \begin{cases} \dfrac{1}{2}\sqrt{(m-1)m} = \langle |x|^2 , e_{m-2} \rangle, & \text{if} \quad 2 \leq m \leq \theta K \\ \alpha^K_{\lfloor \theta K \rfloor} \left(1 - \left(\dfrac{m - \lfloor \theta K \rfloor}{K - \lfloor \theta K \rfloor}\right)\right), & \text{if} \quad \theta K < m \leq K \\ 0, \quad \text{otherwise} \end{cases} \tag{3.22}$$

$$\beta_m^K = \begin{cases} \dfrac{1}{2}(2m+1) = \langle |x|^2 , e_m \rangle, & \text{if} \quad 0 \leq m \leq \theta K \\ \beta^K_{\lfloor \theta K \rfloor} \left(1 - \left(\dfrac{m - \lfloor \theta K \rfloor}{K - \lfloor \theta K \rfloor}\right)\right), & \text{if} \quad \theta K < m \leq K \\ 0, \quad \text{otherwise} \quad K < m \end{cases} \tag{3.23}$$

$$\forall m \in \mathbb{N}, \forall K \in \mathbb{N}^*, \quad \gamma_m^K = \alpha_{m+2}^K. \tag{3.24}$$

We recall that Equation (3.24) implies that B_K is symmetric. The extension to the d-dimensional case, with $d \geq 2$, is done by setting for $i_1, i_2, \ldots, i_d \leq K$,

$$B_K\left(\bigotimes_{k=1}^{d} e_{i_k}\right) = \sum_{j=1}^{d} \bigotimes_{k=1}^{d} \left(\mathbb{1}_{\{j=k\}} B_K + \mathbb{1}_{\{j\neq k\}} Id\right) e_{i_k}. \tag{3.25}$$

We refer to [147] for basics about tensorisation. For example, with $d = 2$, $B_K(e_{i_1} \otimes e_{i_2}) = (B_K e_{i_1}) \otimes e_{i_2} + e_{i_1} \otimes (B_K e_{i_2})$. This extension is natural since it holds for the operator $|x|^2$.

In this paper, stability is proved by an extensive use of the boundedness of the operators B_K. Convergence toward the solution of (3.1) will be achieved by making δt go to zero and K and L to infinity. We will show that stability holds if K does not tend to infinity faster than $\delta t^{-1/4}$, and that convergence of order at least one in probability holds if K actually scales like $\delta t^{-1/4}$.

3.3 Well-posedness of the numerical scheme

We begin by explaining how to choose C_0 in the definition 3.8 of $\tau(\delta t, C_0)$. We require in the following the operators $S_{\delta t,K,n+1,C_0}$ and $T^{-1}_{\delta t,K,n+1,C_0}$ to be uniformly bounded in K. It is possible to choose C_0 to satisfy this requirement as stated by the following lemma:

Lemma 3.8. *Under Assumption 3.6, for all $j \in \mathbb{N}$, there exists $C_0(j) > 0$ such that for all $K \in \mathbb{N}^*$, $\phi \in \Sigma_K(\mathbb{R}^d)$, $\delta t \in (0,1)$ and $n \in \mathbb{N}$,*

$$\left\|S_{\delta t,K,n+1,C_0(j)}\phi\right\|_{\Sigma^j}^2 \leq 2\left\|\phi\right\|_{\Sigma^j}^2, \quad a.s.$$
$$\left\|T^{-1}_{\delta t,K,n+1,C_0(j)}\phi\right\|_{\Sigma^j}^2 \leq 2\left\|\phi\right\|_{\Sigma^j}^2, \quad a.s.$$

From now on, we suppose that $C_0(j)$ satisfies this lemma for a given j. We simply denote by $S_{\delta t,K,n+1,j}$ and $T^{-1}_{\delta t,K,n+1,j}$ the operators $S_{\delta t,K,n+1,C_0(j)}$ and $T^{-1}_{\delta t,K,n+1,C_0(j)}$. This technical difficulty (of truncating the Brownian increments) did not appear in [51] where the noise is multiplicative. The difference comes from the fact that in our case the potential reduces the regularity of the noise term. One can note that this truncation on the noise would have been also necessary in order to prove Claim 3.3 for operators $T_{\delta t,n+1}$ and $S_{\delta t,n+1}$ as stated before.

Proof of Lemma 3.8. Let $\phi \in \Sigma_K(\mathbb{R}^d)$, $n \in \mathbb{N}$ and $C_0 > 0$. Then, to simplify notation, we set $\psi = S_{\delta t,K,n+1,C_0}\phi$. Thus, for almost every $\omega \in \{\tau(\delta t) \leq n+1\}$, we have $\psi = \phi$ by Equation (3.10), where $\tau(\delta t)$ is defined by (3.8), and there is nothing to prove. Then, we

suppose that $\omega \in \{\tau(\delta t) > n+1\}$, which implies that $\sqrt{\delta t}\left|\chi^{n+1}\right| < C_0$. The expression of $S_{\delta t,K,n+1,C_0(j)}$ enables to write,

$$\psi = \phi - i\left(A_K \delta t + B_K \sqrt{\delta t}\chi\right)\left(\frac{\phi+\psi}{2}\right). \tag{3.26}$$

Then, by taking the L^2 inner product of both sides of this equation with $A^j(\phi+\psi)$, one finds the following implicit expression of the growth of the solution :

$$\begin{aligned}\|\psi\|^2_{\Sigma^j} - \|\phi\|^2_{\Sigma^j} =& \Re\langle A^j(\phi+\psi), \psi-\phi\rangle \\ =& -\frac{\sqrt{\delta t}}{2}\chi\Im\langle A^j(\phi+\psi), B_K(\phi+\psi)\rangle.\end{aligned}$$

Then, using first (3.12), and then (3.15),

$$\|\psi\|^2_{\Sigma^j} - \|\phi\|^2_{\Sigma^j} = \frac{\sqrt{\delta t}}{4}\chi\Im\langle [A^j, B_K](\phi+\psi), \phi+\psi\rangle \tag{3.27}$$

$$\leq \sqrt{\delta t}\,|\chi|\,C(j)\left(\|\psi\|^2_{\Sigma^j} + \|\phi\|^2_{\Sigma^j}\right). \tag{3.28}$$

The last line exhibits the reason why we require the truncation on the Gaussian increment. Indeed, since $\omega \in \{\tau(\delta t) > n+1\}$, we get that

$$\|\psi\|^2_{\Sigma^j} - \|\phi\|^2_{\Sigma^j} \leq C_0 C(j)\left(\|\psi\|^2_{\Sigma^j} + \|\phi\|^2_{\Sigma^j}\right), \tag{3.29}$$

and thus by choosing $C_0(j) = \frac{1}{3C(j)}$, we get that

$$\|\psi\|^2_{\Sigma^j} \leq 2\,\|\phi\|^2_{\Sigma^j}. \tag{3.30}$$

The estimate on $T^{-1}_{\delta t,K,n+1,C_0(j)}$ is proven in a similar way. □

The next proposition states the well-posedness of the numerical scheme defined by (3.11). In this proposition we consider K as a function of δt, so that K may tend to infinity when δt tends to zero. Since the non-linearity is implicit, well-posedness is not immediate. It is classically shown by the use of a Banach fixed point theorem, which can be easily implemented. Moreover, this proof relies on the Lipschitz truncation of the nonlinearity.

Proposition 3.9. *Under Assumption 3.6, for all $k > d/2$, $j \geq k$, $L \in \mathbb{N}^*$, $T > 0$ and $K_0 > 0$, there exist $\delta t_0(j, L, K_0) > 0$ and $C(j, L, K_0, T) > 0$ such that for all $\phi_0 \in \Sigma^j(\mathbb{R}^d)$, for all $\delta t = T/N \leq \delta t_0(j, L, K_0)$, and $K \leq K_0 \delta t^{-1/4}$, there exists a unique discrete solution $(\mathcal{F}_n)$-adapted $\phi = (\phi^n)_{n=0,\dots,N}$ satisfying (3.11), almost surely in $L^2_\omega L^\infty([0,T];\Sigma^j(\mathbb{R}^d))$,*

and such that

$$\mathbb{E}\left[\sup_{n\leq N}\|\phi^n\|_{\Sigma^j}^2\right] \leq C(j,L,K_0,T)\,\|\phi_0\|_{\Sigma^j}^2\,.$$

It is important to notice that the norm in which we truncate the nonlinearity is not linked to the space in which we prove existence and uniqueness, but to the space in which we will prove the convergence of the numerical scheme. The fact that ϕ belongs a.s. in $L^\infty(0,T;\Sigma^j(\mathbb{R}^d))$ is understood supposing ϕ constant on every interval $[n\delta t,(n+1)\delta t[$ for $0\leq n\leq N-1$.

The most technical part in the proof of this proposition lies in showing the stability in Σ^j for the linear part of the scheme, given by

$$\begin{cases}\phi^{n+1} = S_{\delta t,K,n+1,C_0}\phi^n,\\ \phi^0 = P_K\phi_0.\end{cases} \tag{3.31}$$

Stability of the linear equation (3.31) is a main issue in this article. We recall that we were not able to prove stability for the semi-discrete Crank-Nicolson scheme defined by Equation (3.5). The proof we present uses extensively the boundedness of operators B_K, and the truncation of the noise. The next lemma states the stability result for Equation (3.31).

Lemma 3.10. *For all $j\in\mathbb{N}$, $T>0$, $K\in\mathbb{N}^*$ and $N\in\mathbb{N}^*$ such that $\delta t = T/N$, there exists a constant $C(j,T,K^4\delta t)>0$ such that for all $\phi_0\in\Sigma^j(\mathbb{R}^d)$,*

$$\mathbb{E}\left[\sup_{n\leq N}\|\phi^n\|_{\Sigma^j}^2\right] \leq C(j,T,K^4\delta t)\,\left\|\phi^0\right\|_{\Sigma^j}^2\,, \tag{3.32}$$

where $(\phi^n)_{n\leq N}$ denotes the solution of the linear scheme given by (3.31).

This Lemma enables to show the stability for the linear equation on a fixed interval $[0,T]$, scaling K as $\delta t^{-1/4}$.

Remark 3.11. *The proof of Lemma 3.10 relies on the discrete analogous of the Itô lemma for the continuous process $(\phi_t)_{t\geq 0}$ solution of the linear part of (3.1),*

$$d\phi = -iA\phi dt - i|x|^2\phi\circ dW_t.$$

Indeed, to show the stability of the numerical scheme, the idea is to compute the variation of the $\Sigma^j(\mathbb{R}^d)$ norm of the solution of the numerical scheme on one time step. The computation

is done in the proof of Lemma 3.10 (see Equation (3.40)) and gives,

$$\begin{aligned}\left\|\phi^{n+1}\right\|_{\Sigma^j}^2 - \|\phi^n\|_{\Sigma^j}^2 =& \sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle[A^j,|x|^2]\phi^n,\phi^n\rangle \\ &+ \frac{\sqrt{\delta t}}{2}\widetilde{\chi}^{n+1}\Im\langle[A^j,|x|^2](\phi^{n+1}+\phi^n),(\phi^{n+1}-\phi^n)\rangle \\ &- \frac{\sqrt{\delta t}}{4}\widetilde{\chi}^{n+1}\Im\langle[A^j,|x|^2](\phi^{n+1}-\phi^n),(\phi^{n+1}-\phi^n)\rangle,\end{aligned} \tag{3.33}$$

where $\widetilde{\chi}^{n+1}$ is defined by Equation (3.38). For the time continuous process, a similar computation using Itô's lemma gives,

$$\left\|\phi_{t_{n+1}}\right\|_{\Sigma^j}^2 = \|\phi_{t_n}\|_{\Sigma^j}^2 + \int_{t_n}^{t_{n+1}} \Im\langle[A^j,|x|^2]\phi_s,\phi_s\rangle dW_s + \int_{t_n}^{t_{n+1}} \Re\langle[A^j,|x|^2]\phi_s,|x|^2\,\phi_s\rangle ds. \tag{3.34}$$

The first term in the right-hand side of Equation (3.33) is an approximation of the stochastic integral of Equation (3.34), while the second term in the right-hand side is an approximation of the Itô correction in Equation (3.34). It is possible to estimate these two terms in the same way as would be done to show at most exponential growth for (3.34) (using Gronwall's inequality). Yet, the main problem is the apparition of the third term in the right-hand side of Equation (3.33). We do not know how to estimate it. It is the reason why we make use of the approximation B_K of the operator $|x|^2$.

We now first prove Proposition 3.9 and then Lemma 3.10.

Proof of Proposition 3.9. First, we recall that the approximation of the nonlinearity is Lipschitz as stated by the following lemma,

Lemma 3.12. *Let $k > d/2$. Then for all $L \in \mathbb{N}^*$, there exists $C(L,k) > 0$ such that,*

$$\forall u_1,u_2,v_1,v_2 \in \Sigma^k(\mathbb{R}^d), \quad \left\|g_L^k(u_1,v_1) - g_L^k(u_2,v_2)\right\|_{\Sigma^k} \leq C(L,k)(\|u_1-u_2\|_{\Sigma^k} + \|v_1-v_2\|_{\Sigma^k}).$$

Moreover, for $d = 1,2,3$, for all $j \geq k$, and for all $L \in \mathbb{N}^*$, there exists $C(j,L) > 0$ such that,

$$\forall u,v \in \Sigma^j(\mathbb{R}^d), \quad \left\|g_L^k(u,v)\right\|_{\Sigma^j} \leq C(j,L)\left(\|u\|_{\Sigma^j} + \|v\|_{\Sigma^j}\right).$$

where g_L^k is defined by Equation (3.6).

The proof of Lemma 3.12 is classical, and is omitted in this article. Let $\Delta < T$. We denote by $X(j,\Delta)$ the space of $(\mathcal{F}_n)$-measurable processes that belong to $L^2_\omega L^\infty(0,\Delta;\Sigma^j(\mathbb{R}^d))$. To prove well-posedness of Equation (3.11) in $X(j,\Delta)$, we use a fixed-point method. It is clear that the solutions of (3.11) are the fixed point of the application $\mathcal{T}$ defined for all

$\phi \in X(j,\Delta)$ and for all $n \leq N$ by

$$\mathcal{T}(\phi)(t_n) = \left(\prod_{l=1}^{n} S_{\delta t,K,l,j}\right)\phi_0 - i\lambda\delta t \sum_{m=1}^{n}\left(\prod_{l=m+1}^{n} S_{\delta t,K,l,j}\right) P_K T^{-1}_{\delta t,K,m,j} g^k_L(\phi(t_{m-1}),\phi(t_m)). \tag{3.35}$$

We can show using Lemma 3.10, that for all $\phi,\psi \in X(j,\Delta)$,

$$\|\mathcal{T}\phi\|^2_{X(j,\Delta)} \leq 2e^{C(j,K_0)\Delta}\|\phi_0\|^2_{\Sigma^j} + 4\Delta^2 C(j,L)e^{C(j,K_0)\Delta}\|\phi\|^2_{X(j,\Delta)} \tag{3.36}$$

$$\|\mathcal{T}(\phi-\psi)\|^2_{X(k,\Delta)} \leq 4\Delta^2 C(k,L)e^{C(k,K_0)\Delta}\|\phi-\psi\|^2_{X(k,\Delta)}\,. \tag{3.37}$$

The second estimate is established in $X(k,\Delta)$ since we need g^k_L to be Lipschitz. Choosing Δ such that $4\Delta^2 C(j,L)e^{C(j,K_0)\Delta} \leq 1/2$ implies that $\mathcal{T}$ is a strict contraction in a ball $B_M^{X(j,\Delta)}$ of $X(j,\Delta)$ defined by $B_M^{X(j,\Delta)} := \{\phi \in X(j,\Delta), \|\phi\|^2_{X(j,\Delta)} \leq M\}$ where we chose $M := 4e^{C(j,K_0)\Delta}\|\phi_0\|^2_{X(j,\Delta)}$. This allows to show local existence and uniqueness in $X(j,\cdot)$. The key point is that the choice of Δ does not depend on ϕ_0. Thus, we can then iterate this process in $[\Delta,2\Delta]$ in the space $L^2_\omega L^\infty(\Delta,2\Delta;\Sigma^j(\mathbb{R}^d))$, and so on to get existence and uniqueness of a solution $X(j,T)$. Moreover since at each iteration the radius M is multiplied by $4e^{C(j,K_0)\Delta}$, we eventually get after repeating this construction on $\frac{T}{\Delta}$ intervals that,

$$\mathbb{E}\left[\sup_{t\leq T}\|\phi(t)\|^2_{\Sigma^j}\right] \leq 4^{T/\Delta}e^{C(j,K_0)T}\|\phi_0\|^2_{\Sigma^j}\,.$$

□

Proof of Lemma 3.10. To simplify the notation, we set $\phi^{n+1} = S_{\delta t,K,n+1,j}\phi^n$. We set $\widetilde{\chi}^{n+1}$ the random variable, independent of $\mathcal{F}_n$, defined by:

$$\widetilde{\chi}^{n+1} = \begin{cases} \widetilde{\chi}^{n+1} & \text{if} \quad \sqrt{\delta t}\left|\widetilde{\chi}^{n+1}\right| \leq C_0(j), \\ 0 & \text{otherwise.} \end{cases} \tag{3.38}$$

Thanks to the symmetry of B_K, one can write for almost every ω,

$$\left\|\phi^{n+1}\right\|^2_{\Sigma^j} = \|\phi^n\|^2_{\Sigma^j} + \frac{\sqrt{\delta t}}{4}\widetilde{\chi}^{n+1}\Im\langle[A^j,B_K](\phi^{n+1}+\phi^n),(\phi^{n+1}+\phi^n)\rangle. \tag{3.39}$$

We notice then that,

$$\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle[A^j,B_K](\phi^{n+1}+\phi^n),(\phi^{n+1}+\phi^n)\rangle = I^1_{n+1} + I^2_{n+1} + I^3_{n+1}, \tag{3.40}$$

with

$$\begin{aligned} I^1_{n+1} =& -\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle [A^j, B_K](\phi^{n+1}-\phi^n), (\phi^{n+1}-\phi^n)\rangle, \\ I^2_{n+1} =& 2\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle [A^j, B_K](\phi^{n+1}+\phi^n), (\phi^{n+1}-\phi^n)\rangle, \\ I^3_{n+1} =& 4\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle [A^j, B_K]\phi^n, \phi^n\rangle. \end{aligned}$$

Therefore, by setting $S_{p+1} = \sum_{n=0}^{p} I^3_{n+1}$, Equation (3.39) can be rewritten

$$\left\|\phi^{p+1}\right\|^2_{\Sigma^j} = \left\|\phi^0\right\|^2_{\Sigma^j} + S_{p+1} + \sum_{n=0}^{p}(I^1_{n+1} + I^2_{n+1}). \tag{3.41}$$

We begin by giving some estimates on I^1_n, I^2_n and I^3_n,

Lemma 3.13. *There exists $C(j) > 0$, depending only on j, such that*

$$\begin{aligned} \mathbb{E}\left[\left|I^1_{n+1}\right|^2 + \left|I^2_{n+1}\right|^2 \middle| \mathcal{F}_n\right]^{1/2} &\leq C(j)\delta t(1+K^2\delta t^{1/2})\left\|\phi^n\right\|^2_{\Sigma^j}, \\ \mathbb{E}\left[\left|I^3_{n+1}\right|^2 \middle| \mathcal{F}_n\right] &\leq C(j)\delta t\left\|\phi^n\right\|^4_{\Sigma^j}, \\ \mathbb{E}\left[I^3_{n+1} \middle| \mathcal{F}_n\right] &= 0. \end{aligned}$$

The proof of this Lemma is postponed after the end of this proof of Lemma 3.10.

We now give a bound on $\mathbb{E}\left[\left\|\phi^{n+1}\right\|^4_{\Sigma^j}\right]$ in terms of $\mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right]$. First, Equation (3.39) gives

$$\begin{aligned} \mathbb{E}\left[\left\|\phi^{n+1}\right\|^4_{\Sigma^j}\right] =& \mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right] + \frac{1}{16}\mathbb{E}\left[\left(I^1_{n+1}+I^2_{n+1}+I^3_{n+1}\right)^2\right] && (3.42) \\ &+\frac{1}{2}\mathbb{E}\left[\left\|\phi^n\right\|^2_{\Sigma^j}\left(I^1_{n+1}+I^2_{n+1}+I^3_{n+1}\right)\right]. && (3.43) \end{aligned}$$

Then, by using Lemma 3.13, we obtain the following estimates for $\delta t \leq 1$,

$$\mathbb{E}\left[\left(I^1_{n+1}+I^2_{n+1}+I^3_{n+1}\right)^2\right] \leq C(j)\delta t(1+K^2\delta t^{1/2})^2\mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right], \tag{3.44}$$

$$\begin{aligned} \mathbb{E}\left[\left\|\phi^n\right\|^2_{\Sigma^j}\left(I^1_{n+1}+I^2_{n+1}+I^3_{n+1}\right)\right] =& \mathbb{E}\left[\left\|\phi^n\right\|^2_{\Sigma^j}\left(I^1_{n+1}+I^2_{n+1}\right)\right] \\ \leq& \mathbb{E}\left[\left\|\phi^n\right\|^2_{\Sigma^j}\mathbb{E}\left[(I^1_{n+1}+I^2_{n+1})^2\middle|\mathcal{F}_n\right]^{1/2}\right] \\ \leq& C(j)\delta t(1+K^2\delta t^{1/2})\mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right]. \end{aligned} \tag{3.45}$$

Collecting (3.42), (3.44) and (3.45), we obtain for $\delta t \leq 1$,

$$\mathbb{E}\left[\left\|\phi^{n+1}\right\|^4_{\Sigma^j}\right] \leq C(j)\delta t(1+K^2\delta t^{1/2})^2\mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right].$$

Then by using Gronwall's inequality, we can estimate $\mathbb{E}\left[\left\|\phi^n\right\|^4_{\Sigma^j}\right]$ uniformly in $n \leq N$ and

$\delta t \leq 1$, as follows:

$$\mathbb{E}\left[\|\phi^n\|_{\Sigma^j}^4\right]^{1/2} \leq e^{C(j)T(1+K^2\delta t^{1/2})^2} \|\phi_0\|_{\Sigma^j}^2 . \tag{3.46}$$

We now turn to the estimate of $\mathbb{E}\left[\sup_{p\leq N} \|\phi^p\|_{\Sigma^j}^2\right]$. We proceed by triangular inequality by bounding first $\mathbb{E}\left[\sup_{p\leq N} S_p\right]$ and then $\mathbb{E}\left[\sup_{p\leq N} \sum_{n=0}^p (I_{n+1}^1 + I_{n+1}^2)\right]$ in Equation (3.41). Noticing that $(S_p)_{p\in\mathbb{N}}$ is a $(\mathcal{F}_p)$-martingale in L^2, and using Doob's inequality,

$$\mathbb{E}\left[\sup_{p\leq N} S_p\right] \leq 2\mathbb{E}\left[S_N^2\right]^{1/2} .$$

Since, for all $n \in \mathbb{N}$, $\mathbb{E}\left[I_{n+1}^3 \middle| \mathcal{F}_n\right] = 0$, using Lemma 3.13 and Equation (3.46),

$$\mathbb{E}\left[S_N^2\right] = \sum_{n=0}^{N-1} \mathbb{E}\left[(I_{n+1}^3)^2\right] \leq C(j)Te^{C(j)T(1+K^2\delta t^{1/2})^2} \|\phi_0\|_{\Sigma^j}^4 .$$

Then,

$$\mathbb{E}\left[\sup_{p\leq N} S_p\right] \leq C(j)Te^{C(j)T(1+K^2\delta t^{1/2})} \|\phi_0\|_{\Sigma^j}^2 . \tag{3.47}$$

The estimate of $\mathbb{E}\left[\sup_{p\leq N} \sum_{n=0}^p (I_{n+1}^1 + I_{n+1}^2)\right]$ is done using triangular inequality, Lemma 3.13 and Equation (3.46),

$$\begin{aligned}\mathbb{E}\left[\sup_{p\leq N} \sum_{n=0}^p I_{n+1}^1 + I_{n+1}^2\right]^2 &\leq \mathbb{E}\left[\left(\sup_{p\leq N} \sum_{n=0}^p I_{n+1}^1 + I_{n+1}^2\right)^2\right] \\ &\leq N\mathbb{E}\left[\sum_{n=0}^N (I_{n+1}^1 + I_{n+1}^2)^2\right],\end{aligned}$$

and using Lemma 3.13 and the uniform bound given by (3.46),

$$\mathbb{E}\left[\sup_{p\leq N} \sum_{n=0}^p I_{n+1}^1 + I_{n+1}^2\right]^2 \leq C(j)N^2\delta t^2(1+K^2\delta t^{1/2})^2 e^{C(j)T(1+K^2\delta t^{1/2})^2} \|\phi_0\|_{\Sigma^j}^2 ,$$

that is to say,

$$\mathbb{E}\left[\sup_{p\leq N} \sum_{n=0}^p I_{n+1}^1 + I_{n+1}^2\right] \leq C(j,T,K^4\delta t) \|\phi_0\|_{\Sigma^j}^2 . \tag{3.48}$$

Collecting Equations (3.41), (3.47) and (3.48), we obtain that for $\delta t \leq 1$, there exists

$C(j, T, K^4\delta t)$ such that

$$\mathbb{E}\left[\sup_{n\leq N} \|\phi^n\|^2_{\Sigma^j}\right] \leq C(j, T, K^4\delta t)\, \|\phi_0\|^2_{\Sigma^j}\,.$$

□

We now prove the technical Lemma 3.13 introduced in the proof of the Lemma 3.10.

Proof of Lemma 3.13. We prove estimates on I^1_{n+1} and I^2_{n+1}. Equation (3.15) enables us to estimate the inner product that appears in I^1_{n+1},

$$\begin{aligned}\left|I^1_{n+1}\right| =& \left|\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle[A^j, B_K](\phi^{n+1}-\phi^n), (\phi^{n+1}-\phi^n)\rangle\right| \\ \leq& C(j)\sqrt{\delta t}\left|\widetilde{\chi}^{n+1}\right| \left\|\phi^{n+1}-\phi^n\right\|^2_{\Sigma^j}\,,\end{aligned}$$

and using again the implicit relation (3.26), we obtain

$$\left|I^1_{n+1}\right| \leq C(j)\sqrt{\delta t}\left|\widetilde{\chi}^{n+1}\right| \left(\left\|\delta t A_K \phi^{n+1/2}\right\|^2_{\Sigma^j} + \left\|\sqrt{\delta t}\widetilde{\chi}^{n+1} B_K \phi^{n+1/2}\right\|^2_{\Sigma^j}\right).$$

Then, Equation (3.14), with the fact that $\left\|\phi^{n+1/2}\right\|^2_{\Sigma^j} \leq \sqrt{2}\,\|\phi^n\|^2_{\Sigma^j}$, given by Equation (3.30), yields for $\delta t \leq 1$

$$\mathbb{E}\left[(I^1_{n+1})^2 \middle| \mathcal{F}_n\right]^{1/2} \leq C(j)\delta t(1+K^2\delta t^{1/2})\,\|\phi^n\|^2_{\Sigma^j}\,.$$

We now prove estimate on I^2_{n+1}. We recall,

$$I^2_{n+1} = 2\sqrt{\delta t}\widetilde{\chi}^{n+1}\Im\langle[A^j, B_K](\phi^{n+1}+\phi^n), (\phi^{n+1}-\phi^n)\rangle.$$

Again, we replace $\phi^{n+1}-\phi^n$ by the implicit formulation given by (3.26),

$$\begin{aligned}I^2_{n+1} =& \delta t^{3/2}\widetilde{\chi}^{n+1}\Re\langle[A^j, B_K](\phi^{n+1}+\phi^n), A_K(\phi^{n+1}+\phi^n)\rangle \\ &+ \delta t(\widetilde{\chi}^{n+1})^2\Re\langle[A^j, B_K](\phi^{n+1}+\phi^n), B_K(\phi^{n+1}+\phi^n)\rangle \\ =& I^{2,a}_{n+1} + I^{2,b}_{n+1}.\end{aligned}$$

Then, by (3.15)

$$\left|I^{2,a}_{n+1}\right| \leq C(j)\left|\widetilde{\chi}^{n+1}\right| \delta t^{3/2} \left\|\phi^{n+1/2}\right\|_{\Sigma^j} \left\|P_K\phi^{n+1/2}\right\|_{\Sigma^{j+2}}.$$

Using (3.19) and Lemma 3.8, we get,

$$\left\|\phi^{n+1/2}\right\|_{\Sigma^j} \left\|P_K\phi^{n+1/2}\right\|_{\Sigma^{j+2}} \leq CK\,\|\phi^n\|^2_{\Sigma^j}\,,$$

which leads to,

$$\left|I_{n+1}^{2,a}\right| \leq C(j)\delta t^{3/2} K \left|\widetilde{\chi}^{n+1}\right| \|\phi^n\|_{\Sigma^j}^2 ,$$

and it follows that,

$$\mathbb{E}\left[\left|I_{n+1}^{2,a}\right|^2 \middle| \mathcal{F}_n\right]^{1/2} \leq C(j)\delta t^{3/2} K \|\phi^n\|_{\Sigma^j}^2 . \tag{3.49}$$

The estimate of $I_{n+1}^{2,b}$ follows from (3.16) and Lemma 3.8,

$$\mathbb{E}\left[\left|I_{n+1}^{2,b}\right|^2 \middle| \mathcal{F}_n\right]^{1/2} \leq C(j)\delta t \|\phi^n\|_{\Sigma^j}^2 . \tag{3.50}$$

Eventually, collecting (3.49) and (3.50) we get that

$$\mathbb{E}\left[(I_{n+1}^2)^2 \middle| \mathcal{F}_n\right]^{1/2} \leq C(j)\delta t(1 + K^2\delta t^{1/2}) \|\phi^n\|_{\Sigma^j}^2 ,$$

which concludes the proof. □

3.4 Convergence of the numerical scheme

The main objective of this section is to prove that the solution of equation (3.11) converges to the solution of Equation (3.1) in a probabilistic sense that we define later. Our numerical scheme $(\phi_{K,L,\delta t}^n)_{n\leq N}$ (3.11) is parametrized by the three variables K, L and δt. Morally, we expect $\phi_{K,L,\delta t}^n$ to tend to $\phi(t_n)$ as we make δt tend to zero, and K and L to infinity. The goal is now to prove the convergence result. To do so, we specify the divergence of K and L with respect to δt. We consider them as functions of the time step, and we denote them by $K(\delta t)$ and $L(\delta t)$. We are able to make the function $K(\delta t)$ explicit, but not the function $L(\delta t)$. Indeed, we recall that on one hand, in order to prove stability, we had to impose a maximal growth on K with respect to $\delta t^{-1/4}$. Now, we actually require it to grow at this speed $\delta t^{-1/4}$ (in other words we suppose that there exists $K_0, K_1 > 0$ such that $K_1 \leq K(\delta t)\delta t^{-1/4} \leq K_0$ for all δt). This condition on $K(\delta t)$ is sufficient to prove the convergence. Nevertheless Theorem 3.14 only establishes the existence of a choice of $L(\delta t)$ for all δt that enables to prove the convergence. Since we do not track explicitly the dependences on L in our computations, we are not able to define explicitly this function $L(\delta t)$. This function is chosen according to the following criteria. First, for δt small enough, $L(\delta t)$ must be sufficiently small to ensure the existence of a solution for the numerical scheme. This remark defines a maximal speed of divergence on $L(\delta t)$. Then, since we compute all the error bounds in a mean-square sense up to a multiplicative constant that depends on L, and increases with L (see Proposition 3.16 and 3.17), this parameter should diverge slowly enough with respect to δt to still ensure the convergences stated by these propositions. This last point is explained in detail in the

proof of Theorem 3.14.

We now state our result on the convergence of $(\phi^n_{K(\delta t),L(\delta t),\delta t})_{n\in\mathbb{N}}$ toward $(\phi(t_n))_{n\in\mathbb{N}}$, when δt tends to zero.

Theorem 3.14. *We suppose that Assumption 3.6 and 3.7 are satisfied. For all $T > 0$, $k \in \mathbb{N}^*$, $\phi_0 \in \Sigma^{k+12}(\mathbb{R}^d)$, $K_0 > K_1 > 0$, $C > 0$, and all $\alpha < 1$, there exists a choice of $L(\delta t)$ such that*

$$\lim_{\delta t\to 0} \sup_{n\delta t\leq T} \mathbb{P}\left(\left\|\phi^n_{K(\delta t),L(\delta t),\delta t} - \phi(t_n)\right\|_{\Sigma^k} \geq C\delta t^\alpha\right) = 0. \tag{3.51}$$

This theorem ensures that the numerical error is small for small δt on large sets of Ω. This kind of convergence result has been studied in [42] for a numerical scheme for the stochastic cubic Schrödinger equation with multiplicative noise of Stratonovich type.

Remark 3.15. *The high regularity of the initial condition ϕ_0 that we assume in Theorem 3.14 (namely that $\phi_0 \in \Sigma^{k+12}$) is classical for this kind of result. For instance, in [20], the authors assume that the initial condition belongs to the Sobolev space $H^{s+12}(\mathbb{R}^d)$ to be able to show a convergence in the Sobolev space $H^s(\mathbb{R}^d)$.*

In order to explain how this result is obtained, we introduce some notation. We consider the following truncated and projected equation for a fixed $L > 0$.

$$\begin{cases} id\phi_{K(\delta t),L} - \lambda P_{K(\delta t)} f^k_L(\phi_{K(\delta t),L})dt - A_{K(\delta t)}\phi_{K(\delta t),L}dt = B_{K(\delta t)}\phi_{K(\delta t),L} \circ dW_t, \\ \phi_{K(\delta t),L}(0) = P_{K(\delta t)}\phi_0. \end{cases} \tag{3.52}$$

We denote by $(\phi_{K(\delta t),L}(t))_{t\leq T}$, the solution of this equation. We also set $(\phi_L(t))_{t\leq T}$, the solution of the following truncated equation

$$\begin{cases} id\phi_L - \lambda f^k_L(\phi_L)dt - A\phi_L dt = |x|^2\phi_L \circ dW_t, \\ \phi_L(0) = \phi_0. \end{cases} \tag{3.53}$$

We finally introduce the $\mathcal{F}_n$-measurable process $(\psi^n_{K(\delta t),L,\delta t})_{n\leq N}$ which is essentially the stopped version of $(\phi_{K(\delta t),L}(t))_{t\leq T}$ at stopping time $\tau(\delta t)$:

$$\psi^n_{K(\delta t),L,\delta t} = \begin{cases} P_K\phi_0, & \text{if} \quad n = 0, \\ \phi_{K(\delta t),L}(t_n), & \text{if} \quad 0 < n < \tau(\delta t), \\ \psi^{n-1}_{K(\delta t),L,\delta t}, & \text{otherwise.} \end{cases} \tag{3.54}$$

The convergence is obtained in several steps. We start by proving that $(\phi^n_{K(\delta t),L,\delta t})$ can be as close as we want (for the $L^\infty L^2_\omega \Sigma^k$ norm) to $(\psi^n_{K(\delta t),L,\delta t})$ by choosing δt small enough. This result is stated in Proposition 3.16. Then we use the fact that $(\psi^n_{K(\delta t),L,\delta t})$ is equal to $(\phi_{K(\delta t),L}(t_n))$ with high probability when δt is small. We next use the fact that

$(\phi_{K(\delta t),L}(t_n))$ converges to $(\phi_L(t_n))$ for the $L^\infty L^2_\omega \Sigma^k$ norm at order one in δt, provided that $K(\delta t)$ grows at least at speed $\delta t^{-1/4}$. This result is stated in Proposition 3.17. Eventually, we use the fact that $(\phi_L(t_n))$ is equal to $(\phi(t_n))$ with high probability when L is large.

Proposition 3.16. *Suppose that Assumption 3.6 is satisfied. For all $T > 0$, for all $k \in \mathbb{N}^*$, for all $L \in \mathbb{N}^*$ and for all $K_0 > 0$, there exist $N_0(T,k,L) \in \mathbb{N}^*$ and $C(T,k,L,K_0) > 0$, such that, for all $N \geq N_0(T,k,L)$ and $\delta t = T/N$, for all $K \leq K_0 \delta t^{-1/4}$, and for all $\phi_0 \in \Sigma^{k+12}(\mathbb{R}^d)$,*

$$\sup_{n \leq N} \mathbb{E}\left[\left\| \phi^n_{K(\delta t),L,\delta t} - \psi^n_{K(\delta t),L,\delta t} \right\|^2_{\Sigma^k} \right] \leq C(T,k,L,K_0) \delta t^2 \left\| \phi_0 \right\|^2_{\Sigma^{k+12}}, \tag{3.55}$$

where $(\phi^n_{K(\delta t),L,\delta t})_{n \leq N}$ is defined by (3.11), and $(\psi^n_{K(\delta t),L,\delta t})_{n \leq N}$ is defined by (3.54).

Proposition 3.17. *We suppose that Assumption 3.6 and 3.7 are satisfied. For all $T > 0$, for all $k \in \mathbb{N}^*$, for all $L \in \mathbb{N}^*$ and for all $K_1 > 0$, there exists $C(T,k,L,K_1) > 0$, such that, for all $\delta t = T/N$, for all $K \geq K_1 \delta t^{-1/4}$, and for all $\phi_0 \in \Sigma^{k+12}(\mathbb{R}^d)$,*

$$\sup_{n \leq N} \mathbb{E}\left[\left\| \phi_{K(\delta t),L}(t_n) - \phi_L(t_n) \right\|^2_{\Sigma^k} \right] \leq C(T,k,L,K_1) \delta t^2 \left\| \phi_0 \right\|^2_{\Sigma^{k+12}}. \tag{3.56}$$

The proof of Proposition 3.16 is quite technical. It is done by controlling the numerical error at one step by the error at the previous step, in order to use Gronwall's inequality to conclude. This proof uses extensively the boundedness of operators B_K as in Lemma 3.10. The proof of Proposition 3.17 relies on Itô's lemma, Assumption 3.7 and the fact that K grows as speed $\delta t^{-1/4}$.

We now prove first Theorem 3.14, and then Propositions 3.16 and 3.17. The proof of Theorem 3.14 relies on these two propositions. We choose to show this proof first since the arguments are quite simple and follow from the convergence results of Propositions 3.16 and 3.17.

Proof of Theorem 3.14. Our goal is now to show that for all $\varepsilon > 0$ there exists a pair $(\delta t_1, L)$ such that $\delta t_1 \leq \delta t_0(j,L,K_0)$ and for all $\delta t \leq \delta t_1$,

$$\mathbb{P}\left(\left\| \phi^n_{K(\delta t),L,\delta t} - \phi(t_n) \right\|_{\Sigma^k} \geq C \delta t^\alpha \right) \leq \varepsilon.$$

Then, we can set $L(\delta t_1)$ to be such a L associated with δt_1. For $L > 0$ and $\delta t \leq \delta t_0(j,L,K_0)$, we upper bound $\mathbb{P}\left(\left\| \phi^n_{K(\delta t),L,\delta t} - \phi(t_n) \geq C \delta t^\alpha \right\|_{\Sigma^k} \right)$ in the following way:

$$\begin{aligned}\mathbb{P}\left(\left\|\phi^n_{K(\delta t),L,\delta t}-\phi(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha\right)\leq&\mathbb{P}\left(\left\|\phi^n_{K(\delta t),L,\delta t}-\psi^n_{K(\delta t),L,\delta t}\right\|_{\Sigma^k}\geq C\delta t^\alpha/4\right)\\&+\mathbb{P}\left(\left\|\psi^n_{K(\delta t),L,\delta t}-\phi_{K(\delta t),L}(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha/4\right)\\&+\mathbb{P}\left(\left\|\phi_{K(\delta t),L}(t_n)-\phi_L(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha/4\right)\\&+\mathbb{P}\left(\left\|\phi_L(t_n)-\phi(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha/4\right)\\\leq& i+ii+iii+iv.\end{aligned}\tag{3.57}$$

Choosing L large enough ensures that $\mathbb{P}\left(\sup_{t\in[0,T]}\left\|\phi(t_n)\right\|_{\Sigma^k}\geq L\right)\leq\varepsilon/4$ since $(\phi(t))_{t\geq0}\in\mathcal{C}(\mathbb{R}^+,\Sigma^k)$. In this case, $\mathbb{P}\left(\left\|\phi_L(t_n)-\phi(t_n)\right\|_{\Sigma^k}>0\right)\leq\varepsilon/4$ and *a fortiori*

$$\mathbb{P}\left(\left\|\phi_L(t_n)-\phi(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha/4\right)\leq\varepsilon/4.$$

Then it comes $iv\leq\varepsilon/4$. Obviously, this choice of truncation parameter L implies the existence of a maximal time-step, that we denote by $\delta t_0(j,L,K_0)$ that ensures existence of a solution for the numerical scheme. We do not have any estimate on the smallness of this maximal time-step, but we actually do not require one to prove this theorem. The other terms in the right-hand side of Equation (3.57) are made smaller than $\varepsilon/4$ with sufficiently small time-steps δt, and especially smaller than $\delta t_0(j,L,K_0)$.

At the cost of supposing δt smaller than $\delta t_0(j,L,K_0)$, we can assume that $\mathbb{P}\left(\tau(\delta t)\leq N\right)\leq\varepsilon/4$, which leads to $ii\leq\varepsilon/4$. This follows from Lemma 3.5.

We now use Proposition 3.16 and 3.17 with Markov's inequality to choose C such that $i+iii\leq\varepsilon/2$:

$$\begin{aligned}i+iii\leq&\frac{1}{C^2\delta t^{2\alpha}}\left(\mathbb{E}\left[\left\|\phi^n_{K(\delta t),L,\delta t}-\psi^n_{K(\delta t),L,\delta t}\right\|^2_{\Sigma^k}\right]+\mathbb{E}\left[\left\|\phi_{K(\delta t),L}(t_n)-\phi_L(t_n)\right\|^2_{\Sigma^k}\right]\right)\\\leq&\frac{\delta t^{2(1-\alpha)}}{C^2}\left\|\phi_0\right\|^2_{\Sigma^{k+12}}\left(C(T,k,L,K_0)+C(T,k,L,K_1)\right).\end{aligned}$$

Since $\delta t^{2(1-\alpha)}$ tends to zero when δt tends to zero, then it is possible to take δt even smaller to ensure that $i+iii\leq\varepsilon/2$.

It is clear that for all ε we can construct a pair $(\delta t_1,L(\delta t_1))$ with δt_1 that tends to zero when ε tends to zero, and such that

$$\mathbb{P}\left(\left\|\phi^n_{K(\delta t_1),L(\delta t_1),\delta t_1}-\phi(t_n)\right\|_{\Sigma^k}\geq C\delta t_1^\alpha\right)\leq\varepsilon.$$

This implies the result,

$$\lim_{\delta t\to0}\mathbb{P}\left(\left\|\phi^n_{K(\delta t),L(\delta t),\delta t}-\phi(t_n)\right\|_{\Sigma^k}\geq C\delta t^\alpha\right)=0.$$

□

Proof of Proposition 3.16. To simplify the notation, we get rid of the subscripts $K(\delta t)$, L and δt in this step. For example, we denote $\phi^n_{K(\delta t),L,\delta t}$ by ϕ^n. We also use g and f instead of g^k_L and f^k_L defined by (3.6) and (3.7). We set for all $n \leq N$, $e^n = \phi^n - \psi^n$.

Our goal is to show that there exists $C(T,k,L,K_0) > 0$, such that,

$$\mathbb{E}\left[\left\|e^{n+1}\right\|^2_{\Sigma^k}\right] \leq (1+\delta t C(T,k,L,K_0))\mathbb{E}\left[\left\|e^n\right\|^2_{\Sigma^k}\right] + \delta t^3 C(T,k,L,K_0)\left\|\phi_0\right\|^2_{\Sigma^{k+12}}. \quad (3.58)$$

This will enable us to conclude the proof of this proposition by using the discrete Gronwall's lemma. To show Equation (3.58), we prove that,

$$\left|\Re\mathbb{E}\left[\langle e^{n+1}-e^n, A^k e^{n+1/2}\rangle\right]\right| \leq C(T,k,L,K_0)\delta t\left(\mathbb{E}\left[\left\|e^n\right\|^2_{\Sigma^k}\right] + \delta t^2\left\|\phi_0\right\|^2_{\Sigma^{k+12}}\right). \quad (3.59)$$

To obtain these estimates, we begin by splitting $e^{n+1}-e^n$ as a sum of several terms. Using Equations (3.54), (3.52) and (3.11), we get that

$$\phi^{n+1} = \phi^n + \mathbb{1}_{\{\tau(\delta t)>n+1\}}\left(-i\delta t A_K\phi^{n+1/2} - i\sqrt{\delta t}\chi^{n+1}B_K\phi^{n+1/2} - i\delta t P_K g(\phi^{n+1},\phi^n)\right),$$

and

$$\begin{aligned}\psi(t_{n+1}) =& \psi(t_n) + \mathbb{1}_{\{\tau(\delta t)>n+1\}}\left(-i\int_{t_n}^{t_{n+1}} A_K\phi(s)ds - i\int_{t_n}^{t_{n+1}} B_K\phi(s)\circ dW_s\right.\\ &\left. -i\int_{t_n}^{t_{n+1}} P_K f(\phi(s))ds\right).\end{aligned}$$

We split $e^{n+1}-e^n$ in the following way,

$$e^{n+1} = e^n + \mathbb{1}_{\{\tau(\delta t)>n+1\}}(a_1+a_2+a_3), \quad (3.60)$$

with

$$a_1 = -i\delta t A_K\phi^{n+1/2} + i\int_{t_n}^{t_{n+1}} A_K\phi(s)ds, \quad (3.61)$$

$$a_2 = -i\sqrt{\delta t}\chi^{n+1}B_K\phi^{n+1/2} + i\int_{t_n}^{t_{n+1}} B_K\phi(s)\circ dW_s, \quad (3.62)$$

$$a_3 = -i\delta t P_K g(\phi^{n+1},\phi^n) + i\int_{t_n}^{t_{n+1}} P_K f(\phi(s))ds. \quad (3.63)$$

The term a_1 comes from the term $A_K\phi_L dt$ in (3.53), a_2 comes from the term $B_K\phi_L\circ dW_t$, and a_3 comes from the nonlinear term. We now split these terms again. We set $a_1 =$

$a_{1,1} + a_{1,2} + a_{1,3}$ with,

$$\begin{aligned} a_{1,1} &= -i\delta t A_K e^{n+1/2}, \\ a_{1,2} &= -\frac{i}{2}\delta t A_K(\phi(t_{n+1}) - \phi(t_n)), \\ a_{1,3} &= i\int_{t_n}^{t_{n+1}} A_K(\phi(s) - \phi(t_n))ds. \end{aligned}$$

We define $R(t,s)$ for $t \geq s$ by

$$R(t,s) = \int_s^t A_K\phi(r)dr + \frac{1}{2}\int_s^t B_K^2\phi(r)dr + \int_s^t P_K f(\phi(r))dr. \tag{3.64}$$

a_2 can be split in the following way.

$$\begin{aligned} a_2 =& -i\sqrt{\delta t}\chi^{n+1}B_K e^{n+1/2} - i\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K(\phi(t_{n+1}) - \phi(t_n)) \\ &+ i\int_{t_n}^{t_{n+1}} B_K(\phi(s) - \phi(t_n)) \circ dW_s, \end{aligned}$$

and changing the Stratonovich integral into an Itô integral gives

$$\begin{aligned} a_2 =& -i\sqrt{\delta t}\chi^{n+1}B_K e^{n+1/2} + \frac{1}{2}\int_{t_n}^{t_{n+1}} B_K^2(\phi(s) - \phi(t_n))ds \\ &- i\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K(\phi(t_{n+1}) - \phi(t_n)) + i\int_{t_n}^{t_{n+1}} B_K(\phi(s) - \phi(t_n))dW_s. \end{aligned}$$

We split again the two last terms in the right-hand side of the second equality by using again Equation (3.52) to explicit $\phi(t_{n+1}) - \phi(t_n)$ and $\phi(s) - \phi(t_n)$. Then, we can split a_2 by setting $a_2 = a_{2,1} + a_{2,2} + a_{2,3} + a_{2,4}$ with,

$$\begin{aligned} a_{2,1} &= -i\sqrt{\delta t}\chi^{n+1}B_K e^{n+1/2}, \\ a_{2,2} &= -\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K\int_{t_n}^{t_{n+1}} B_K(\phi(s) - \phi(t_n))dW_s \\ &\quad + \int_{t_n}^{t_{n+1}} B_K\int_{t_n}^{s} B_K(\phi(r) - \phi(t_n))dW_r dW_s, \\ a_{2,3} &= \frac{1}{2}\int_{t_n}^{t_{n+1}} B_K^2(\phi(s) - \phi(t_n))ds, \\ a_{2,4} &= -\frac{\sqrt{\delta t}}{2}\chi^{n+1}B_K R(t_{n+1}, t_n) + \int_{t_n}^{t_{n+1}} B_K R(s, t_n)dW_s. \end{aligned}$$

We set $a_3 = a_{3,1} + a_{3,2} + a_{3,3}$ with,

$$\begin{aligned} a_{3,1} &= -i\delta t P_K \left(g(\phi^{n+1}, \phi^n) - g(\phi(t_{n+1}), \phi(t_n)) \right), \\ a_{3,2} &= -i\delta t P_K \left(g(\phi(t_{n+1}), \phi(t_n)) - f(\phi(t_n)) \right), \\ a_{3,3} &= i \int_{t_n}^{t_{n+1}} P_K \left(f(\phi(s)) - f(\phi(t_n)) \right) ds. \end{aligned}$$

The terms $a_{1,2}$, $a_{1,3}$, $a_{2,2}$, $a_{2,3}$, $a_{2,4}$, $a_{3,2}$ and $a_{3,3}$ have similar *smallness* properties. This is why we defined

$$b = a_{1,2} + a_{1,3} + a_{2,2} + a_{2,3} + a_{2,4} + a_{3,2} + a_{3,3}.$$

Then, Equation (3.60) can be written as

$$e^{n+1} - e^n = \mathbb{1}_{\{\tau(\delta t) > n+1\}} \left(a_{1,1} + a_{2,1} + a_{3,1} + b \right). \tag{3.65}$$

Plugging this expression of $e^{n+1} - e^n$ into Equation (3.59) and using the triangular inequality it becomes clear that, in order to show the estimate (3.59), it is enough to show that for all $1 \leq p \leq 3$,

$$\left| \Re\mathbb{E} \left[\mathbb{1}_{\{\tau(\delta t) > n+1\}} \langle a_{p,1}, A^k e^{n+1/2} \rangle \right] \right| \leq C(k, j, T) \delta t \left(\mathbb{E} \left[\|e^n\|_{\Sigma^k}^2 \right] + \delta t^2 \|\phi_0\|_{\Sigma^{k+12}}^2 \right), \tag{3.66}$$

and

$$\left| \Re\mathbb{E} \left[\mathbb{1}_{\{\tau(\delta t) > n+1\}} \langle b, A^k e^{n+1/2} \rangle \right] \right| \leq C(k, j, T) \delta t \left(\mathbb{E} \left[\|e^n\|_{\Sigma^k}^2 \right] + \delta t^2 \|\phi_0\|_{\Sigma^{k+12}}^2 \right). \tag{3.67}$$

If we try instead to take the square of each side of Equation (3.60), then we will not be able to estimate properly $\mathbb{E}\left[\|a_2\|_{\Sigma^k}^2 \right]$.

Before starting to prove Equations (3.66) and (3.67), we explain why we chose to decompose $e^{n+1} - e^n$ as in Equation (3.65). As we said, it comes from the *smallness* properties of b, which is quantified by the following Lemma,

Lemma 3.18. *There exists $C(L, j, T) > 0$ such that,*

$$\mathbb{E} \left[\|b\|_{\Sigma^j}^4 \right]^{1/2} \leq C(L, j, T) \delta t^3 \|\phi_0\|_{\Sigma^{j+8}}^2.$$

Proof of Lemma 3.18. The proof of this Lemma is a direct consequence of Lemmas 3.19 and 3.20. We show for instance the estimate on $\mathbb{E}\left[\|a_{2,3}\|_{\Sigma^j}^4 \right]^{1/2}$. First, Hölder's inequality gives,

$$\begin{aligned} \|a_{2,3}\|_{\Sigma^j}^4 \leq & C\delta t^3 \int_{t_n}^{t_{n+1}} \left\| B_K^2(\phi(s) - \phi(t)) \right\|_{\Sigma^j}^4 ds \\ & C(j)\delta t^3 \int_{t_n}^{t_{n+1}} \|\phi(s) - \phi(t)\|_{\Sigma^{j+4}}^4 ds. \end{aligned}$$

Taking the expectation and using Lemma 3.20, it comes,

$$\mathbb{E}\left[\|a_{2,3}\|_{\Sigma^j}^4\right] \leq C(T,L,j)\delta t^3 \left(\int_{t_n}^{t_{n+1}} (s-t_n)^2 ds\right) \|\phi_0\|_{\Sigma^{j+8}}^4$$
$$C(T,L,j)\delta t^6 \|\phi_0\|_{\Sigma^{j+8}}^4 ,$$

which gives the kind of estimate we are looking for. □

Lemma 3.19. *For all $T > 0$, $L \in \mathbb{N}^*$, $K(\delta t) \in \mathbb{N}^*$, $j \in \mathbb{N}$ if $\phi_0 \in \Sigma^j(\mathbb{R}^d)$, then $\phi_{K(\delta t),L} \in L^\infty(0,T;L^4(\Omega;\Sigma^j(\mathbb{R}^d)))$, where $\phi_{K(\delta t),L}$ is the solution of (3.52), and there exists $C(T,L,j) > 0$, independent of ϕ_0, such that,*

$$\sup_{t\leq T} \mathbb{E}\left[\left\|\phi_{K(\delta t),L}(t)\right\|_{\Sigma^j}^4\right]^{1/2} \leq C(T,L,j)\|\phi_0\|_{\Sigma^j}^2 .$$

Lemma 3.20. *For all $T > 0$, $L \in \mathbb{N}^*$, $j \in \mathbb{N}$, $K(\delta t) \in \mathbb{N}^*$ and for all $\phi_0 \in \Sigma^j(\mathbb{R}^d)$, there exists $C(T,L,j) > 0$, independent of ϕ_0, such that, for all $s \leq t$*

$$\mathbb{E}\left[\left\|\phi_{K(\delta t),L}(t) - \phi_{K(\delta t),L}(s)\right\|_{\Sigma^j}^4\right]^{1/2} \leq C(T,L,j)(t-s)\|\phi_0\|_{\Sigma^{j+4}}^2 ,$$

and

$$\mathbb{E}\left[\|R(t,s)\|_{\Sigma^j}^4\right]^{1/2} \leq C(T,L,j)(t-s)^2\|\phi_0\|_{\Sigma^{j+4}}^2 ,$$

where $\phi_{K(\delta t),L}$ is the solution of (3.53) and $R(t,s)$ is defined by (3.64).

Their proof is left to the reader and follows from Itô's Lemma.

Yet, in the remainder of the proof, we need another stronger result.

Lemma 3.21. *The term b can be split in two terms b_1 and b_2, such that there exists $C(L,j,T) > 0$ for which*

$$\mathbb{E}\left[\|b_1\|_{\Sigma^j}^4\right]^{1/2} \leq C(L,j,T)\delta t^4\|\phi_0\|_{\Sigma^{j+8}}^2 , \tag{3.68}$$

$$\mathbb{E}\left[\|b_2\|_{\Sigma^j}^4\right]^{1/2} \leq C(L,j,T)\delta t^3\|\phi_0\|_{\Sigma^{j+8}}^2 ,$$

and

$$\mathbb{E}\left[\langle b_2, A^k e_n\rangle\right] = 0.$$

Proof of Lemma 3.21. The proof of Lemma 3.21 is quite computational, but not difficult. The idea is to develop the terms $\phi(s)-\phi(t_n)$ that appear in the expressions of the $a_{i,j}$ using Equation (3.53) in the Itô form. Then one has to collect the terms with zero expectation when taking their L^2 inner product with $A^k e_n$. These terms form b_2. The remaining terms

form b_1 and their estimate (3.68) follows from a direct application of Lemmas 3.19 and 3.20. □

We can state the following Lemma, that gives a naive estimate of $e^{n+1} - e^n$,

Lemma 3.22. *For all $j \in \mathbb{N}^*$, $L > 0$ and $K_0 > 0$, there exists $C(L, j, T, K_0) > 0$ such that, for all $\delta t \leq \delta t_0(j+2, L, K_0)$ (where $\delta t_0(j+2, L, K_0)$ is given by Proposition 3.9)*

$$\mathbb{E}\left[\left\|e^{n+1} - e^n\right\|_{\Sigma^j}^2\right] \leq C(L, j, T, K_0)\left(\delta t \mathbb{E}\left[\|e^n\|_{\Sigma^{j+2}}^2\right] + \delta t^3 \mathbb{E}\left[\|b\|_{\Sigma^{j+10}}^2\right]\right).$$

Proof of Lemma 3.22. The proof starts with the following triangular inequality,

$$\mathbb{E}\left[\left\|e^{n+1} - e^n\right\|_{\Sigma^j}^2\right] \leq 3\left(\mathbb{E}\left[\|a_{1,1}\|_{\Sigma^j}^2\right] + \mathbb{E}\left[\|a_{2,1}\|_{\Sigma^j}^2\right] + \mathbb{E}\left[\|a_{3,1}\|_{\Sigma^j}^2\right] + \mathbb{E}\left[\|b\|_{\Sigma^j}^2\right]\right). \tag{3.69}$$

Then, it comes directly that,

$$\mathbb{E}\left[\|a_{1,1}\|_{\Sigma^j}^2\right] \leq \delta t^2 \mathbb{E}\left[\left\|e^{n+1/2}\right\|_{\Sigma^{j+2}}^2\right],$$

and using Lemma 3.23,

$$\mathbb{E}\left[\|a_{1,1}\|_{\Sigma^j}^2\right] \leq C(j, L)\delta t^2\left(\mathbb{E}\left[\|e^n\|_{\Sigma^{j+2}}^2\right] + \mathbb{E}\left[\|b\|_{\Sigma^{j+2}}^2\right]\right). \tag{3.70}$$

Similarly,

$$\mathbb{E}\left[\|a_{2,1}\|_{\Sigma^j}^2\right] \leq \delta t \mathbb{E}\left[(\chi^{n+1})^2\left\|e^{n+1/2}\right\|_{\Sigma^{j+2}}^2\right],$$

and using Lemma 3.23,

$$\mathbb{E}\left[\|a_{2,1}\|_{\Sigma^j}^2\right] \leq C(j, L)\delta t\left(\mathbb{E}\left[\|e^n\|_{\Sigma^{j+2}}^2\right] + \mathbb{E}\left[\|b\|_{\Sigma^{j+2}}^2\right]\right). \tag{3.71}$$

The third term is estimated from the fact that g is Lipschitz,

$$\mathbb{E}\left[\|a_{3,1}\|_{\Sigma^j}^2\right] \leq C(j, L)\delta t^2\left(\mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^j}^2\right] + \mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^j}^2\right]\right),$$

and using Lemma 3.23 yields,

$$\mathbb{E}\left[\|a_{3,1}\|_{\Sigma^j}^2\right] \leq C(j, L)\delta t^2\left(\mathbb{E}\left[\|e^n\|_{\Sigma^j}^2\right] + \mathbb{E}\left[\|b\|_{\Sigma^j}^2\right]\right). \tag{3.72}$$

Eventually, collecting Equations (3.69), (3.70), (3.71) and (3.72), and using Lemma 3.18 yields,

$$\mathbb{E}\left[\left\|e^{n+1} - e^n\right\|_{\Sigma^j}^2\right] \leq C(L, j, T, K_0)\left(\delta t \mathbb{E}\left[\|e^n\|_{\Sigma^{j+2}}^2\right] + \delta t^3 \mathbb{E}\left[\|b\|_{\Sigma^{j+10}}^2\right]\right).$$

□

We need a last result that makes use of the truncation of the discretized Brownian motion at each time step.

Lemma 3.23. *For all $j \in \mathbb{N}^*$, $L > 0$ and $K_0 > 0$, there exists $\delta t_1(j, L, K_0) \leq \delta t_0(j, L, K_0)$ (where $\delta t_0(j, L, K_0)$ is given by Proposition 3.9) and $C(j, L) > 0$ such that for all $\delta t \leq \delta t_1(j, L)$, almost surely,*

$$\left\|e^{n+1}\right\|_{\Sigma^k}^2 \leq C(j, L)\left(\|e^n\|_{\Sigma^k}^2 + \|b\|_{\Sigma^k}^2\right).$$

Proof of Lemma 3.23. If $\omega \in \{\tau(\delta t)(\omega) <= n+1\}$, then there is nothing to prove. Otherwise, by definition $\sqrt{\delta t}\left|\chi^{n+1}\right| \leq C_0$ and we have by (3.60),

$$e^{n+1} = e^n + (a_1 + a_2 + a_3),$$

Then,

$$\left\|e^{n+1}\right\|_{\Sigma^k}^2 - \|e^n\|_{\Sigma^k}^2 = \Re\langle A^k(e^{n+1} + e^n), a_{1,1} + a_{2,1} + a_{3,1} + b\rangle.$$

Since $a_{1,1} = -i\delta t A_K e^{n+1/2}$, it comes that

$$\Re\langle A^k(e^{n+1} + e^n), a_{1,1}\rangle = 0.$$

Since $a_{2,1} = -i\sqrt{\delta t}\chi^{n+1}B_K e^{n+1/2}$,

$$\begin{aligned}\Re\langle A^k(e^{n+1} + e^n), a_{2,1}\rangle &= -2\sqrt{\delta t}\chi^{n+1}\Im\langle A^k e^{n+1/2}, B_K e^{n+1/2}\rangle \\ &= \sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, e^{n+1/2}\rangle.\end{aligned}$$

Then, using (3.15), we get in the same way of Equation (3.27) and (3.29),

$$\begin{aligned}\Re\langle A^k(e^{n+1} + e^n), a_{2,1}\rangle &\leq C(j)\sqrt{\delta t}\left|\chi\right|\left\|e^{n+1/2}\right\|_{\Sigma^k}^2 \\ &\leq C_0 C(j)\left\|e^{n+1/2}\right\|_{\Sigma^k}^2 \\ &\leq C_0 C(j)\frac{1}{2}\left(\left\|e^{n+1}\right\|_{\Sigma^k}^2 + \|e^n\|_{\Sigma^k}^2\right).\end{aligned}$$

Note that the constant $C(j)$ is the same here as in Equation (3.27). Then, following the proof of Lemma 3.8, we get that $C_0 C(j) = 1/3$.

Since $a_{3,1} = -i\delta t P_K\left(g(\phi^{n+1}, \phi^n) - g(\phi(t_{n+1}), \phi(t_n))\right)$, and using Lemma 3.12,

$$\Re\langle A^k(e^{n+1} + e^n), a_{3,1}\rangle \leq C(j, L)\delta t\left(\left\|e^{n+1}\right\|_{\Sigma^k}^2 + \|e^n\|_{\Sigma^k}^2\right).$$

Next, by the Cauchy-Schwarz and the Young inequalities, for all $\varepsilon > 0$,

$$\Re\langle A^k(e^{n+1}+e^n), b\rangle \leq \varepsilon \left\|e^{n+1}\right\|_{\Sigma^k}^2 + \frac{4}{\varepsilon}\|b\|_{\Sigma^k}^2 .$$

To sum up, we obtain that,

$$\left\|e^{n+1}\right\|_{\Sigma^k}^2 \left(1 - \frac{C_0C(j)}{2} + C(j,L)\delta t + \varepsilon\right) \leq \|e^n\|_{\Sigma^k}^2 \left(1 + \frac{C_0C(j)}{2} + C(j,L)\delta t\right) + \frac{4}{3}\|b\|_{\Sigma^k}^2 .$$

Then, by taking ε and δt small enough, one can ensure that

$$\left(1 - \frac{C_0C(j)}{2} + C(j,L)\delta t + \varepsilon\right) > 0,$$

which proves that there exists $C(j,L) > 0$, such that,

$$\left\|e^{n+1}\right\|_{\Sigma^k}^2 \leq C(j,L)\left(\|e^n\|_{\Sigma^k}^2 + \|b\|_{\Sigma^k}^2\right). \tag{3.73}$$

□

Thanks to these lemmas, we now are ready to prove Equation (3.66) for $p \in \{1,2,3\}$, and Equation (3.67). We begin by Equation (3.66).

— Estimate of $\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{1,1}, A^k e^{n+1/2}\rangle\right]$. We recall that

$$a_{1,1} = -i\delta t A_K e^{n+1/2}.$$

Then $\langle a_{1,1}, A^k e^{n+1/2}\rangle$ is purely imaginary, and it follows that

$$\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{1,1}, A^k e^{n+1/2}\rangle\right] = 0. \tag{3.74}$$

— Estimate of $\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{2,1}, A^k e^{n+1/2}\rangle\right]$. We recall that

$$a_{2,1} = -i\sqrt{\delta t}\chi^{n+1} B_K e^{n+1/2}.$$

Then,

$$\begin{aligned}\Re\langle A^k e^{n+1/2}, a_{2,1}\rangle &= -\sqrt{\delta t}\chi^{n+1}\Im\langle A^k e^{n+1/2}, B_K e^{n+1/2}\rangle \\ &= \frac{1}{2}\sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, e^{n+1/2}\rangle.\end{aligned}$$

We split this expression in the following way.

$$\begin{aligned}\Re\langle a_{2,1}, A^k e^{n+1/2}\rangle \leq & \frac{1}{4}\sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, e^{n+1}-e^n\rangle \\ & + \frac{1}{4}\sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n}, e^{n+1}-e^n\rangle \\ & + \frac{1}{2}\sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n}, e^{n}\rangle \\ = & I + II + III.\end{aligned}$$

Using the following triangular inequality,

$$\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{2,1}, A^k e^{n+1/2}\rangle\right]\right| \leq \mathbb{E}\left[|I|\right] + \mathbb{E}\left[|II|\right] + \left|\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}} III\right]\right|, \tag{3.75}$$

we now estimate each terms in the right-hand side of Equation (3.75).

— Estimate of $\mathbb{E}\left[|I|\right]$. We recall that

$$|I| = \left|\frac{1}{4}\sqrt{\delta t}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, e^{n+1}-e^n\rangle\right|.$$

Plugging the expression of $e^{n+1}-e^n$ given by (3.60) into this expression, and using the triangular inequality gives,

$$\begin{aligned}|I| \leq & \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, a_{1,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, a_{2,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, a_{3,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, b\rangle\right| \\ = & i + ii + iii + iv.\end{aligned}$$

By replacing $a_{1,1}$, $a_{2,1}$ and $a_{3,1}$ by their definitions gives,

$$\begin{aligned}|I| \leq & \left|\frac{1}{4}\delta t^{3/2}\chi^{n+1}\Re\langle [A^k, B_K]e^{n+1/2}, A_K e^{n+1/2}\rangle\right| \\ & + \left|\frac{1}{4}\delta t(\chi^{n+1})^2\Re\langle [A^k, B_K]e^{n+1/2}, B_K e^{n+1/2}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{3/2}\chi^{n+1}\Re\langle [A^k, B_K]e^{n+1/2}, f(\phi^n)-f(\phi(t_n))\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle [A^k, B_K]e^{n+1/2}, b\rangle\right| \\ = & i + ii + iii + iv.\end{aligned}$$

By the use of Assumption 3.6, we can easily get,

$$\mathbb{E}[i] + \mathbb{E}[ii] \leq C(k,L)\delta t \mathbb{E}\left[((\chi^{n+1})^2 + \sqrt{\delta t} K |\chi^{n+1}|)\right] \left\|e^{n+1/2}\right\|_{\Sigma^k}^2 .$$

Since $e^{n+1/2}$ is not $\mathcal{F}_n$-measurable, we now use the almost sure bound on $\left\|e^{n+1/2}\right\|_{\Sigma^k}^2$ given by Lemma 3.23, followed by Cauchy-Schwarz inequality,

$$\begin{aligned}\mathbb{E}[i] + \mathbb{E}[ii] \leq & C(k)\delta t \mathbb{E}\left[((\chi^{n+1})^2 + \sqrt{\delta t} K |\chi^{n+1}|)\left(\|e^n\|_{\Sigma^k}^2 + \|b\|_{\Sigma^k}^2\right)\right] \\ \leq & C(k)\delta t \mathbb{E}\left[(\chi^{n+1})^2 + \sqrt{\delta t} K |\chi^{n+1}|\right] \mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] \\ & + C(k)\delta t \mathbb{E}\left[((\chi^{n+1})^2 + \sqrt{\delta t} K |\chi^{n+1}|)^2\right]^{1/2} \mathbb{E}\left[\|b\|_{\Sigma^k}^4\right]^{1/2} .\end{aligned}$$

To estimate expression iii, we use the fact that f is Lipschitz, as well as the Cauchy-Schwarz and Young inequalities, to get,

$$\begin{aligned}\mathbb{E}[iii] \leq & C(k,L)\delta t^{3/2} \mathbb{E}\left[|\chi^{n+1}| \left\|e^{n+1/2}\right\|_{\Sigma^k} \|e^n\|_{\Sigma^k}\right] \\ \leq & C(k,L)\delta t^{3/2} \mathbb{E}\left[|\chi^{n+1}|^2\right] \mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] \\ & + C(k,L)\delta t^{3/2} \mathbb{E}\left[\left\|e^{n+1/2}\right\|_{\Sigma^k}^2\right] .\end{aligned}$$

Eventually, the last term iv is estimated by Young inequality followed by Cauchy-Schwarz inequality,

$$\mathbb{E}[iv] \leq C(k)\delta t \mathbb{E}\left[\left\|e^{n+1/2}\right\|_{\Sigma^k}^2\right] + C(k)\mathbb{E}\left[(\chi^{n+1})^2\right]^{1/2} \mathbb{E}\left[\|b\|_{\Sigma^k}^4\right]^{1/2} .$$

By collecting the estimates on i, ii, iii and iv, and using Lemmas 3.18 and 3.23, it comes the following estimate of $|I|$,

$$\mathbb{E}[|I|] \leq C(k, K^2\delta t^{1/2}, L, T)\left(\delta t \mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \delta t^3 \|\phi_0\|_{\Sigma^{k+8}}^2\right) . \tag{3.76}$$

— Estimate of $\mathbb{E}[|II|]$. We recall that

$$|II| = \left|\frac{1}{4}\sqrt{\delta t}\chi^{n+1} \Im\langle [A^k, B_K]e^n, e^{n+1} - e^n\rangle\right| .$$

Similarly to the estimate of $\mathbb{E}[|I|]$, we plug the expression of $e^{n+1} - e^n$ given by

(3.60) into this expression, and using the triangular inequality gives,

$$\begin{aligned}|II| \leq & \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle[A^k, B_K]e^n, a_{1,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle[A^k, B_K]e^n, a_{2,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle[A^k, B_K]e^n, a_{3,1}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle[A^k, B_K]e^n, b\rangle\right| \\ =& i + ii + iii + iv.\end{aligned}$$

By replacing $a_{1,1}$, $a_{2,1}$ and $a_{3,1}$ by their definitions gives,

$$\begin{aligned}|II| \leq & \left|\frac{1}{4}\delta t^{3/2}\chi^{n+1}\Re\langle[A^k, B_K]e^n, A_K e^{n+1/2}\rangle\right| \\ & + \left|\frac{1}{4}\delta t(\chi^{n+1})^2\Re\langle[A^k, B_K]e^n, B_K e^{n+1/2}\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{3/2}\chi^{n+1}\Re\langle[A^k, B_K]e^n, f(\phi^n) - f(\phi(t_n))\rangle\right| \\ & + \left|\frac{1}{4}\delta t^{1/2}\chi^{n+1}\Im\langle[A^k, B_K]e^n, b\rangle\right| \\ =& i + ii + iii + iv.\end{aligned}$$

Terms i, iii and iv can be estimated with the same techniques we used to estimate expression I, and we get,

$$\mathbb{E}\left[i + iii + iv\right] \leq C(k, K^2\delta t^{1/2}, L, T)\left(\delta t\mathbb{E}\left[\|e^n\|^2_{\Sigma^k}\right] + \delta t^3\|\phi_0\|^2_{\Sigma^{k+8}}\right). \quad (3.77)$$

To estimate ii we use the following technique. We split ii in the following way,

$$\begin{aligned}ii \leq & \left|\frac{1}{8}\delta t(\chi^{n+1})^2\Re\langle[A^k, B_K]e^n, B_K(e^{n+1} - e^n)\rangle\right| \\ & + \left|\frac{1}{4}\delta t(\chi^{n+1})^2\Re\langle[A^k, B_K]e^n, B_K e^n\rangle\right|.\end{aligned}$$

The expectation of the second term in the right-hand side is easily estimated using Assumption 3.6 since e^n is independent of χ^{n+1}, and gives,

$$\mathbb{E}\left[\left|\frac{1}{4}\delta t(\chi^{n+1})^2\Re\langle[A^k, B_K]e^n, B_K e^n\rangle\right|\right] \leq C(k)\delta t\mathbb{E}\left[\|e^n\|^2_{\Sigma^k}\right].$$

To estimate the first term in the right-hand side, we substitute again $e^{n+1} - e^n$

by its expression given by (3.60), and use the triangular inequality,

$$\begin{aligned}
&\left|\frac{1}{8}\delta t(\chi^{n+1})^2\Im\langle[A^k,B_K]e^n,B_K(e^{n+1}-e^n)\rangle\right| \\
&\quad\leq\left|\frac{1}{8}\delta t^2(\chi^{n+1})^2\Im\langle[A^k,B_K]e^n,B_KA_Ke^{n+1/2}\rangle\right| \\
&\quad+\left|\frac{1}{8}\delta t^{3/2}(\chi^{n+1})^3\Im\langle[A^k,B_K]e^n,B_K^2e^{n+1/2}\rangle\right| \\
&\quad+\left|\frac{1}{8}\delta t^2(\chi^{n+1})^2\Im\langle[A^k,B_K]e^n,B_K(f(\phi^n)-f(\phi(t_n)))\rangle\right| \\
&\quad+\left|\frac{1}{8}\delta t^{3/2}(\chi^{n+1})^3\Re\langle[A^k,B_K]e^n,B_Kb\rangle\right| \\
&\quad=\alpha+\beta+\gamma+\delta.
\end{aligned}$$

We only show the estimate of β that contains all the arguments required to estimate α, γ and δ for which the estimates are left to the reader. To estimate β, we use first Cauchy-Schwarz inequality and Assumption 3.14,

$$\begin{aligned}
\mathbb{E}\left[\beta\right]\leq&C(k)\delta t^{3/2}\mathbb{E}\left[\left|\chi^{n+1}\right|^3\left\|e^n\right\|_{\Sigma^k}\left\|B_K^2e^{n+1/2}\right\|_{\Sigma^k}\right]\\
\leq&C(k)\delta t(\delta t^{1/2}K^2)\mathbb{E}\left[\left|\chi^{n+1}\right|^3\left\|e^n\right\|_{\Sigma^k}\left\|e^{n+1/2}\right\|_{\Sigma^k}\right].
\end{aligned}$$

Now using Young's inequality and Lemma 3.23, followed by Lemma 3.18,

$$\begin{aligned}
\mathbb{E}\left[\beta\right]\leq&C(k)\delta t(\delta t^{1/2}K^2)\left(\mathbb{E}\left[\left|\chi^{n+1}\right|^6\left\|e^n\right\|_{\Sigma^k}^2\right]+\mathbb{E}\left[\left\|b\right\|_{\Sigma^k}^2\right]\right)\\
\leq&C(k)\delta t(\delta t^{1/2}K^2)\left(\mathbb{E}\left[\left|\chi^{n+1}\right|^6\right]\mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right]+C(L,k,T)\delta t^3\left\|\phi_0\right\|_{\Sigma^{k+8}}^2\right).
\end{aligned}$$

This last estimate can be written as,

$$\mathbb{E}\left[\beta\right]\leq C(k,\delta t^{1/2}K^2,L,T)\left(\delta t\mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right]+\delta t^3\left\|\phi_0\right\|_{\Sigma^{k+8}}^2\right), \tag{3.78}$$

and combining (3.77) and (3.78) leads to the following estimate for $\mathbb{E}\left[|II|\right]$,

$$\mathbb{E}\left[|II|\right]\leq C(k,\delta t^{1/2}K^2,L,T)\left(\delta t\mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right]+\delta t^3\left\|\phi_0\right\|_{\Sigma^{k+8}}^2\right). \tag{3.79}$$

— Estimate of $\left|\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}III\right]\right|$. It follows from the definition of $\tau(\delta t)$ (see (3.8)) that,

$$\mathbb{1}_{\{\tau(\delta t)>n+1\}}=\mathbb{1}_{\{\tau(\delta t)>n\}}\mathbb{1}_{\left\{|\chi^{n+1}|<C_0\delta t^{-1/2}\right\}}.$$

Using now the fact that $\chi^{n+1}\mathbb{1}_{\left\{|\chi^{n+1}|<C_0\delta t^{-1/2}\right\}}$ is independent of e^n and $\mathbb{1}_{\{\tau(\delta t)>n\}}$,

it follows

$$\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}} III\right] = \mathbb{E}\left[\chi^{n+1}\mathbb{1}_{\left\{|\chi^{n+1}|<C_0\delta t^{-1/2}\right\}}\right] \times \mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n\}}\frac{1}{2}\sqrt{\delta t}\Im\langle[A^k, B_K]e^n, e^n\rangle\right].$$

Since

$$\mathbb{E}\left[\chi^{n+1}\mathbb{1}_{\left\{|\chi^{n+1}|<C_0\delta t^{-1/2}\right\}}\right] = 0,$$

it follows that,

$$\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}} III\right] = 0. \tag{3.80}$$

Eventually, combining (3.76), (3.79) and (3.80) gives,

$$\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{2,1}, A^k e^{n+1/2}\rangle\right]\right| \leq C(k, \delta t^{1/2}K^2, L, T)\left(\delta t\mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \delta t^3\|\phi_0\|_{\Sigma^{k+8}}^2\right). \tag{3.81}$$

— Estimate of $\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{3,1}, A^k e^{n+1/2}\rangle\right]$. We recall that

$$a_{3,1} = -i\delta t P_K\left(g(\phi^{n+1}, \phi^n) - g(\phi(t_{n+1}), \phi(t_n))\right).$$

Using Cauchy-Schwarz inequality and the fact that g is Lipschitz for each variable in $\Sigma^k(\mathbb{R}^d)$, it easily comes

$$\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{3,1}, A^k e^{n+1/2}\rangle\right]\right| \leq C(k, L)\delta t\left(\mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right]\right).$$

Using now Lemma 3.23, it comes,

$$\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{3,1}, A^k e^{n+1/2}\rangle\right]\right| \leq C(k, L)\delta t\left(\mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \mathbb{E}\left[\|b\|_{\Sigma^k}^2\right]\right).$$

Eventually using Lemma 3.18, it follows that,

$$\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle a_{3,1}, A^k e^{n+1/2}\rangle\right]\right| \leq C(k, L, T)\left(\delta t\mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \delta t^3\|\phi_0\|_{\Sigma^{k+8}}^2\right). \tag{3.82}$$

We recall that, at this stage of the proof, Equations (3.74), (3.81) and (3.82) prove Equation (3.66) for $p \in \{1, 2, 3\}$. It remains to prove Equation (3.67), which is done now:

— Estimate of $\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle b, A^k e^{n+1/2}\rangle\right]$ This estimate uses extensively the Lemma 3.21.

We introduce b_1 and b_2, given this Lemma, and the triangular inequality to get,

$$\begin{aligned}\left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle b, A^k e^{n+1/2}\rangle\right]\right| \leq& \left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle b_1, A^k e^{n+1/2}\rangle\right]\right| \\ &+ \left|\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle b_2, A^k e^{n+1/2}\rangle\right]\right| \\ =&(i)+(ii).\end{aligned}$$

We now estimate (i) and (ii).

— Estimate of (i). By Cauchy-Schwarz and Young inequalities,

$$(i) \leq C\left(\frac{\delta t}{2}\mathbb{E}\left[\left\|e^{n+1/2}\right\|_{\Sigma^k}^2\right] + \frac{1}{2\delta t}\mathbb{E}\left[\|b_1\|_{\Sigma^k}^2\right]\right),$$

and by Lemma 3.21,

$$(i) \leq C(L,j,T)\left(\frac{\delta t}{2}\mathbb{E}\left[\left\|e^{n+1/2}\right\|_{\Sigma^k}^2\right] + \delta t^3\|\phi_0\|_{\Sigma^{k+8}}^2\right). \tag{3.83}$$

— Estimate of (ii). By triangular inequality, we split (ii) in the following way,

$$\begin{aligned}(ii) \leq& \frac{1}{2}\mathbb{E}\left[\left|\langle A^2 b_2, A^{k-2}(e^{n+1}-e^n)\rangle\right|\right] \\ &+ \left|\mathbb{E}\left[\langle b_2, A^k e^{n+1/2}\rangle\right]\right| \\ &+ \left|\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)\leq n+1\}}\langle b_2, A^k e^{n+1/2}\rangle\right]\right| \\ \leq&(a)+(b)+(c).\end{aligned}$$

Thanks to Lemma 3.21 we get

$$(b) = 0, \tag{3.84}$$

and using in addition Young's inequality,

$$(a) \leq C(L,j,T)\left(\mathbb{E}\left[\|e^n\|_{\Sigma^k}^2\right] + \delta t^3\|\phi_0\|_{\Sigma^{k+12}}^2\right). \tag{3.85}$$

We use Young's and Cauchy-Schwarz inequalities to estimate (c) ,

$$\begin{aligned}(c) &\leq \delta t\mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right] + \frac{1}{\delta t}\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)\leq n+1\}}\|b_2\|_{\Sigma^k}^2\right] \\ &\leq \delta t\mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right] + \frac{1}{\delta t}\mathbb{P}\left(\tau(\delta t)\leq n+1\right)^{1/2}\mathbb{E}\left[\|b_2\|_{\Sigma^k}^4\right]^{1/2}.\end{aligned}$$

Using again Lemma 3.21 and the fact that, by Lemma 3.5 there exists $C(T)>0$ such that

$$\mathbb{P}\left(\tau(\delta t)\leq n+1\right)^{1/2} \leq C(T)\delta t$$

holds uniformly in δt,

$$(c) \leq \delta t \mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right] + C(L,k,T)\delta t^3 \left\|\phi_0\right\|_{\Sigma^{k+8}} . \tag{3.86}$$

Indeed, a direct consequence of Lemma 3.5 gives that the probability $\mathbb{P}(\tau(\delta t) \leq n+1)$ can be bounded above by any power of δt asymptotically. Then, by collecting (3.83), (3.84), (3.85), and (3.86), we get that

$$\Re\mathbb{E}\left[\mathbb{1}_{\{\tau(\delta t)>n+1\}}\langle b, A^k e^{n+1/2}\rangle\right] \leq C(k,L,T)\left(\delta t \mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right] + \delta t^3 \left\|\phi_0\right\|_{\Sigma^{k+8}}^2\right), \tag{3.87}$$

which proves Equation (3.67).

Since Equations (3.66) and (3.67) have been proved, it immediately follows Equation (3.58):

$$\mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right] \leq (1+\delta t C(T,k,L,K_0))\mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right] + \delta t^3 C(T,k,L,K_0)\left\|\phi_0\right\|_{\Sigma^{k+12}}^2 ,$$

and the discrete Gronwall's Lemma gives that,

$$\mathbb{E}\left[\left\|e^{n+1}\right\|_{\Sigma^k}^2\right] \leq e^{C(T,k,L,K_0)n\delta t}\delta t^2 C(T,k,L,K_0)\left\|\phi_0\right\|_{\Sigma^{k+12}}^2 ,$$

and

$$\sup_{n\leq N}\mathbb{E}\left[\left\|e^n\right\|_{\Sigma^k}^2\right] \leq C(T,k,L,K_0)\delta t^2 \left\|\phi_0\right\|_{\Sigma^{k+12}}^2 .$$

□

Proof of Proposition 3.17. We begin by recalling that since the nonlinearity is truncated to be Lipschitz in $\Sigma^k(\mathbb{R}^d)$, it follows that for all $L>0$ and $T>0$, there exists $C(k,L,T)>0$ such that for all $\phi_0 \in \Sigma^{k+12}(\mathbb{R}^d)$,

$$\sup_{t\leq T}\mathbb{E}\left[\left\|\phi_L(t)\right\|_{\Sigma^{k+12}}^2\right] + \sup_{t\leq T}\mathbb{E}\left[\left\|\phi_{K(\delta t),L}(t)\right\|_{\Sigma^{k+12}}^2\right] \leq C(k,L,T)\left\|\phi_0\right\|_{\Sigma^{k+12}}^2 , \tag{3.88}$$

where ϕ_L is the solution of Equation (3.53) and $\phi_{K(\delta t),L}$ is the solution of Equation (3.52).

We define the $\mathcal{F}_t$-measurable process $\Phi_{K(\delta t),L}$ for all $t\geq 0$ by $\Phi_{K(\delta t),L}(t) := \phi_L(t) - \phi_{K(\delta t),L}(t)$. With this notation, we aim to show that

$$\sup_{t\leq T}\mathbb{E}\left[\left\|\Phi_{K(\delta t),L}(t)\right\|_{\Sigma^k}^2\right] \leq C(T,k,L,K_1)\delta t^2. \tag{3.89}$$

The proof is done by using Itô's lemma, followed by Gronwall's inequality. More precisely, we apply Itô's lemma to ϕ_L and $\phi_{K(\delta t),L}$ for the functional $(\phi_1,\phi_2) \mapsto \Re\langle A^k(\phi_1-\phi_2),(\phi_1-$

$\phi_2)\rangle = \|\phi_1 - \phi_2\|^2_{\Sigma^k}$. To make clear how the computation is done, we begin by recalling Itô's formulations of (3.53) and (3.52) where we denote K instead of $K(\delta t)$ to simplify the notation,

$$d\phi_{K,L} = -iA_K\phi_{K,L}dt - iB_K\phi_{K,L}dW_t - \frac{1}{2}B_K^2\phi_{K,L}dt - i\lambda P_K f_L^k(\phi_{K,L})dt,$$

and

$$\begin{aligned} d\phi_L =& -iA\phi_L dt - i\left|x\right|^2\phi_L dW_t - \frac{1}{2}\left|x\right|^4\phi_L dt - i\lambda f_L^k(\phi_L)dt \\ =& -iA_K\phi_L dt - iA(Id-P_K)\phi_L dt - i\left|x\right|^2\phi_L dW_t \\ & - \frac{1}{2}B_K^2\phi_L dt - \frac{1}{2}(|x|^4 - B_K^2)\phi_L dt - i\lambda P_K f_L^k(\phi_L)dt - i\lambda(Id-P_K)f_L^k(\phi_L)dt. \end{aligned}$$

Then, the multidimensional Itô's lemma gives,

$$\begin{aligned} \|\Phi_{K,L}(t)\|^2_{\Sigma^k} =& \|\Phi_{K,L}(0)\|^2_{\Sigma^k} \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -iA_K\Phi_{K,L}(s)\rangle ds \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -iA(Id-P_K)\phi_L(s)\rangle ds \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -i(\left|x\right|^2\phi_L(s) - B_K\phi_{K,L}(s))\rangle dW_s \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -\frac{1}{2}B_K^2\Phi_{K,L}(s)\rangle ds \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -\frac{1}{2}(|x|^4 - B_K^2)\phi_L(s)\rangle ds \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -i\lambda P_K(f_L^k(\phi_L(s)) - f_L^k(\phi_{K,L}(s)))\rangle ds \\ & + 2\int_0^t \Re\langle A^k\Phi_{K,L}(s), -i\lambda(Id-P_K)f_L^k(\phi_L(s))\rangle ds \\ & + \int_0^t \Re\langle A^k(-i\left|x\right|^2\phi_L(s)), -i\left|x\right|^2\phi_L(s)\rangle ds \\ & + \int_0^t \Re\langle A^k(-iB_K\phi_{K,L}(s)), -iB_K\phi_{K,L}(s)\rangle ds \\ & - 2\int_0^t \Re\langle A^k(-i\left|x\right|^2\phi_L(s)), -iB_K\phi_{K,L}(s)\rangle ds. \end{aligned} \tag{3.90}$$

The last three terms in the right-hand side are the Itô correction. The last line takes into account the cross derivatives of the functional. We write this correction in another way,

that will enable us to use Cauchy-Schwarz inequality, by noticing that,

$$\begin{aligned}\Re\langle A^k |x|^2 \phi_L(s), |x|^2 \phi_L(s)\rangle + \Re\langle A^k B_K \phi_{K,L}(s), B_K \phi_{K,L}(s)\rangle - 2\Re\langle A^k |x|^2 \phi_L(s), B_K \phi_{K,L}(s)\rangle \\ = \Re\langle A^k B_K \Phi_{K,L}(s), B_K \phi_{K,L}(s)\rangle + \Re\langle A^k(|x|^2 - B_K)\phi_L(s), (|x|^2 - B_K)\phi_L(s)\rangle \\ + 2\Re\langle A^k B_K \Phi_{K,L}(s), (|x|^2 - B_K)\phi_L(s)\rangle.\end{aligned}$$

Then, Equation (3.90) can be written as,

$$\begin{aligned}\|\Phi_{K,L}(t)\|^2_{\Sigma^k} = \|\Phi_{K,L}(0)\|^2_{\Sigma^k} \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -iA_K \Phi_{K,L}(s)\rangle ds \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -iA(Id - P_K)\phi_L(s)\rangle ds \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -i(|x|^2 \phi_L(s) - B_K \phi_{K,L}(s))\rangle dW_s \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -\frac{1}{2} B_K^2 \Phi_{K,L}(s)\rangle ds \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -\frac{1}{2}(|x|^4 - B_K^2)\phi_L(s)\rangle ds \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -i\lambda P_K(f_L^k(\phi_L(s)) - f_L^k(\phi_{K,L}(s)))\rangle ds \\ + 2\int_0^t \Re\langle A^k \Phi_{K,L}(s), -i\lambda(Id - P_K) f_L^k(\phi_L(s))\rangle ds \\ + \int_0^t \Re\langle A^k B_K \Phi_{K,L}(s), B_K \Phi_{K,L}(s)\rangle ds \\ + \int_0^t \Re\langle A^k(|x|^2 - B_K)\phi_L(s), (|x|^2 - B_K)\phi_L(s)\rangle ds \\ + 2\int_0^t \Re\langle A^k B_K \Phi_{K,L}(s), (|x|^2 - B_K)\phi_L(s)\rangle ds.\end{aligned} \tag{3.91}$$

We now proceed to estimate all the terms that appear in the right-hand side of Equation (3.91),

— the second term is equal to zero since $\Re\langle A^k \Phi_{K,L}(s), -iA_K \Phi_{K,L}(s)\rangle = 0$.
— Using Cauchy-Schwarz and Young inequalities, and Equation (3.19), it follows that the third term can be bounded in the following way,

$$\Re\langle A^k \Phi_{K,L}(s), -iA(Id - P_K)\phi_L(s)\rangle \leq C\left(\|\Phi_{K,L}(s)\|^2_{\Sigma^k} + K^{-8}\|\phi_L(s)\|^2_{\Sigma^{k+10}}\right).$$

— The expectation of the stochastic integral vanishes.

— Collecting the fifth and the ninth terms,

$$2\Re\langle A^k\Phi_{K,L}(s), -\frac{1}{2}B_K^2\Phi_{K,L}(s)\rangle + \Re\langle A^k B_K\Phi_{K,L}(s), B_K\Phi_{K,L}(s)\rangle \\ = \Re\langle [A^k, B_K]\Phi_{K,L}(s), B_K\Phi_{K,L}(s)\rangle,$$

and by (3.16),

$$2\Re\langle A^k\Phi_{K,L}(s), -\frac{1}{2}B_K^2\Phi_{K,L}(s)\rangle + \Re\langle A^k B_K\Phi_{K,L}(s), B_K\Phi_{K,L}(s)\rangle \\ \leq C\left\|\Phi_{K,L}(s)\right\|_{\Sigma^k}^2.$$

— Using the fact that $|x|^4 - B_K^2 = |x|^2\left(|x|^2 - B_K\right) + \left(|x|^2 - B_K\right)B_K$, the sixth term can be bounded by,

$$\Re\langle A^k\Phi_{K,L}(s), -\frac{1}{2}(|x|^4 - B_K^2)\phi_L(s)\rangle \leq C\left(\left\|\Phi_{K,L}(s)\right\|_{\Sigma^k}^2 + K^{-8}\left\|\phi_L(s)\right\|_{\Sigma^{k+12}}^2\right),$$

similarly to the third term.

— The estimate of the seventh term uses the truncation of nonlinearity,

$$\Re\langle A^k\Phi_{K,L}(s), -i\lambda P_K(f_L^k(\phi_L(s)) - f_L^k(\phi_{K,L}(s)))\rangle \leq C(L,k)\left\|\Phi_{K,L}(s)\right\|_{\Sigma^k}^2.$$

The estimate of the eighth term uses in addition the Cauchy-Schwarz and Young inequalities, and (3.19),

$$\Re\langle A^k\Phi_{K,L}(s), -i\lambda(Id - P_K)f_L^k(\phi_L(s))\rangle \leq C(L,k)\left(\left\|\Phi_{K,L}(s)\right\|_{\Sigma^k}^2 + K^{-8}\left\|\phi_L(s)\right\|_{\Sigma^{k+8}}^2\right).$$

— The tenth term is estimated by Assumption 3.7,

$$\Re\langle A^k(|x|^2 - B_K)\phi_L(s), (|x|^2 - B_K)\phi_L(s)\rangle \leq CK^{-8}\left\|\phi_L(s)\right\|_{\Sigma^{k+10}}^2.$$

— The last term is estimated by replacing $A^k B_K$ by $B_K A^k + [A^k, B_K]$, and using the triangular inequality and Assumption 3.7,

$$\Re\langle A^k B_K\Phi_{K,L}(s), (|x|^2 - B_K)\phi_L(s) \leq C\left(\left\|\Phi_{K,L}(s)\right\|_{\Sigma^k}^2 + K^{-8}\left\|\phi_L(s)\right\|_{\Sigma^{k+12}}^2\right).$$

Collecting all these estimates, it follows that, for all $t \leq T$,

$$\mathbb{E}\left[\left\|\Phi_{K(\delta t),L}(t)\right\|_{\Sigma^k}^2\right] \leq \left\|\Phi_{K(\delta t),L}(0)\right\|_{\Sigma^k}^2 + C(k,L)\int_0^t \mathbb{E}\left[\left\|\Phi_{K(\delta t),L}(s)\right\|_{\Sigma^k}^2\right]ds \\ + C(k,L)K^{-8}\int_0^t \mathbb{E}\left[\left\|\phi_L(s)\right\|_{\Sigma^{k+12}}^2\right]ds.$$

Using now (3.88) and the fact that $\left\|\Phi_{K(\delta t),L}(0)\right\|_{\Sigma^k}^2 \leq CK^{-8}\left\|\phi_0\right\|_{\Sigma^{k+8}}^2$, it comes that for

all $t \leq T$,

$$\mathbb{E}\left[\left\|\Phi_{K(\delta t),L}(t)\right\|_{\Sigma^k}^2\right] \leq C(k,L,T)K^{-8}\left\|\phi_0\right\|_{\Sigma^{k+12}}^2 + C(k,L)\int_0^t \mathbb{E}\left[\left\|\Phi_{K(\delta t),L}(s)\right\|_{\Sigma^k}^2\right] ds.$$

The application of Gronwall's inequality with the fact that we choose $K^{-8} \leq K_1^{-4}\delta t^2$ enables to conclude the proof. □

3.5 Numerical experiments

This part is devoted to numerical experiments in space dimension one with the scheme defined above. We aim to show that we obtain good results using a spectral discretisation based on Hermite functions. We also compare the Crank-Nicolson time discretisation against splitting methods. Four different schemes are compared in the following. We begin by explaining how to implement the scheme defined by (3.11), and then we present the other schemes, which are some adaptations of classical schemes for the deterministic Gross-Pitaevskii equation.

Some of the schemes we will present in the following rely on the computation of a discrete Hermite transform at each time-step. This transform decomposes a function $f \in L^2(\mathbb{R})$ on the basis of $L^2(\mathbb{R})$ composed by Hermite functions, which are the eigenvectors of operator $-\Delta + |x|^2$. The discrete transform on the K first modes enables to exactly compute the components of any function $f \in \Sigma_K(\mathbb{R})$ in this basis. It can be computed using the Gauss–Hermite quadrature. Obviously, for a general element $f \in L^2(\mathbb{R})$, the discrete Hermite transform does not exactly compute these projections. The error between the values of the components of f on the K first Hermite functions, and the values given by the discrete transform can be bounded by inverse polynomials of K depending on the regularity of those functions (see [121]). A naive implementation (by a classical matrix-vector product) of the discrete Hermite transform algorithm (in 1D) using the Gauss-Hermite quadrature on the K first Hermite functions leads to a computational cost of order K^2. Such an implementation can be generalized to the d-dimensional case with computational costs of order K^{d+1}. Nevertheless, there exists efficient implementations with computational costs of order $K^d \log^2(K)$ (see [69, 112, 143]). In the following we will say that a computation is *almost exact* when the only approximations come from the fact that we use the discrete Hermite transform to compute the projections on the K first Hermite functions for a general function $f \in L^2(\mathbb{R}^d)$, which does not necessarily belong to $\Sigma_K(\mathbb{R}^d)$. The same denomination will be used for the Discrete Fourier Transform.

With the explicit example of operator B_K we gave (with Equations (3.21), (3.22), (3.23) and (3.24)), the solutions of Equation (3.11) can be computed *almost exactly* thanks to a fixed-point iterative algorithm. The convergence of such an algorithm is ensured if δt is small enough with respect to L, the level of truncation of the nonlinearity, as it is shown in the proof of Proposition 3.9. Nevertheless, the smooth spectral cut-off of operator $|x|^2$

does not seem to have a significant impact in practical cases. Thus, in practice one can simply choose the symmetric operator $P_K |x|^2 P_K$ for B_K.

Moreover, in practical cases, the truncations of the non-linearity and the normal deviates $(\chi^n)_{n\in\mathbb{N}^*}$ do not seem to be significant and can be omitted in the implementation. In this case, we denote by (ϕ_K^n) the numerical approximation such that for all $n \in \mathbb{N}$, $\phi_K^n \in M_{K,1}(\mathbb{R})$ (column vector of size K) and its components are the components in the Hermite basis of $L^2(\mathbb{R})$. Then, Equation (3.11) is given by

$$D(\delta t, \chi^{n+1})\phi_K^{n+1} = C(\delta t, \chi^{n+1})\phi_K^n - i\delta t P_K g(\phi_K^{n+1}, \phi_K^n), \tag{3.92}$$

where C and D are the operators defined on $\Sigma_K(\mathbb{R})$ by

$$C(\delta t, \chi^{n+1}) = \left(Id - i\frac{\delta t}{2}A - i\frac{\sqrt{\delta t}\chi^{n+1}}{2} P_K |x|^2 \right), \tag{3.93}$$

and,

$$D(\delta t, \chi^{n+1}) = \left(Id + i\frac{\delta t}{2}A + i\frac{\sqrt{\delta t}\chi^{n+1}}{2} P_K |x|^2 \right). \tag{3.94}$$

We recall that, by denoting $(e_k)_{k\in\mathbb{N}}$ the orthonormal basis of $L^2(\mathbb{R})$ composed by Hermite functions, we have the recursive relation,

$$\forall x \in \mathbb{R}, \qquad x e_k(x) = \sqrt{\frac{k}{2}} e_{k-1}(x) + \sqrt{\frac{k+1}{2}} e_{k+1}(x).$$

Thus, in 1D the matrices of the operators C and D are pentadiagonal in the Hermite basis. The computation of the Hermite modes of $P_K g(\phi_K^{n+1}, \phi_K^n)$ can be done almost exactly using the discrete Hermite transform. Indeed, knowing ϕ_K^n and ϕ_K^{n+1} one can compute $g(\phi_K^{n+1}, \phi_K^n)(x)$ exactly for all $x \in \mathbb{R}$. Thus one can use a discrete Hermite transform to compute *almost exactly* the components of $P_K g(\phi_K^{n+1}, \phi_K^n)$ in the Hermite basis.

We now present three other schemes which are compared to the one presented above.

Splitting scheme with spectral-Fourier discretisation. For this scheme, the time discretisation is based on a splitting method, and the space discretisation is based on a spectral discretisation on the Fourier modes. In other words, we look for a solution defined on $[-L_x, L_x]$, with periodic boundary conditions, supported by the K first Fourier modes. A similar scheme has been proposed for the deterministic case in [8, 13] for solving (3.1) (without the Stratonovich noise). We generalize this scheme to our stochastic setting. The

idea is to solve the following equation,

$$\begin{cases} d\phi = -i\Delta\phi dt, \\ \phi(t_n) = \phi_K^n, \end{cases} \tag{3.95}$$

and set $\phi_K^{n+1/2} = \phi(t_{n+1})$. Then, one solves

$$\begin{cases} d\psi = -i\,|x|^2\,\psi dt - i\,|x|^2\,\psi \circ dW_t - ig\,|\psi|^2\,\psi dt, \\ \psi(t_n) = \phi_K^{n+1/2}, \end{cases} \tag{3.96}$$

and sets $\phi_K^{n+1} = P_K\psi(t_{n+1})$. Here P_K denotes the projection onto the K first Fourier modes. The first step is computed exactly in the Fourier space. To compute the other one, we use a discrete Fourier transform. An inverse discrete Fourier transform enables to compute the value of $\phi_K^{n+1/2}$ on a uniform grid of $[-L_x, L_x]$ of $(K+1)$ points. Then, the value of $\psi(t_{n+1})$ on those points x is explicitly given by

$$\psi(t_{n+1})(x) = e^{-i|x|^2(\delta t + \sqrt{\delta t}\chi^{n+1}) - ig\left|\phi_K^{n+1/2}\right|^2\delta t}\phi_K^{n+1/2}(x).$$

This computation relies on the fact that $t \mapsto |\psi(t)|^2$ is constant on $[t_n, t_{n+1}]$. Then, the computation of the first Fourier modes of ϕ_K^{n+1} is done *almost exactly* by a discrete Fourier transform.

Splitting scheme with spectral-Hermite discretisation. This scheme is similar to the previous one, but we chose this time a spectral discretisation based on the first Hermite functions. Equations (3.95) and (3.96) are replaced by

$$\begin{cases} d\phi = -iA\phi dt, \\ \phi(t_n) = \phi_K^n, \end{cases} \tag{3.97}$$

and

$$\begin{cases} d\psi = -i\,|x|^2\,\psi \circ dW_t - ig\,|\psi|^2\,\psi \circ dt, \\ \psi(t_n) = \phi_K^{n+1/2}, \end{cases} \tag{3.98}$$

and P_K denotes the projection onto the first Hermite functions. Similarly, both equations (3.97) and (3.98) are solved exactly, and the projection of $\psi(t_{n+1})$ onto the K first Hermite functions is done by a discrete Hermite Transform, and thus is not exact.

Crank-Nicolson scheme with finite differences discretisation. In this case, we suppose that the solution almost vanishes outside the space interval $[-L_x, L_x]$, so that we can set homogeneous Dirichlet conditions on the boundaries of this domain. The space discretisation of the Laplacian is classical, and the nonlinearity can be truncated to ensure well-posedness of the scheme. We discretize $[-L_x, L_x]$ with a regular grid of $2K+1$ points. We set $\delta x = L_x/K$. We denote by ϕ_k^n an approximation of $\phi(n\delta t, k\delta x)$ for $k \in \{-K, -K+$

$1, \dots, K\}$, and set $\phi_k^{n+1/2} = \frac{\phi_k^{n+1}+\phi_k^n}{2}$. We define the scheme by induction for $n \in \mathbb{N}$ and $k \in \{-K+1, \dots, K-1\}$ by setting $\phi_k^0 = \phi(0, k\delta x)$ and

$$\begin{aligned}\phi_k^{n+1} = \phi_k^n - i\frac{\delta t}{\delta x^2}(\phi_{k+1}^{n+1/2} - 2\phi_k^{n+1/2} + \phi_{k-1}^{n+1/2}) - i(\delta t + \sqrt{\delta t}\chi^{n+1})\,|k\delta x|^2\,\phi_k^{n+1/2} \\ - i\delta t g(\phi_k^{n+1}, \phi_k^n).\end{aligned} \tag{3.99}$$

In the following, we compare the convergence properties for these four schemes. We begin by estimating the pathwise speed of convergence for the space discretisation. To do that, we choose a fine enough time-step for the effect of the time discretisation to be negligible, and a geometrically increasing sequence of degrees of freedom for the space discretisation. Then we compare the error between the solutions computed for two successive levels of space discretisation, for one trajectory. More precisely, we choose $T = 10$, $\delta t = 10 \times 2^{-20}$. For the Fourier and finite differences approximations, we solve the equation on the domain $[-10, 10]$ and we set respectively periodic boundary conditions and Dirichlet homogeneous. Moreover, we use a coefficient $\alpha = 0.3$ in front of the noise, to decrease its effects. It enables to reduce significantly the error coming from the time discretisation. The result is given in Figure 3.1 We can observe first that the discretisation by

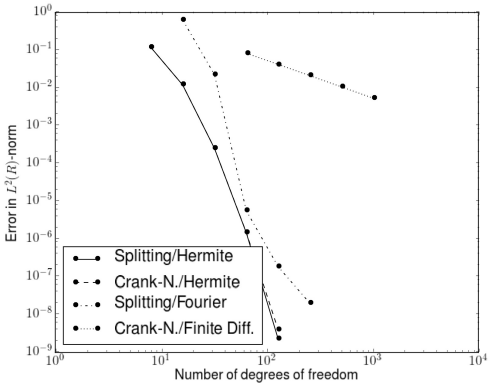

Figure 3.1 – Pathwise speed of convergence for each discretisation toward their respective limit.

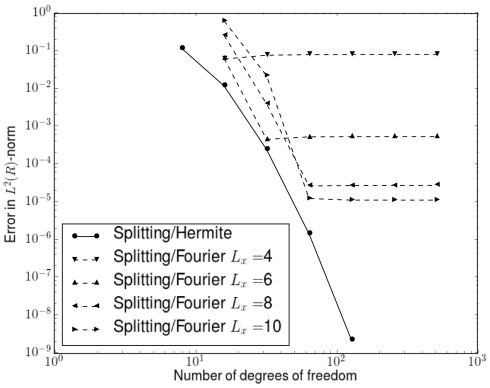

Figure 3.2 – Pathwise speed of convergence toward the exact solution

finite differences is less efficient than the spectral discretisation for the same number of degrees of freedom. We can also observe that for the two schemes using the spectral-Hermite discretisation behave similarly. This is expected since they essentially use the same space discretisation with different time approximations, and the time step is chosen small enough to neglect the errors coming from the time approximations. We can also observe that the Fourier discretisation converges at about the same speed as the one based on Hermite functions. In some case (especially for small T), we observed that the convergence can be even quicker. We guess that this can be explained by the fact that the n^{th} eigenvalue

of the Laplacian (on a bounded domain) is of order n^2, whereas it is of order n for the harmonic oscillator. Nevertheless, the Fourier discretisation converges to a biased solution because of the truncation of the level.

To take into account the fact that the Fourier solution is biased, we compute the error between the solutions of the Fourier scheme with the most precise solution computed with the Hermite discretisation. The result is plotted in Figure 3.2 for several domains $[-L_x, L_x]$ (for the Fourier discretisation).

We now observe that, provided the domain of integration is large enough, the Fourier discretisation performs quite well. This remark is all the more interesting since the discrete Fourier transform can be implemented very efficiently. Nevertheless, since the spread of the solution is varying for each realization of the solution of (3.1), the choice of the domain should be adapted to each realization. This is the reason why the Hermite transform can be very interesting (despite its cost). Moreover, as we show thereafter, the time discretisation is actually the limiting approximation. For example, in Figure 3.2, it is useless to plot the error for larger values of L_x since the error generated by the time discretisation for both splitting methods becomes greater than the error generated by the space discretisation.

We now look at the convergence with respect to the time discretisation in the case for a spectral Hermite discretisation. We compare the Crank-Nicolson discretisation with the above time-splitting discretisation. We choose a geometrically decreasing sequence of time steps, and we plot the average over 100 samples of the square of the error in $L^2(\mathbb{R})$-norm between the solutions computed with two successive time steps. The parameters are the number of modes $K \in \{40, 80, 120, 160, 200\}$, and the coefficient that we use in front of the Stratonovich integral (we call it α and choose $\alpha \in \{0.2, 0.4, 0.6, 1\}$). In practice, α should be chosen small since it models the intensity of the anomalous fluctuations of the confining potential. Since the stochastic integral raised some difficulties about the convergence analysis, we display some numerical results with large values of α which are not physically relevant. We plot the results in Figure 3.3 for $T = 4$. Each subplot corresponds to a value of α, and in each subplot we plot the mean square error for all the values of K, and for the two kinds of discretisation. First, for $\alpha = 0.2$ we can observe that the Crank-Nicolson scheme performs well, which justifies its practical use. Nevertheless, we can observe that the Splitting method is consistently doing better than the Crank-Nicolson discretisation, even for small values of α. This is surprising since the Crank-Nicolson scheme performs usually better than the splitting scheme in the deterministic setting (see [7]). Moreover, the Crank-Nicolson scheme is performing badly for high values of α, contrary to the splitting method which is much less sensitive to this parameter. In addition, the error is increasing with the number of modes for the largest α.

We believe that this rather bad numerical behaviour, when α is large, comes from the fact that the matrices $C(\delta t, \chi^{n+1})$ and $D(\delta t, \chi^{n+1})$ respectively defined by (3.93) and (3.94) have diagonal terms of order $1 \pm i\sqrt{\delta t}K/2$. For $K = 200$ and $\delta t = 5 \cdot 2^{-18}$ (which are the values used in Figure 3.3), we obtain that $\sqrt{\delta t}K/2 \approx 0.44$, which is not much smaller that one. Thus it is possible that the Crank-Nicolson approximation is not precise enough

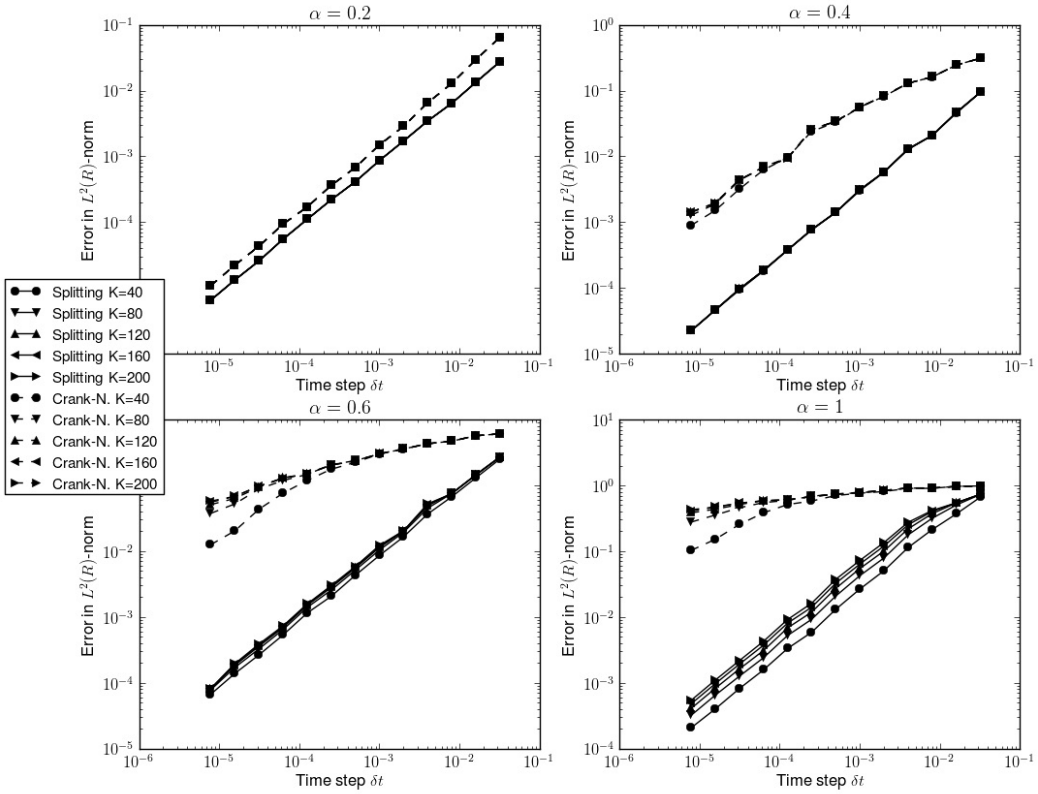

Figure 3.3 – Mean square convergence with respect to the time step for the Crank-Nicolson and the splitting discretisations

in this regime for the highest modes (that is to say that K is too large with respect to δt). We recall that the theoretical analysis imposes that $K\delta t^{1/4}$ stays bounded. In practice, this assumption imposes that the matrices $C(\delta t, \chi^{n+1})$ and $D(\delta t, \chi^{n+1})$ tend to be diagonal when the time step vanishes.

Acknowledgement : This work was supported by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH. The author is grateful to A. de Bouard for helpful discussions.

3.6 Appendix

Proof of Proposition 3.1. We begin by showing local existence and uniqueness in the space $\Sigma^k(\mathbb{R}^d)$, $\forall k \in \mathbb{N}^*$, $k \geq 2$, assuming that $\phi_0 \in \Sigma^k(\mathbb{R}^d)$. Let $T_0 > 0$. We use Equation (3.1) with the following gauge transformation $\phi(t,x) = e^{-iG(t,x)}\psi(t,x)$ with

$G(t,x) = \frac{1}{2}|x|^2 (t + \varepsilon W(t))$. The equation becomes

$$i\partial_t \psi = -(\nabla - ix(t + \varepsilon W(t)))^2 \psi + \lambda |\psi|^2 \psi. \tag{3.100}$$

We denote by $U^\omega(t,s)$ the propagator for the linear equation,

$$i\partial_t \psi = -(\nabla - ix(t + \varepsilon W(t)))^2 \psi. \tag{3.101}$$

We introduce T_ω, a positive random constant introduced in Proposition 4 [56], such that Propositions 6, 7 and Lemma 4.1 in [56] holds. In order to show the local existence and uniqueness in $\Sigma^k(\mathbb{R}^d)$, we use a classical fixed point argument based on application $\mathcal{T}^\omega$ defined by:

$$(\mathcal{T}^\omega \psi)(t) = U^\omega(t,0)\phi_0 - i\lambda \int_0^t U^\omega(t,s) |\psi|^2 \psi(s) ds.$$

To define the domains, we suppose that $(r,4)$ is an admissible pair, *i.e.* $r = 8/d$, and we set $\theta = d/4$. We define the following spaces for $T \leq T_0 \wedge T_\omega$ where T_0 is fixed.

$$\begin{aligned} X_T &= L^\infty(I, L^2) \cap L^r(I, L^4), \\ Y_T^k &= \{v \in X_T / \quad x^\alpha \partial^\beta v \in X_T, |\alpha| + |\beta| \leq k\}, \\ \widetilde{Y}_T^k &= \{v / \quad x^\alpha \partial^\beta v \in L^1(I, L^2) + L^{r'}(I, L^{\frac{4}{3}}), |\alpha| + |\beta| \leq k\}, \end{aligned}$$

where $I = [0,T]$, and where α and β are two multi-indices. We now show that there exist $C_1(T_0,k,\omega) > 0$, $C_2(T_0,k,\omega) > 0$ and $C_3(T_0,k,\omega) > 0$ so that for all $\psi_1, \psi_2 \in Y_T^k$:

$$\|\mathcal{T}^\omega \psi_1\|_{Y_T^k} \leq C_1 \|\phi_0\|_{\Sigma^k} + C_2 T^{1-\theta} \|\psi_1\|_{Y_T^k}^3 \tag{3.102}$$

$$\|\mathcal{T}^\omega \psi_1 - \mathcal{T}^\omega \psi_2\|_{L^r(I;L^4)} \leq C_3 T^{1-\theta} (\|\psi_1\|_{Y_T^k}^2 + \|\psi_2\|_{Y_T^k}^2) \|\psi_1 - \psi_2\|_{L^r(I;L^4)}, \tag{3.103}$$

To prove this, we use the following lemmas:

Lemma 3.24. *Let $k \in \mathbb{N}^*$, and assume $d \leq 3$. Then there exists $C(k) > 0$ such that for all $\phi \in \Sigma^{k,4}(\mathbb{R}^d) \cap \Sigma^{k,2}(\mathbb{R}^d)$,*

$$\left\| |\phi|^2 \phi \right\|_{\Sigma^{k,\frac{4}{3}}} \leq C(k) \|\phi\|_{\Sigma^{k,4}} \|\phi\|_{\Sigma^{k,2}}^2, \tag{3.104}$$

and for all $\phi \in \Sigma^{k+1}(\mathbb{R}^d)$,

$$\left\| |\phi|^2 \phi \right\|_{\Sigma^{k+1}} \leq C(k) \|\phi\|_{\Sigma^{k,4}}^2 \|\phi\|_{\Sigma^{k+1}}. \tag{3.105}$$

Lemma 3.25. *Let $k \in \mathbb{N}^*$. Then, for all $\phi_0 \in \Sigma^k(\mathbb{R}^d)$, $t \mapsto U^\omega(t,0)\phi_0 \in \mathcal{C}(I, \Sigma^k(\mathbb{R}^d)) \cap Y_T^k$. Moreover,*

$$\|U^\omega(\cdot,0)\phi_0\|_{Y_T^k} \leq C_1(T_0,k,\omega) \|\phi_0\|_{\Sigma^k}.$$

For all $\phi \in Y_T^k$, $|\phi|^2 \phi \in \widetilde{Y}_T^k$ *and*

$$\left\| \Lambda_\omega(\cdot,0) |\phi|^2 \phi \right\|_{Y_T^k} \leq C_2(T_0,k,\omega) \left\| |\phi|^2 \phi \right\|_{\widetilde{Y}_T^k},$$

where $\Lambda_\omega(\cdot,0)|\phi|^2\phi$ *is defined for all* $t \in I$ *by,*

$$\Lambda_\omega(t,0) = \int_0^t U^\omega(t,s) |\phi_s|^2 \phi_s ds.$$

Proof of Lemma 3.24. Equation (3.105) can be shown using Sobolev embedding theorems and Holder's inequality. To show Equation (3.104), we use the fact that there exists $C(k) > 0$ such that for all multi-indices α and β such that $|\alpha| + |\beta| \leq k$,

$$\left| x^\alpha \partial^\beta (|\phi|^2 \phi) \right| \leq C(k) \sum_{|l| \leq |\beta|, |j| \leq k-1} \left| x^\alpha \partial^l \phi \right| \left| \partial^j \phi \right|^2,$$

which leads, using Hölder's inequality and the Sobolev embedding $\Sigma^k(\mathbb{R}^d) \hookrightarrow H^k(\mathbb{R}^d) \hookrightarrow W^{k-1,4}(\mathbb{R}^d)$, to

$$\left\| x^\alpha \partial^\beta (|\phi|^2 \phi) \right\|_{L^{4/3}} \leq C(k) \|\phi\|_{\Sigma^{k,4}} \|\phi\|^2_{\Sigma^{k,2}}.$$

□

Proof of Lemma 3.25. By using k times the Proposition 6 of [56], we get that for all multi-index α, β such that $|\alpha| + |\beta| \leq k$,

$$\left\| x^\alpha \partial^\beta U^\omega(\cdot,0)\phi_0 \right\|_{X_T} \leq C(k,T_0) \sum_{|\gamma|+|\delta| \leq k} \left\| U^\omega(\cdot,0) x^\gamma \partial^\delta \phi_0 \right\|_{X_T},$$

and using Proposition 7 [56] that relies on the fact that U^ω in an isometry in $L^2(\mathbb{R}^d)$ it comes,

$$\left\| x^\alpha \partial^\beta U^\omega(\cdot,0)\phi_0 \right\|_{X_T} \leq C(k,T_0) \sum_{|\gamma|+|\delta| \leq k} \left\| x^\gamma \partial^\delta \phi_0 \right\|_{L^2},$$

which proves the first point. The second one can be shown in a similar way. □

To show (3.102), Lemma 3.24 and Hölder's inequality give

$$\left\| |\phi|^2 \phi \right\|_{\widetilde{Y}_T^k} \leq C(k) T^{1-\theta} \|\phi\|_{L^r(I;\Sigma^{k,4})} \|\phi\|^2_{L^\infty(I;\Sigma^{k,2})}$$
$$\left\| |\phi|^2 \phi \right\|_{\widetilde{Y}_T^k} \leq C(k) T^{1-\theta} \|\phi\|^3_{Y_T^k}$$

To show (3.103), we use the fact that there exists $C > 0$ so that for all $u, v \in \mathbb{C}$,

$$\left||u|^2 u - |v|^2 v\right| \leq C(|u|^2 + |v|^2)\,|u - v|$$

Using Strichartz estimates given in Proposition 7 [56] and Hölder inequality,

$$\begin{aligned}\|\mathcal{T}^\omega\psi - \mathcal{T}^\omega\phi\|_{L^r(I;L^4)} &\leq C(\omega, T_0)\left\||\psi|^2\psi - |\phi|^2\phi\right\|_{L^{r'}(I;L^{4/3})}\\ &\leq C(\omega,T_0)(\|\psi\|^2_{L^\infty(I;L^4)} + \|\phi\|^2_{L^\infty(I;L^4)})\,\|\psi-\phi\|_{L^{r'}(I;L^4)}\\ &\leq T^{1-\theta}C(\omega,T_0)(\|\psi\|^2_{Y_T^k} + \|\phi\|^2_{Y_T^k})\,\|\psi-\phi\|_{L^r(I;L^4)}\end{aligned}$$

Then T can be chosen small enough so that $\mathcal{T}^\omega$ is a contraction in

$$B_M := \left\{\phi \in Y_T^k / \quad \|\phi\|_{Y_T^k} \leq M\right\}$$

with $M = 2C(T_0, k, \omega)\,\|\phi_0\|_{\Sigma^k}$.

We now can show by induction that this local solution is actually global in time. Let $m \in \mathbb{N}$ such that $2 \leq m < k$. Since we know that the solution is global in the case $m = 2$ (see Proposition 4 [57]), we are going to show that if the solution is global in Σ^m, then it is global in Σ^{m+1}. To prove it, we begin by showing that it is global in $\Sigma^{m,4}(\mathbb{R}^d)$. It can be shown by induction using proposition 6 [56] for $\phi \in Y_T^k$ and all multi-index α, β so that $|\alpha| + |\beta| \leq m$ and $\forall (s,t) \in [0,T]^2$, $s < t$:

$$\left\|x^\alpha\partial^\beta U^\omega(t,s)\phi(s)\right\|_{L^4} \leq C(\omega,T_0,k)\,((t-s)^m \vee 1)\sum_{|\delta|+|\gamma|\leq m}\left\|U^\omega(t,s)x^\delta\partial^\gamma\phi(s)\right\|_{L^4}$$

Then, using lemma 4.1 [56],

$$\begin{aligned}&\left\|x^\alpha\partial^\beta\int_0^t U^\omega(t,s)\,|\phi(s)|^2\,\phi(s)ds\right\|_{L^4}\\ &\leq C(\omega,T_0,k)\int_0^t((t-s)^m\vee 1)\sum_{|\delta|+|\gamma|\leq m}|t-s|^{-d/4}\left\|x^\delta\partial^\gamma\,|\phi(s)|^2\,\phi(s)\right\|_{L^{4/3}}ds, \quad (3.106)\\ &\leq C(\omega,T_0,k)\int_0^t((t-s)^m\vee 1)\,|t-s|^{-d/4}\left\||\phi(s)|^2\,\phi(s)\right\|_{\Sigma^{m,4/3}}ds.\end{aligned}$$

We notice that $|t-s|^{-d/4}$ is integrable in s since $d \leq 3$. Lemma 3.24 and Gronwall's inequality enables us to conclude that the local solution is actually global in $\Sigma^{m,4}(\mathbb{R}^d)$.

We now show that the local solution is actually global in $\Sigma^{m+1}(\mathbb{R}^d)$ using again the mild formulation:

$$\|\psi(t)\|_{\Sigma^{m+1}} \leq \|U^\omega(t,0)\phi_0\|_{\Sigma^{m+1}} + \left\|\Lambda^\omega(t,0)(|\psi|^2\,\psi)\right\|_{\Sigma^{m+1}}.$$

Using proposition 7 [56], and reasoning in a similar way than previously, we can show:

$$\|\psi(t)\|_{\Sigma^{m+1}} \leq C(\omega, T_0, k)\, \|\phi_0\|_{\Sigma^{m+1}} + C(\omega, T_0, k) \int_0^t \left\| |\psi(s)|^2 \psi(s) \right\|_{\Sigma^{m+1}} ds.$$

We now use Lemma 3.24 and Gronwall's inequality to conclude that the solution is global in Σ^{m+1}. □

Chapitre 4

Vortex solutions in BEC under a trapping potential varying randomly in time

The aim of this chapter is to perform a theoretical and numerical study of the dynamics of vortices in Bose-Einstein condensates in the case when the trapping potential varies randomly in time. We take a deterministic vortex solution as an initial condition for the stochastically fluctuated Gross-Pitaevskii equation, and we observe the influence of the stochastic perturbation on the evolution. We theoretically prove that up to times of the order of ε^{-2}, the solution having the same symmetry properties as the vortex decomposes into the sum of a randomly modulated vortex solution and a small remainder, and we derive the equations for the modulation parameter. In addition, we show that the first order of the remainder, as ε goes to zero, converges to a Gaussian process. Finally, some numerical simulations on the dynamics of the vortex solution in the presence of noise are presented.

This chapter corresponds to the article [57] "Vortex solutions in Bose-Einstein condensation under a trapping potential varying randomly in time", published in *Discrete and Continuous Dynamical Systems - Series B*.

4.1 Introduction

We study the Gross-Pitaevskii equation for a two dimensional Bose gas in a randomly varying confinement. The first experimental realization of Bose-Einstein condensation in weakly interacting gases (e.g., [31, 50]) sparked off many theoretical and experimental studies on coherent atomic matter. The Schrödinger equation with cubic nonlinearity and a harmonic potential,

$$i\partial_t u = -\frac{\hbar^2}{2M}\Delta u + V(x)u + \lambda|u|^{2\sigma}u, \quad V(x) = |x|^2, \quad \sigma = 1 \tag{4.1}$$

called Gross-Pitaevskii equation, was initially used as a model equation. The Bose gas is described by u, the wave function of the condensate, $\hbar$ is Planck's constant, M the atomic mass of atoms in the condensate, and λ an interaction strength parameter. For example, in [139] a quantitative argument of the quantum ground state for magnetically trapped Bose gas can be found. Since then, Bose-Einstein condensates have been extensively studied and various model equations depending on the experimental situations have been introduced. We will treat mathematically a model of the condensate in, so-called, all-optical far-off resonance laser trap [1].

From the point of view of nonlinear waves, the interesting phenomenon is that the Gross-Pitaevskii equation, similarly to other nonlinear dispersive equations, supports various types of solitary wave solutions. In the two-dimensional setting we will study in particular, there are vortex solutions of the form

$$u(t, r, \theta) = e^{-i\mu t} e^{im\theta} \psi(r), \tag{4.2}$$

where r, θ are polar coordinates, m is the vortex degree, μ is the chemical potential and $\psi(r)$ is the radial positive vortex profile. Stability of vortex solutions to diverse forms of nonlinear Schrödinger equations has drawn much attention in recent years. For example, in the case $V \equiv 0$, Mizumachi in [126, 127] investigated the orbital stability and instability of the vortex solution to (4.1) with $\lambda = -1$, making use of the perturbation with the same symmetry as the vortex, which will be called in this chapter "m-equivariant" perturbation. Also it was proved in the case of general perturbations that the vortex solution is unstable for any $\sigma > 0$ if m is sufficiently large. Some numerical observations are available too, e.g., see [41] for the computation of the spectrum of the associated linearized operator at the vortex, [135, 145] for the case of other nonlinearities and [77] for the blow-up profile. For the case where $V \neq 0$, especially in the physically important case $V(x) = |x|^2$, a variational approach is used to give stability results in [120, 156]. It was proved in [120] that in the case $\lambda = 1$, there is a stable vortex solution for any $m \geq 1$ by m-equivariant perturbations. The author [156] showed that in the defocusing case $\lambda = 1$, if $m \geq 2$, there is a direction in which the second derivative of the action functional at the vortex is negative. However this is not sufficient to conclude to the nonlinear instability. None of these results give a satisfactory answer for the case $m = 1$ where we expect intuitively the stability of the vortex under any sort of perturbation. Here, we note that the vortex solution seems to be stable for $m = 1$ and unstable for $m \geq 2$ in the case $\lambda = 1$ from the numerical studies in [16]. More recent studies can be found in [108, 109], which deal with the spectrum of the linearized operator at the vortex solution in the case $\lambda = 1$. A bound for azimuthal Fourier modes which can cause the instability is given there. Numerical computations combined with the help of Krein signature allow to detect all unstable eigenvalues and infer a collision process between eigenvalues for $m = 2, 3$. This complexity of the spectrum structure makes it difficult to prove theoretically the stability, and we believe it is still challenging and important to push further and complete rigorously all those discussions.

In the present chapter, we are interested in the influence of the noise on the vortex solution (4.2), in the Gross-Pitaevskii equation with a stochastic perturbation of the following form:

$$i\partial_t u = -\Delta u + V(x)u + \lambda|u|^{2\sigma}u + \varepsilon K(x)u\dot{\xi}(t), \quad t \geq 0, \quad x \in \mathbb{R}^d, \tag{4.3}$$

where $\lambda = \pm 1$ and $\dot{\xi}$ is a white noise in time with correlation function $\mathbb{E}(\dot{\xi}(t)\dot{\xi}(s)) = \delta_0(t-s)$. Here, δ_0 denotes the Dirac measure at the origin, $\sigma > 0$ and $\varepsilon > 0$. Remark that we may set $\hbar = 2M = 1$ in (4.1) for our analysis without loss of generality. The product arising in the right hand side is interpreted in the Stratonovich sense, since the noise here naturally arises as the limit of processes with nonzero correlation length. We moreover assume that the noise is real valued. The term $\varepsilon\dot{\xi}(t)$ represents the deviations of the laser intensity $E(t)$ around its mean value (see [1]). Also, the sign of λ is related to the sign of the atomic scattering length, which may be positive or negative. It may be assumed without loss of generality that $\lambda = \pm 1$. This model is proposed in [1], possibly with the addition of a damping term, to describe Bose-Einstein condensate wave function in an all-optical far-off resonance laser trap, arguing that some fluctuations of the laser intensity are observed in this case, and that one should take into account stochasticity in the dynamical behavior of the condensate in real situations (see also [85, 155]). Related equations are also found in the context of optical fibers [2] in order to model the propagation of the optical soliton in fibers with random inhomogeneities.

Our aim in this chapter is to investigate the influence of random perturbations on the propagation of deterministic vortex solutions (4.2) theoretically and numerically. Not only rigorous mathematical results about the stability questions on deterministic vortices in nonlinear Schrödinger equations are rather few, but also the studies on the effect of stochastic perturbations on vortices are quite rare. Because of the presence of noise, a stable vortex would not persist in its form for all time. Thus an interesting question is how long the stable vortex can persist, compared to the noise strength ε. The method we will use, so called collective coordinate approach, consists in writing that the main part of the solution is given by a modulated vortex and finding then the modulation equations for the vortex parameters. Such ideas to analyze the asymptotic behavior have been used by many authors in the physics literature, as well as in the study of mathematical problems, but mainly for ground states (see, for example, in the deterministic case [79, 104, 168] and in the stochastic case [52, 53, 55, 58]). The stability property of the vortex solutions in the deterministic equation is required in order to apply such collective coordinate approaches, and we will here take advantage of the fact that the m-equivariance property of the solutions is preserved by the noise. Finally, we use numerical simulations to investigate the sharpness of the bounds that we obtain theoretically, both with respect to time T and to the strength of the noise ε.

The chapter is organized as follows: in Section 2, we state precisely our results. In Section 3, the existence of the modulation parameter is justified and we give an estimate on the time up to which the modulation procedure is available. In Section 4 we give the

equations of the modulation parameter. Section 5 is devoted to some estimates on the remainder term. Using these estimates, we will see the convergence of the remainder term as ε goes to zero to a limit process. Note that most arguments follow the ideas of [55], and we will give some technical explanations concerning the differences with these previous works in Section 8 (Appendix). The numerical results are presented in Section 6. We sometimes denote all through the chapter by $C_{\theta,\dots}$ a constant which depends on θ and so on.

4.2 Preliminaries and main results

We consider the following stochastic nonlinear Schrödinger equation

$$idu + (\Delta u - |x|^2 u - \lambda |u|^{2\sigma} u)dt = \varepsilon |x|^2 u \circ dW, \tag{4.4}$$

where $\circ$ stands for a Stratonovich product in the right hand side of (4.4), $\sigma > 0$, $\varepsilon > 0$, and $\lambda = \pm 1$. The unknown function u is a random process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a standard filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathcal{F}_0$ is complete, $W(t)$ is a standard real valued Brownian motion on $\mathbb{R}^+$ associated with the filtration $(\mathcal{F}_t)_{t\geq 0}$. We have set $\dot{\xi} = \frac{dW}{dt}$, $V(x) = K(x) = |x|^2$ in the equation (4.3). It will be useful to introduce here the equivalent Itô equation which may be written as

$$idu + \left(\Delta u - |x|^2 u + \frac{i}{2}\varepsilon^2 |x|^4 u - \lambda |u|^{2\sigma} u\right) dt = \varepsilon |x|^2 u dW. \tag{4.5}$$

We define

$$\Sigma^k = \left\{ v \in L^2(\mathbb{R}^d), \sum_{|\alpha|+|\beta|\leq k} |x^\beta \partial_x^\alpha v|^2_{L^2(\mathbb{R}^d)} = |v|^2_{\Sigma^k} < +\infty \right\}$$

for $k \in \mathbb{N}$, and write Σ^{-k} for the dual space of Σ^k in the L^2 sense. In particular we denote Σ^1 by Σ.

We define the energy

$$H(u) = \frac{1}{2}|\nabla u|^2_{L^2} + \frac{1}{2}|xu|^2_{L^2} + \frac{\lambda}{2\sigma+2}|u|^{2\sigma+2}_{L^{2\sigma+2}}, \tag{4.6}$$

which is a conserved quantity of the deterministic equation i.e., (4.4) with $\varepsilon = 0$. We will consider solutions in the space Σ since $H(u)$ is well defined in Σ, thanks to the Sobolev embedding $\Sigma \subset H^1(\mathbb{R}^d) \subset L^{2\sigma+2}(\mathbb{R}^d)$, for $\sigma < \frac{2}{d-2}$ if $d \geq 3$ or $\sigma < +\infty$ if $d = 1, 2$. It is worth adding a remark that there is also another conserved quantity, namely L^2 norm

$$Q(u) = \frac{1}{2}|u|^2_{L^2}.$$

In the case where $\varepsilon = 0$, it is known that in the energy space Σ, equation (4.4) is locally well posed for $\lambda = \pm 1$, $\sigma < \frac{2}{d-2}$ if $d \geq 3$ or $\sigma < +\infty$ if $d = 1, 2$ and globally well posed if either $\lambda = 1$ or $\lambda = -1$ and $\sigma < 2/d$ (see [130]). In the case where $\varepsilon \neq 0$, it was established in [56] a random Strichartz estimate and the authors proved the local and global well-posedness under the same condition for the nonlinear power as the deterministic case $\varepsilon = 0$. To rigorously state our results, we give some notation. If I is an interval of $\mathbb{R}$, E is a Banach space, and $1 \leq r \leq \infty$, then $L^r(I; E)$ is the space of strongly Lebesgue measurable functions v from I into E such that the function $t \to |v(t)|_E$ is in $L^r(I)$. We define similarly the spaces $C(I; E)$ and $L^r(\Omega; E)$.

Proposition 4.1. *Assume $\sigma > 0$, $\varepsilon > 0$ and $\lambda = \pm 1$. Let $u_0 \in \Sigma$ and $\sigma < 2/(d-2)$ if $d \geq 3$, $\sigma < +\infty$ if $d = 1, 2$. Then there exist a stopping time $\tau^*_{u_0,\omega} > 0$ and a unique solution $u(t)$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ of (4.4) with $u(0) = u_0$ almost surely in $C([0, \tau^*_{u_0,\omega}); \Sigma)$. Moreover, we have almost surely,*

$$\tau^*_{u_0,\omega} = +\infty \ \text{ or } \ \limsup_{t \nearrow \tau^*_{u_0,\omega}} |u(t)|_\Sigma = +\infty,$$

and the L^2 norm is conserved:

$$Q(u(t)) = Q(u_0), \ \text{ a.s. in } \omega, \ \text{ for all } t \in [0, \tau^*_{u_0,\omega}).$$

In particular, if $\lambda = 1$, then there exists a unique global solution $u(t)$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ of (4.4) with $u(0) = u_0$ almost surely in $C(\mathbb{R}^+; \Sigma)$.

The result of Proposition 4.1 holds, not only for a quadratic potential, but for more general potentials $K(x)$ and $V(x)$ in (4.3) (see [56] for details). The key point of the proof of Proposition 4.1 is the following gauge transformation: let $w(t)$ be a solution of

$$i\partial_t w = -\left(\nabla - ix(t + \varepsilon W(t))\right)^2 w, \tag{4.7}$$

then

$$u(t, x) = \exp\Big\{ -\frac{i}{2}|x|^2(t + \varepsilon W(t))\Big\} w(t, x) \tag{4.8}$$

satisfies Equation (4.4) with $\lambda = 0$. Therefore, the Cauchy problem for Eq.(4.4) (or equivalently (4.5)) is deduced from the Cauchy problem for

$$i\partial_t w = -\left(\nabla - ix(t + \varepsilon W(t))\right)^2 w + \lambda |w|^{2\sigma} w, \tag{4.9}$$

which is a *deterministic* equation for each fixed $\omega \in \Omega$.

In addition to the results of Proposition 4.1, in this chapter, we will make use of the following fact. There is an explicit representation of the solution of the linear part of

Eq.(4.4), which is a derivation from the result in [160]. To introduce it, let $T_0 > 0$ be fixed and consider $(q(t), v(t)) \in \mathbb{R}^2$ solution of the system:

$$\begin{cases} \dot{q}(t) &= v(t) - (t + \varepsilon W(t))q(t), \\ \dot{v}(t) &= (t + \varepsilon W(t))(v(t) - (t + \varepsilon W(t))q(t)) \end{cases}$$

with initial data $(q(0), v(0)) = (0, 1) \in \mathbb{R}^2$. There exists a unique solution $(q, v) \in C([0, T_0], \mathbb{R}^2)$ for each $\omega \in \Omega$ satisfying $W(\cdot, \omega) \in C^\alpha([0, T_0])$ with $0 < \alpha < 1/2$. Using this solution, we define the following system of ODEs:

$$\begin{cases} 2\alpha(t) &= v(t)/q(t), \\ \dot{\beta}(t) &= (t + \varepsilon W(t) - \alpha(t))\,\beta(t), \\ \dot{\gamma}(t) &= -\frac{1}{2}\beta^2(t) \end{cases}$$

Remark that $\alpha(t), \beta(t), \gamma(t)$ are well defined for any $t \leq \tilde{T} \wedge T_0$, where $\tilde{T} = \inf\{s > 0, q(s) = 0\}$. Note that $\tilde{T} > 0$, thanks to the initial condition $(q(0), v(0)) = (0, 1)$. Then we have the following representation formula for the solution of (4.7).

Proposition 4.2. *Let $T_0 > 0$ be fixed. For each $\omega \in \Omega$ satisfying $W(\cdot, \omega) \in C^\alpha([0, T_0])$ with $0 < \alpha < 1/2$, the solution of (4.7) with initial data $w(0) = u_0 \in C_0^\infty(\mathbb{R}^d)$ may be expressed as follows.*

$$w(t, x) = \frac{1}{(2\pi i q(t))^{d/2}} \int_{\mathbb{R}^d} e^{i(\alpha(t)|x|^2 + \beta(t)x\cdot y + \gamma(t)|y|^2)} u_0(y) dy, \tag{4.10}$$

for any $t \leq T_0 \wedge \tilde{T}$.

Remind that our interest is in the vortex solutions. From now on we fix $\lambda = 1$, so that we consider the equation;

$$idu + \left(\Delta u - |x|^2 u + \frac{i}{2}\varepsilon^2 |x|^4 u - |u|^{2\sigma} u\right) dt = \varepsilon |x|^2 u dW. \tag{4.11}$$

This equation can be regarded as a stochastically perturbed version of the following deterministic equation;

$$i\partial_t u + \Delta u - |x|^2 u - |u|^{2\sigma} u = 0. \tag{4.12}$$

We also restrict ourselves to the two dimensional case $d = 2$, and make use of the polar coordinates : for $x = (x_1, x_2) \in \mathbb{R}^2$, we define $x_1 = r\cos\theta$, $x_2 = r\sin\theta$. We recall that the vortex solutions are the solutions of (4.12) of the form,

$$u(t, x) = e^{-i\mu t}\phi_{\mu,m}(x) = e^{-i\mu t} e^{im\theta} \psi_{\mu,m}(r), \quad \mu \in \mathbb{R}, \quad x \in \mathbb{R}^2, \tag{4.13}$$

where $m \in \mathbb{Z}^*$ is the winding number of the vortex, and we may assume $m \geq 1$ without loss of generality by the reflection symmetry. Substituting (4.13) in (4.12), $\phi_{\mu,m}(x) \in \Sigma$

satisfies

$$-\Delta\phi + |x|^2\phi - \mu\phi + |\phi|^{2\sigma}\phi = 0. \tag{4.14}$$

Before stating more details about the properties of vortex solutions (4.13), we introduce the closed subspace X_m of Σ that has the same symmetry, i.e., m-equivariant symmetry as the vortex solutions $\phi_{\mu,m}(x) = e^{im\theta}\psi_{\mu,m}(r)$:

$$X_m := \{v \in \Sigma, \quad e^{-im\theta}v(x) \text{ does not depend on } \theta\}.$$

We remark that, thanks to Proposition 4.2, one may check that the solution $u(t)$ of (4.11) belongs to X_m if the initial data $u(0) = u_0$ belong to X_m. This preservation of the structure of Eq.(4.11) in X_m leads to the following result.

Lemma 4.3. *Let $\sigma > 0$ and $u_0 \in X_m$. Then there exists a unique global solution $u(t)$ of (4.11) which is adapted to $(\mathcal{F}_t)_{t\geq 0}$, that satisfies the initial condition $u(0) = u_0$, and that belongs almost surely to $C(\mathbb{R}^+; X_m)$.*

We give a proof of Lemma 4.3 in Appendix using Proposition 4.2.

The existence of the vortex solutions is easily proved (see e.g. [120]), with the help of the compact embedding $X_m \subset \Sigma \subset L^2$, for any $\mu > \lambda_m$ where λ_m is the solution of the linear eigenvalue problem

$$\lambda_m = \inf\Big\{ \int_0^\infty \left(-\frac{d^2}{dr^2}v + r^2 v + \frac{1}{r^2}\left(m^2 - \frac{1}{4}\right)v\right)\bar{v}dr;\ v \in D,\ \int_0^\infty |v(r)|^2 dr = 1\Big\}, \tag{4.15}$$

where

$$D = \Big\{v \in L^2(0,\infty),\ \frac{d^2}{dr^2}v,\ r^2 v,\ \frac{1}{r^2}v \in L^2(0,\infty)\Big\}.$$

It is known that $\lambda_m = 2m + 2$ (see [163]). More precisely, let S_μ be the action functional, i.e.,

$$S_\mu(u) = H(u) - \mu Q(u), \quad u \in \Sigma.$$

The variational problem

$$\inf\{S_\mu(u), u \in X_m \setminus \{0\}\} \tag{4.16}$$

is attained by $\phi_{\mu,m}$ on X_m for any $\mu > \lambda_m$. Namely, $\phi_{\mu,m}$ is characterized as a global minimizer of S_μ on X_m. The uniqueness of minimizers is established as in [120]. Such a variational characterization allows to show the orbital stability in X_m for any $m \geq 1$. This fact is related to Proposition 4.4 (i) below.

The solutions ϕ of (4.14) have regularity properties by the usual elliptic bootstrap arguments, i.e.,

$$\lim_{|x|\to\infty} |\phi_{\mu,m}(x)| = 0, \text{ and } \phi_{\mu,m} \in \bigcap_{2\leq q<\infty} W^{2,q} \cap C^2.$$

The radial profile $\psi_{\mu,m}(r)$ is thus of class $C^2(0,\infty)$ and satisfies the equation

$$-\frac{d^2}{dr^2}\psi - \frac{1}{r}\frac{d}{dr}\psi + \frac{m^2}{r^2}\psi + r^2\psi - \mu\psi + |\psi|^{2\sigma}\psi = 0, \quad r > 0, \tag{4.17}$$

with

$$\lim_{r\to\infty} |\psi_{\mu,m}(r)| = 0.$$

In particular, the solution $\psi_{\mu,m}$ associated to the global minimizer $\phi_{\mu,m}$ of (4.16) is non-negative. This comes from the fact that if ψ is a minimizer of (4.16), then $|\psi|$ as well as seen from the radial form of (4.16), and

$$\int_0^\infty \left|\frac{d}{dr}|\psi|\right|^2 r dr = \int_0^\infty \left|\frac{d}{dr}\psi\right|^2 r dr.$$

Remark that every regular solution of (4.17) should satisfy $\psi(r) \to 0$, as $r \to 0^+$ (see [126]), which we will see again later in Proposition 4.4 below.

Now we summarize here some properties of $\psi_{\mu,m}(r)$. The inner product in the Hilbert space $L^2(\mathbb{R}^2)$ is denoted by $(\cdot,\cdot)_{L^2(\mathbb{R}^2,dx)}$, i.e.,

$$(u,v)_{L^2(\mathbb{R}^2,dx)} = \int_{\mathbb{R}^2} u(x)\overline{v(x)}\, dx, \quad \text{for} \quad u, v \in L^2(\mathbb{R}^2).$$

Moreover we denote $\langle u, v\rangle := \mathrm{Re}(u,v)_{L^2(\mathbb{R}^2,dx)}$. Taking account of the property of the space X_m, we here introduce the radial norm and the radial inner products, too.

$$(u,v)_{L^2_r} = \int_0^\infty u(r)\overline{v(r)}\, r dr, \quad |u|^2_{L^2_r} = \int_0^\infty |u(r)|^2\, r dr.$$

Proposition 4.4. *Let $\sigma > 0$, $m \geq 1$ and $\mu > \lambda_m$. Let $\phi_{\mu,m}(x)$ be the minimizer of (4.16), and consider the associated radial profile $\psi_{\mu,m}(r) := e^{-im\theta}\phi_{\mu,m}(x)$. Then the following properties hold.*

(i) There exist $\nu = \nu(\mu,m) > 0$, such that for any $v \in X_m$ satisfying

$$\mathrm{Re}(v, i\phi_{\mu,m})_{L^2(\mathbb{R}^2,dx)} = 0,$$

we have

$$\langle S''_\mu(\phi_{\mu,m})v, v\rangle \geq \nu|v|^2_\Sigma.$$

(ii) The radial profile $\psi_{\mu,m}(r)$ has the following asymptotics.

$$\psi_{\mu,m}(r) = O(r^{\mu/2}e^{-r^2/2}), \quad r \to +\infty, \quad \text{and} \quad \psi_{\mu,m}(r) = O(r^m), \quad r \to 0^+.$$

(iii) The radial profile $\psi_{\mu,m}(r)$ has a local maximum for some $r_0 \in (m/\sqrt{\mu}, \sqrt{\mu})$, is increasing on $(0, r_0)$ and is decreasing on (r_0, ∞). Moreover, $|\psi_{\mu,m}(r)|^{2\sigma} < \mu - 2m$ for all $r > 0$.

The positivity of $S''_\mu(\phi_{\mu,m})$, that is (i) of Proposition 4.4, is discussed in the Appendix; for the properties (ii) and (iii), see Appendix of [108] or [99].

Note that the second-order differential $S''_\mu(\phi_{\mu,m})$ is related to the linearization problem around $e^{-i\mu t}\phi_{\mu,m}(x)$ in (4.12), which is precisely written as

$$\frac{dy}{dt} = J\mathcal{L}_{\mu,m}y \quad \text{in} \quad \Sigma^{-1},$$

where

$$J = -i : \begin{pmatrix} \text{Re } u \\ \text{Im } u \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \text{Re } u \\ \text{Im } u \end{pmatrix},$$

$$\mathcal{L}_{\mu,m} = S''_\mu(\phi_{\mu,m}) = \begin{pmatrix} L^1_{\mu,m} & 0 \\ 0 & L^2_{\mu,m} \end{pmatrix}, \tag{4.18}$$

$$L^1_{\mu,m} = -\Delta + |x|^2 - \mu + (2\sigma+1)|\phi_{\mu,m}|^{2\sigma}, \quad L^2_{\mu,m} = -\Delta + |x|^2 - \mu + |\phi_{\mu,m}|^{2\sigma}.$$

In the radial notation, we will use, as the linearized operators corresponding to $L^1_{\mu,m}$ and $L^2_{\mu,m}$,

$$L^{1,r}_{\mu,m} = -\frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + r^2 + \frac{m^2}{r^2} - \mu + (2\sigma+1)\psi^{2\sigma}_{\mu,m}, \tag{4.19}$$

$$L^{2,r}_{\mu,m} = -\frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + r^2 + \frac{m^2}{r^2} - \mu + \psi^{2\sigma}_{\mu,m}. \tag{4.20}$$

Our purpose is to investigate the influence of random perturbations of the form given in equation (4.11) on the phase of vortex solutions (4.13).

We fix $\mu_0 > \lambda_m$ and consider for $\varepsilon > 0$ the solution $u^\varepsilon(t,x)$ of equation (4.11) given by Lemma 4.3 with initial data $u^\varepsilon(0,x) = \phi_{\mu_0,m}(x)$, which is a stable vortex in X_m of the deterministic equation.

The first theorem says that we can decompose u^ε in X_m as the sum of a modulated vortex solution and a remainder with small Σ norm, for t less than some stopping time τ^ε, and that this τ^ε goes to infinity in probability as ε goes to zero. We will see then that the remaining part is of order one with respect to ε. The proof of the theorem is rather similar to those in [52, 53, 55], and we only mention the differences with respect to the previous works. The decomposition we consider is of the form

$$u^\varepsilon(t,x) = e^{-i\xi^\varepsilon(t)}(\phi_{\mu_0,m}(x) + \varepsilon\eta^\varepsilon(t,x)) \tag{4.21}$$

for a semi-martingale process $\xi^\varepsilon(t)$ with values in $\mathbb{R}$, and η^ε with value in X_m. Here, we use the following orthogonality condition

$$\text{Re}(\eta^\varepsilon, i\phi_{\mu_0,m})_{L^2(\mathbb{R}^2,dx)} = 0, \quad \text{a.s.}, \quad t \le \tau^\varepsilon, \tag{4.22}$$

for a suitable stopping time τ^ε as long as $|\varepsilon\eta^\varepsilon|_\Sigma$ remains small. The precise result is the following.

Theorem 4.5. *Let $1/2 \leq \sigma < \infty$ and $\mu_0 > \lambda_m$ be fixed. For $\varepsilon > 0$, let $u^\varepsilon(t,x)$, as defined above, be the solution of (4.11) with $u(0,x) = \phi_{\mu_0,m}(x)$. Then there exists $\alpha_0 > 0$ such that, for each α, $0 < \alpha \leq \alpha_0$, there is a stopping time $\tau^\varepsilon_\alpha \in (0,\infty)$ a.s., and a semi-martingale process $\xi^\varepsilon(t)$, defined a.s. for $t \leq \tau^\varepsilon_\alpha$, with values in $\mathbb{R}$, so that if we set $\varepsilon\eta^\varepsilon(t,x) = e^{i\xi^\varepsilon(t)}u^\varepsilon(t,x) - \phi_{\mu_0,m}(x)$, then (4.22) holds. Moreover, a.s. for $t \leq \tau^\varepsilon_\alpha$,*

$$|\varepsilon\eta^\varepsilon(t)|_\Sigma \leq \alpha. \tag{4.23}$$

In addition, there is a constant $C = C_{\alpha,\mu_0} > 0$, such that for any $T > 0$ and any $\alpha \leq \alpha_0$, there is $\varepsilon_0 > 0$, for which

$$\forall \varepsilon < \varepsilon_0, \quad \mathbb{P}(\tau^\varepsilon_\alpha \leq T) \leq \exp\left(-\frac{C}{\varepsilon^2 T}\right). \tag{4.24}$$

This means that the modulation parameter $\xi^\varepsilon(t)$ is a semi-martingale process defined up to times of the order ε^{-2}, i.e. the shape of the vortex solution is preserved over this time scale. We give in the next theorem the behavior of η^ε, and the convergence of the modulation parameter as ε goes to zero.

Theorem 4.6. *Let $1 \leq \sigma < \infty$, $\mu_0 > \lambda_m$ be fixed and η^ε, ξ^ε, for $\varepsilon > 0$ be given by Theorem 4.5, with $\alpha \leq \alpha_0$ fixed. Then, for any $T > 0$, the X_m-valued process $(\eta^\varepsilon(t))_{t\in[0,T\wedge\tau^\varepsilon_\alpha]}$ converges in probability, as ε goes to zero, to a process $\eta = (\eta(t))_{t\in[0,T]}$ satisfying $e^{-im\theta}\eta = \tilde\eta$ and*

$$d\tilde\eta = J\begin{pmatrix} L^{1,r}_{\mu_0,m} & 0 \\ 0 & L^{2,r}_{\mu_0,m}\end{pmatrix}\tilde\eta dt - y\begin{pmatrix}0\\ \psi_{\mu_0,m}\end{pmatrix}dt - \begin{pmatrix}0\\ r^2\psi_{\mu_0,m}\end{pmatrix}dW - z\begin{pmatrix}0\\ \psi_{\mu_0,m}\end{pmatrix}dW, \tag{4.25}$$

with $\tilde\eta(0) = 0$, and

$$y(t) = \frac{2\sigma(\operatorname{Re}\tilde\eta, \psi^{2\sigma+1}_{\mu_0,m})_{L^2_r}}{|\psi_{\mu_0,m}|^2_{L^2_r}}, \quad z(t) = \frac{|r\psi_{\mu_0,m}|^2_{L^2_r}}{|\psi_{\mu_0,m}|^2_{L^2_r}}, \tag{4.26}$$

for all $t \geq 0$. The convergence holds in $C([0,\tau^\varepsilon_\alpha \wedge T], L^2)$.

The above process η satisfies for any $T > 0$ the estimate

$$\mathbb{E}\left(\sup_{t\leq T}|\eta(t)|^2_\Sigma\right) \leq CT$$

for some constant $C > 0$.

Moreover the modulation parameter ξ^ε may be written, for $t \leq \tau^\varepsilon_\alpha$, as

$$d\xi^\varepsilon = \mu_0 dt + \varepsilon y^\varepsilon dt + \varepsilon z^\varepsilon dW,$$

for some adapted processes y^ε, z^ε with values in $\mathbb{R}$, converging in probability in $C([0,T])$ respectively to y, z given in (4.26) as ε goes to zero.

Since z is deterministic, this shows that at first order in ε the noise acts on the phase of vortex in a Gaussian way.

Remark that in [55] the authors assumed $\sigma = 1$ in order to show a similar theorem. In the present chapter, Theorem 4.6 above gives an improvement in this respect, although $\sigma \geq 1$ is still essential due to the treatment of the estimate for the remainder η^ε in $L^4(\Omega, C([0,T],\Sigma))$. We will need the following proposition, whose proof is given in the Appendix.

Proposition 4.7. *Let $\lambda = 1$ and $1/2 \leq \sigma < \infty$. Let $u_0 \in \Sigma^2$. Then there exists a unique global solution $w(t)$ of (4.9) with $w(0) = u_0$, adapted to $(\mathcal{F}_t)_{t\geq 0}$, in $C(\mathbb{R}^+, \Sigma^2)$ almost surely.*

4.3 Proof of Theorem 4.5

In this section, we give the outline of the proof of the existence of modulation parameter and the estimate on the exit time (4.24). The arguments are similar to those in [53, 55]. The following lemma shows the conservation of the charge Q and of the energy H by the evolution of (4.11). For the proof, we refer to Theorem 3 (i) in [54].

Lemma 4.8. *Assume $0 < \sigma < \infty$ and $\mu_0 > \lambda_m$. Let $\phi_{\mu_0,m}$ be the vortex solution and u^ε be the solution of (4.11) given by Proposition 4.1, with $u^\varepsilon(0,x) = \phi_{\mu_0,m}$. Then for any stopping time τ we have*

$$|u^\varepsilon(\tau)|_{L^2} = |\phi_{\mu_0,m}|_{L^2}, \quad a.s., \tag{4.27}$$

$$\begin{aligned} H(u^\varepsilon(\tau)) = H(\phi_{\mu_0,m}) &- 2\varepsilon \mathrm{Im}\left(\int_0^\tau \int_{\mathbb{R}^d} \nabla u^\varepsilon \cdot x\bar{u^\varepsilon}\, dx\, dW(s)\right) \\ &+ 2\varepsilon^2 \int_0^\tau |xu^\varepsilon|_{L^2}^2 ds, \quad a.s. \end{aligned} \tag{4.28}$$

The implicit function theorem ensures the existence of the modulation parameter satisfying the orthogonality condition (4.22) . Indeed, let $B_{\phi_{\mu_0,m}}(2\alpha) = \{v \in X_m, |v - \phi_{\mu_0,m}|_\Sigma \leq 2\alpha\}$ for α with $0 < \alpha < \mu_0/4$. We then consider a C^2 mapping $\mathcal{I}$, from $(-2\alpha, 2\alpha) \times B_{\phi_{\mu_0,m}}(2\alpha)$ to $\mathbb{R}$, defined by

$$\mathcal{I}(\xi, u) = |e^{i\xi}u - \phi_{\mu_0,m}|_{L^2}^2.$$

One easily obtains

$$\partial_\xi \mathcal{I}(0, \phi_{\mu_0,m}) = 0, \quad \partial_\xi^2 \mathcal{I}(0, \phi_{\mu_0,m}) = 2|\phi_{\mu_0,m}|_{L^2}^2 > 0.$$

Applying the implicit function theorem shows that for $\alpha \leq \alpha_0$ where α_0 is sufficiently small, there exists a C^2 mapping $\xi(u)$ defined for $u \in B_{\phi_{\mu_0,m}}(2\alpha)$, such that

$$\partial_\xi \mathcal{I}(\xi(u), u) = 0.$$

Applying this with $u = u^\varepsilon(t)$ given by Lemma 4.3, we get the existence of $\xi^\varepsilon(t) = \xi(u^\varepsilon(t))$ such that the orthogonality condition (4.22) holds with $\varepsilon\eta^\varepsilon(t) = e^{i\xi^\varepsilon(t)}u^\varepsilon(t) - \phi_{\mu_0,m}$. Then the Itô formula shows, as in [55], that ξ^ε is a semi-martingale process. Moreover, since it is clear that $\partial_\xi \mathcal{I}(0, e^{i\xi^\varepsilon(t)}u^\varepsilon(t)) = 0$, the existence of ξ^ε holds as long as

$$|e^{i\xi^\varepsilon(t)}u^\varepsilon(t) - \phi_{\mu_0,m}|_\Sigma < \alpha.$$

We now define a stopping time

$$\tau_\alpha^\varepsilon = \inf\{t \geq 0, |e^{i\xi^\varepsilon(t)}u^\varepsilon(t) - \phi_{\mu_0,m}|_\Sigma \geq \alpha\}.$$

The process $\xi^\varepsilon(t)$ is defined for all $t \leq \tau_{\alpha_0}^\varepsilon$, and satisfies (4.23) for all $t \leq \tau_\alpha^\varepsilon$ and $\alpha \leq \alpha_0$, together with the orthogonality condition (4.22), where again $\varepsilon\eta^\varepsilon(t) = e^{i\xi^\varepsilon(t)}u^\varepsilon(t) - \phi_{\mu_0,m}$.

Let us fix $T > 0$. Setting $\tau = \tau_\alpha^\varepsilon \wedge T$, remark that

$$\mathbb{P}(\tau_\alpha^\varepsilon \leq T) \leq \mathbb{P}(|\varepsilon\eta^\varepsilon(\tau)|_\Sigma \geq \alpha). \tag{4.29}$$

Thus, similar, and in fact even simpler arguments than in [52, 58, 55] lead to the exponential estimate (4.24). Indeed, by Taylor expansion,

$$\begin{aligned} S_{\mu_0}(u^\varepsilon(t,\cdot) - S_{\mu_0}(\phi_{\mu_0,m}) =& S_{\mu_0}(e^{i\xi^\varepsilon(t)}u^\varepsilon(t,\cdot)) - S_{\mu_0}(\phi_{\mu_0,m}) \\ =& \langle S'_{\mu_0}(\phi_{\mu_0,m}), \varepsilon\eta^\varepsilon(t)\rangle + \langle S''_{\mu_0}(\phi_{\mu_0,m})\varepsilon\eta^\varepsilon(t), \varepsilon\eta^\varepsilon(t)\rangle \\ &+ o(|\varepsilon\eta^\varepsilon(t)|_\Sigma^2). \end{aligned}$$

Note that $o(|\varepsilon\eta^\varepsilon(t)|_\Sigma^2)$ is uniform in ε, ω and t for $t \leq \tau_\alpha^\varepsilon$, since $|e^{i\xi^\varepsilon(t)}u^\varepsilon(t,\cdot) - \phi_{\mu_0,m}|_\Sigma = |\varepsilon\eta^\varepsilon(t,\cdot)|_\Sigma \leq \alpha$ for all $t \leq \tau_\alpha^\varepsilon$. We then assume α_0 small enough so that the last term is less than $\frac{\nu}{2}|\varepsilon\eta^\varepsilon(t)|_\Sigma^2$ for all $t \leq \tau_\alpha^\varepsilon$. Since, the positivity of $S''_{\mu_0}(\phi_{\mu_0,m})$ is satisfied in X_m (see Proposition 4.4), for any $\mu_0 > \lambda_m$,

$$\langle S''_{\mu_0}(\phi_{\mu_0,m})\varepsilon\eta^\varepsilon, \varepsilon\eta^\varepsilon\rangle \geq \nu|\varepsilon\eta^\varepsilon|_\Sigma^2$$

holds a.s. for $t \leq \tau_\alpha^\varepsilon$, noticing that $\eta^\varepsilon \in X_m$ and satisfies the orthogonality condition (4.22). It thus follows, since $S'_{\mu_0}(\phi_{\mu_0,m}) = 0$,

$$S_{\mu_0}(u^\varepsilon(t,\cdot)) - S_{\mu_0}(\phi_{\mu_0,m}) \geq \frac{\nu}{2}|\varepsilon\eta^\varepsilon(t)|_\Sigma^2.$$

Again we choose α_0 sufficiently small, then make use of (4.27), (4.28) to get, for any

$\tau \leq \tau_\alpha^\varepsilon$

$$\begin{aligned}\frac{\nu}{2}|\varepsilon\eta^\varepsilon(\tau)|_\Sigma^2 \leq & S_{\mu_0}(u^\varepsilon(\tau,x)) - S_{\mu_0}(\phi_{\mu_0,m}) = H(u^\varepsilon(\tau,x)) - H(\phi_{\mu_0,m}) \\ = & -2\varepsilon \mathrm{Im} \int_0^\tau \int_{\mathbb{R}^d} \nabla u^\varepsilon(s,x)\cdot x\bar{u^\varepsilon}(s,x)\,dx\,dW(s) \\ & + 2\varepsilon^2 \int_0^\tau |xu^\varepsilon(s,\cdot)|_{L^2}^2 ds.\end{aligned} \tag{4.30}$$

The estimate on the right hand side of (4.29) follows from exactly the same arguments as in [55], thanks to the classical exponential tail estimates for 1D stochastic integrals, once we have noticed that, for any $t \in [0,\tau]$,

$$\left|\int_{\mathbb{R}^d} \nabla u^\varepsilon \cdot x\bar{u^\varepsilon}(t,x)dx\right|^2 \leq \sup_{t\in[0,\tau]} |u^\varepsilon(t)|_\Sigma^2 \leq C_{\alpha_0,\mu_0,m}, \quad \text{a.s.},$$

which concludes (4.24). To show such a result, the standard method consists in introducing for some positive γ and t, the exponential martingale $(M_t)_{t\geq 0}$ defined by,

$$\begin{aligned}\log(M_t^\gamma) = & \gamma \int_0^t -2\varepsilon \mathrm{Im} \int_{\mathbb{R}^d} \nabla u^\varepsilon(s,x)\cdot x\bar{u^\varepsilon}(s,x)\,dx\,dW(s) \\ & - \frac{\gamma^2}{2}\int_0^t \left(-2\varepsilon \mathrm{Im}\int_{\mathbb{R}^d} \nabla u^\varepsilon(s,x)\cdot x\bar{u^\varepsilon}(s,x)\,dx\right)^2 ds.\end{aligned}$$

The tail estimate follows from a Markov inequality and an optimisation on the parameter γ. □

4.4 Modulation equations and SDE for the remainder

In this section we formally derive the equation of the modulation parameter ξ^ε, and the remaining term η^ε. A justification of these formal computations is obtained similarly to [55] and thus we omit it. We fix α for which Theorem 4.5 holds and we write τ^ε for τ_α^ε from now on. Since ξ^ε is a semi-martingale process, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by $(W(t))_{t\geq 0}$, we may thus write a priori the equation for ξ^ε in the form

$$d\xi^\varepsilon = \mu_0 dt + \varepsilon y^\varepsilon dt + \varepsilon z^\varepsilon dW \tag{4.31}$$

where y^ε is a real valued adapted process with paths in $L^1(0,\tau^\varepsilon)$ a.s., z^ε is a real valued predictable process, with paths in $L^2(0,\tau^\varepsilon)$ a.s.

We write $\phi_{\mu_0,m}(x) = e^{im\theta}\psi_{\mu_0,m}(r)$ and $\eta^\varepsilon(t,x) = e^{im\theta}\tilde{\eta}^\varepsilon(t,r)$. Recall that $\psi_{\mu_0,m}(r)$ is a real-valued function satisfying (4.17) with $\mu = \mu_0$, and $\tilde{\eta}^\varepsilon(t,r)$ is a complex-valued radial function since $\eta^\varepsilon(t,\cdot) \in X_m$. Note that the orthogonality condition (4.22) may be written as

$$(\mathrm{Im}\,\tilde{\eta}^\varepsilon, \psi_{\mu_0,m})_{L_r^2} = 0. \tag{4.32}$$

Lemma 4.9. *Let $\tilde\eta^\varepsilon = \tilde\eta^\varepsilon_R + i\tilde\eta^\varepsilon_I$, where $\tilde\eta^\varepsilon_R = \operatorname{Re}\tilde\eta^\varepsilon$ and $\tilde\eta^\varepsilon_I = \operatorname{Im}\tilde\eta^\varepsilon$. With the above notation, for $\sigma \geq 1/2$, $\tilde\eta^\varepsilon_R$ and $\tilde\eta^\varepsilon_I$ satisfy the equations*

$$\begin{aligned} d\tilde\eta^\varepsilon_R =& L^{2,r}_{\mu_0,m}\tilde\eta^\varepsilon_I dt \\ &+ \varepsilon\left(-y^\varepsilon\tilde\eta^\varepsilon_I dt - \frac{1}{2}(z^\varepsilon)^2\psi_{\mu_0,m}dt + z^\varepsilon r^2\psi_{\mu_0,m}dt - \frac{1}{2}r^4\psi_{\mu_0,m}dt\right. \\ &\left.+h^\varepsilon_I dt - \frac{1}{2}(z^\varepsilon)^2\psi_{\mu_0,m}dt - z^\varepsilon\tilde\eta^\varepsilon_I dW + r^2\tilde\eta^\varepsilon_I dW\right) \\ &+ \varepsilon^2\left(z^\varepsilon r^2\tilde\eta^\varepsilon_R dt - \frac{1}{2}(z^\varepsilon)^2\tilde\eta^\varepsilon_R dt - \frac{1}{2}r^4\tilde\eta^\varepsilon_R dt\right), \end{aligned} \tag{4.33}$$

$$\begin{aligned} d\tilde\eta^\varepsilon_I =& -L^{1,r}_{\mu_0,m}\tilde\eta^\varepsilon_R dt + y^\varepsilon\psi_{\mu_0,m}dt - r^2\psi_{\mu_0,m}dW + z^\varepsilon\psi_{\mu_0,m}dW \\ &+ \varepsilon\left(y^\varepsilon\tilde\eta^\varepsilon_R dt - h^\varepsilon_R dt - r^2\tilde\eta^\varepsilon_R dW + z^\varepsilon\tilde\eta^\varepsilon_R dW\right) \\ &+ \varepsilon^2\left(-\frac{1}{2}(z^\varepsilon)^2\tilde\eta^\varepsilon_I dt + z^\varepsilon r^2\tilde\eta^\varepsilon_I dt - \frac{1}{2}r^4\tilde\eta^\varepsilon_I dt\right). \end{aligned} \tag{4.34}$$

Note that the terms of order 0 in ε correspond to the terms appearing in Equation (4.6).

Proof. Using the fact that u^ε satisfies Eq.(4.11) and ξ^ε satisfies Eq.(4.31), Itô formula gives, putting $u^\varepsilon = e^{im\theta}v^\varepsilon$,

$$\begin{aligned} d(e^{i\xi^\varepsilon(t)}v^\varepsilon(t)) = e^{i\xi^\varepsilon(t)}\Big(& i\frac{d^2}{dr^2}v^\varepsilon + i\frac{1}{r}\frac{d}{dr}v^\varepsilon - i\frac{m^2}{r^2}v^\varepsilon - ir^2v^\varepsilon - \frac{\varepsilon^2}{2}r^4v^\varepsilon \\ &-i|v^\varepsilon|^{2\sigma}v^\varepsilon + i\mu_0 v^\varepsilon + i\varepsilon y^\varepsilon v^\varepsilon + \varepsilon^2 z^\varepsilon r^2 v^\varepsilon - \frac{\varepsilon^2}{2}(z^\varepsilon)^2 v^\varepsilon\Big)dt \\ &+ie^{i\xi^\varepsilon(t)}(\varepsilon z^\varepsilon v^\varepsilon - \varepsilon r^2 v^\varepsilon)dW. \end{aligned} \tag{4.35}$$

Also, we write for $\sigma \geq 1/2$,

$$|\psi_{\mu_0,m} + \varepsilon\tilde\eta^\varepsilon|^{2\sigma}(\psi_{\mu_0,m} + \varepsilon\tilde\eta^\varepsilon) = \psi^{2\sigma+1}_{\mu_0,m} + \varepsilon(2\sigma+1)\tilde\eta^\varepsilon_R\psi^{2\sigma}_{\mu_0,m} + i\varepsilon\tilde\eta^\varepsilon_I\psi^{2\sigma}_{\mu_0,m} + \varepsilon^2 h^\varepsilon_R + i\varepsilon^2 h^\varepsilon_I,$$

where

$$\varepsilon^2 h^\varepsilon_R + i\varepsilon^2 h^\varepsilon_I = \int_0^1 (1-s)\frac{\partial^2}{\partial s^2}\left(|\psi_{\mu_0,m} + s\varepsilon\tilde\eta^\varepsilon|^{2\sigma}(\psi_{\mu_0,m} + s\varepsilon\tilde\eta^\varepsilon)\right)ds.$$

Using these facts, (4.17), replacing $e^{i\xi^\varepsilon(t)}v^\varepsilon(t)$ by $\psi_{\mu_0,m} + \varepsilon\tilde\eta^\varepsilon(t,r)$ in (4.35), and identifying the real and imaginary parts, we deduce the equations (4.33) and (4.34). □

Remark that if we write h^ε_R and h^ε_I as functions of $\psi_{\mu_0,m}$ and $\varepsilon\tilde\eta^\varepsilon$, we obtain what

follows:

$$
\begin{aligned}
h_R^\varepsilon &= 2\sigma \int_0^1 (1-s)[(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)^2+(s\varepsilon\tilde\eta_I^\varepsilon)^2]^{\sigma-1} \\
&\quad \times\Big\{((\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)\tilde\eta_R^\varepsilon+s\varepsilon(\tilde\eta_I^\varepsilon)^2)\tilde\eta_R^\varepsilon+((\tilde\eta_R^\varepsilon)^2+(\tilde\eta_I^\varepsilon)^2)(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)\Big\}ds \\
&\quad +4\sigma(\sigma-1)\int_0^1 (1-s)[(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)^2+(s\varepsilon\tilde\eta_I^\varepsilon)^2]^{\sigma-2} \\
&\quad \times\{(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)\tilde\eta_R^\varepsilon+s\varepsilon(\tilde\eta_I^\varepsilon)^2\}^2(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)ds, \\
h_I^\varepsilon &= 2\sigma \int_0^1 (1-s)[(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)^2+(s\varepsilon\tilde\eta_I^\varepsilon)^2]^{\sigma-1} \\
&\quad \times\Big\{2((\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)\tilde\eta_R^\varepsilon+s\varepsilon(\tilde\eta_I^\varepsilon)^2)\tilde\eta_I^\varepsilon+((\tilde\eta_R^\varepsilon)^2+(\tilde\eta_I^\varepsilon)^2)s\varepsilon\tilde\eta_I^\varepsilon\Big\}ds \\
&\quad +4\sigma(\sigma-1)\int_0^1 (1-s)[(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)^2+(s\varepsilon\tilde\eta_I^\varepsilon)^2]^{\sigma-2} \\
&\quad \times\{(\psi_{\mu_0,m}+s\varepsilon\tilde\eta_R^\varepsilon)\tilde\eta_R^\varepsilon+s\varepsilon(\tilde\eta_I^\varepsilon)^2\}^2 s\varepsilon\tilde\eta_I^\varepsilon ds.
\end{aligned}
$$

Taking the radial L^2 inner product $(\cdot,\cdot)_{L^2_r}$ of Eq.(4.34) with $\psi_{\mu_0,m}(r)$ and making use of the orthogonality condition (4.32), we obtain the equations for the modulation parameters y^ε, z^ε from the identification of drift parts and that of martingale parts.

Lemma 4.10. *Under the assumptions of Theorem 4.5, the modulation parameters satisfy the following system of equations, for any $t \le \tau^\varepsilon$,*

$$
z^\varepsilon(t)\Big\{|\psi_{\mu_0,m}|^2_{L^2_r}+\varepsilon(\tilde\eta_R^\varepsilon,\psi_{\mu_0,m})_{L^2_r}\Big\} = |r\psi_{\mu_0,m}|^2_{L^2_r}+\varepsilon(\tilde\eta_R^\varepsilon,r^2\psi_{\mu_0,m})_{L^2_r}, \tag{4.36}
$$

and

$$
\begin{aligned}
y^\varepsilon(t)\Big\{|\psi_{\mu_0,m}|^2_{L^2_r}+\varepsilon(\tilde\eta_R^\varepsilon,\psi_{\mu_0,m})_{L^2_r}\Big\} &= -\varepsilon^2 z^\varepsilon(t)(\tilde\eta_I^\varepsilon,r^2\psi_{\mu_0,m})_{L^2_r} \\
&\quad +(L^{1,r}_{\mu_0,m}\tilde\eta_R^\varepsilon,\psi_{\mu_0,m})_{L^2_r}+\varepsilon(h_R^\varepsilon,\psi_{\mu_0,m})_{L^2_r} \\
&\quad +\frac{\varepsilon^2}{2}(\tilde\eta_I^\varepsilon,r^4\psi_{\mu_0,m})_{L^2_r}.
\end{aligned} \tag{4.37}
$$

We deduce from the modulation equations obtained in Lemma 4.10 the following estimates for the modulation parameters. Noting that for $t \le \tau^\varepsilon$, choosing α_0 again sufficiently small if necessary, such that $2\alpha_0 \le \sqrt{2\pi}|\psi_{\mu_0,m}|_{L^2_r}$, we have

$$
\Big||\psi_{\mu_0,m}|^2_{L^2_r}+\varepsilon(\tilde\eta_R^\varepsilon,\psi_{\mu_0,m})_{L^2_r}\Big| \ge |\psi_{\mu_0,m}|^2_{L^2_r}-|\varepsilon\tilde\eta_R^\varepsilon|_{L^2_r}|\psi_{\mu_0,m}|_{L^2_r} \ge (1/2)|\psi_{\mu_0,m}|^2_{L^2_r}
$$

for $\alpha \le \alpha_0$. Other estimates for the right hand sides of (4.37)-(4.36) follow exactly as in Corollary 4.3 of [55], using the expression of h_R^ε and h_I^ε above, the Sobolev embedding, and the properties of $\psi_{\mu_0,m}$ given in Proposition 4.4.

Corollary 4.11. *Under the assumptions of Theorem 4.5, there exists $\alpha_1 > 0$ such that for any $\alpha \le \alpha_1$, there is a constant $C_{\mu_0,\alpha,m}$ with*

$$|z^\varepsilon(t)| \le C_{\mu_0,\alpha,m}, \quad \text{a.s. for all} \quad t \le \tau^\varepsilon, \quad \varepsilon \le \varepsilon_0. \tag{4.38}$$

Moreover, there are constants C_1 and C_2 depending only on σ, α, m and μ_0 such that

$$|y^\varepsilon(t)| \le C_1 |\tilde\eta^\varepsilon(t)|_{L^2_r} + \varepsilon C_2 \quad \text{a.s. for all} \quad t \le \tau^\varepsilon, \quad \varepsilon \le \varepsilon_0. \tag{4.39}$$

4.5 Estimates on the remainder term and convergence

Let $\tilde\eta(t,r) = \tilde\eta_R + i\tilde\eta_I$ with $\tilde\eta_R = \operatorname{Re}\tilde\eta(t,r)$ and $\tilde\eta_I = \operatorname{Im}\tilde\eta(t,r)$. Consider the equation (4.25) for $\tilde\eta$, with $y(t)$, $z(t)$ defined by (4.26). It is not difficult to see that Eq. (4.25) with $\tilde\eta(0) = 0$ has a unique adapted solution $\tilde\eta \in C(\mathbb{R}^+, X_m)$, a.s., and $\tilde\eta$ satisfies $(\tilde\eta_I, \psi_{\mu_0,m})_{L^2_r} = 0$.

In this section we will explain that $\tilde\eta^\varepsilon$ converges to $\tilde\eta$ in probability in $C([0,\tau^\varepsilon \wedge T], L^2_r)$ for any $T > 0$ as ε goes to 0, which means η^ε converges to η in probability in $C([0,\tau^\varepsilon \wedge T], L^2(\mathbb{R}^2))$, setting $\eta(t,x) = e^{im\theta}\tilde\eta(t,r)$. We will list up some estimates to prove this convergence, which are similar to those in [55]. We thus omit the proofs of almost all these estimates except the part where we use Proposition 4.7. To clarify the notation, we remark that for $f \in X_m$ with $f(x) = e^{im\theta}g(r)$,

$$|f|^2_\Sigma = (2\pi)|g|^2_{\Sigma_r} = (2\pi)\left(|g|^2_{L^2_r} + |rg|^2_{L^2_r} + \left|\frac{dg}{dr}\right|^2_{L^2_r} + m^2\left|\frac{g}{r}\right|^2_{L^2_r}\right) \tag{4.40}$$

and the property (i) of Proposition 4.4 is useful to estimate the Σ_r-norm since

$$\langle S''_{\mu_0,m}(\phi_{\mu_0,m})f, f\rangle = 2\pi\{\langle L^{1,r}_{\mu_0,m}(\operatorname{Re} g), (\operatorname{Re} g)\rangle + \langle L^{2,r}_{\mu_0,m}(\operatorname{Im} g), (\operatorname{Im} g)\rangle\}.$$

We note that $\tilde\eta^\varepsilon$, y^ε, z^ε are a priori defined only for $t \le \tau^\varepsilon$. We define them for $t \in \mathbb{R}^+$ by simply setting $\tilde\eta^\varepsilon(t) = \tilde\eta^\varepsilon(\tau^\varepsilon)$ for $t \ge \tau^\varepsilon$ and the same for the others.

Lemma 4.12. *Let $T > 0$ be fixed and $1/2 \le \sigma$. Let $\mu_0 > \lambda_m$ be fixed and $\eta^\varepsilon, \xi^\varepsilon$ $(\varepsilon > 0)$ be given by Theorem 4.5 with $\alpha \le \alpha_0$ fixed. Put $\eta^\varepsilon(t,x) = e^{im\theta}\tilde\eta^\varepsilon(t,r)$. There exist constants C_1 and C_2 depending only on T, α, μ_0, m, N such that*

$$\text{(i)}\quad \mathbb{E}\left(\sup_{t\le\bar\tau^\varepsilon_N\wedge T}|\tilde\eta^\varepsilon(t)|^2_{L^2_r}\right) \le C_1, \quad \text{and} \quad \text{(ii)}\quad \mathbb{E}\left(\sup_{t\le\bar\tau^\varepsilon_N\wedge T}|\tilde\eta^\varepsilon(t)|^4_{L^2_r}\right) \le C_2,$$

where

$$\bar\tau^\varepsilon_N = \inf\{t \le \tau^\varepsilon \wedge T,\ |\varepsilon\eta^\varepsilon|_{\Sigma^2} \ge N\}, \quad \text{for any } N > 0.$$

In fact, in order to prove (i) of Lemma 4.12, we need a bound $|\varepsilon\tilde\eta^\varepsilon(t)|_{L^\infty} \le C'$ for some constant $C' > 0$ for any $t \le \bar\tau^\varepsilon_N$. This is realized by the above definition of the stopping

time $\bar{\tau}_N^\varepsilon$, because $|\varepsilon\tilde{\eta}^\varepsilon(t,r)|_{L^\infty} \le C|\varepsilon\eta^\varepsilon(t)|_{\Sigma^2} \le CN$ for any $t \le \bar{\tau}_N^\varepsilon$.

Lemma 4.13. *Let $T > 0$ be fixed and $1 \le \sigma$. Let $\mu_0 > \lambda_m$ be fixed and $\eta^\varepsilon, \xi^\varepsilon$ $(\varepsilon > 0)$ be given by Theorem 4.5 with $\alpha \le \alpha_0$ fixed. Put $\eta^\varepsilon(t,x) = e^{im\theta}\tilde{\eta}^\varepsilon(t,r)$. There exists a constant C_3 depending only on T, α, μ_0, m, N such that*

$$\mathbb{E}\left(\sup_{t\le\bar{\tau}_N^\varepsilon\wedge T}|\tilde{\eta}^\varepsilon(t)|_{\Sigma_r}^4\right) \le C_3,$$

where $\bar{\tau}_N^\varepsilon$ is defined in Lemma 4.12.

As for the solution of Eq.(4.25), we have the following estimate.

Lemma 4.14. *Let $T > 0$ be fixed and $1/2 \le \sigma$. Let $\mu_0 > \lambda_m$ be fixed and $\eta^\varepsilon, \xi^\varepsilon$ $(\varepsilon > 0)$ be given by Theorem 4.5 with $\alpha \le \alpha_0$ fixed. Put $\eta(t,x) = e^{im\theta}\tilde{\eta}(t,r)$. There exist C_4 and C_5 depending only on T, α, μ_0, m such that*

$$\text{(i)}\quad \mathbb{E}\left(\sup_{t\le T}|\tilde{\eta}(t)|_{\Sigma_r}^2\right) \le C_4, \quad \text{(ii)}\quad \mathbb{E}\left(\sup_{t\le T}|(1+r^4)\tilde{\eta}(t)|_{L_r^2}^2\right) \le C_5.$$

We remark that the assumption $\sigma \ge 1$ is needed only for Lemma 4.13. The reason is the same as in [55], due to the terms involving ∇h_R^ε and ∇h_I^ε. Using these lemmas we obtain the following convergence.

Lemma 4.15. *Let $T > 0$ be fixed. Under the assumptions of Theorem 4.6, $\tilde{\eta}^\varepsilon$ converges to $\tilde{\eta}$ as ε tends to 0 in $L^2(\Omega; C([0, \bar{\tau}_N^\varepsilon \wedge T], L_r^2))$.*

The convergence in probability in the time interval $[0, \tau^\varepsilon \wedge T]$ follows from Lemma 4.15:

Corollary 4.16. *Let $T > 0$ be fixed. Under the assumptions of Theorem 4.6, $\tilde{\eta}^\varepsilon$ converges to $\tilde{\eta}$ as ε tends to 0 in probability, in $C([0, \tau^\varepsilon \wedge T], L_r^2)$.*

We give here, assuming Lemma 4.15, a different proof from [55] for this corollary, which improves the range of admissible exponents σ.

Proof of Corollary 4.16. We prove that for any $\beta > 0$, $\delta > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]}|\mathbb{1}_{[0,\tau^\varepsilon\wedge T]}\tilde{\eta}^\varepsilon - \mathbb{1}_{[0,T]}\tilde{\eta}|_{L_r^2} > \delta\right) \le \beta, \tag{4.41}$$

provided that ε is sufficiently small. We note that

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,T]}|\mathbb{1}_{[0,\tau^\varepsilon\wedge T]}\tilde{\eta}^\varepsilon - \mathbb{1}_{[0,T]}\tilde{\eta}|_{L_r^2} > \delta\right) &\le \mathbb{P}\left(\sup_{t\in[0,T]}|\mathbb{1}_{[0,\tau^\varepsilon\wedge T]}(\tilde{\eta}^\varepsilon - \tilde{\eta})|_{L_r^2} > \delta\right)\\ &+ \mathbb{P}(\tau^\varepsilon \wedge T < T).\end{aligned}$$

It follows from (4.24) that for any $\beta > 0$ there exists $\varepsilon_0 > 0$ such that $\mathbb{P}(\tau^\varepsilon \wedge T < T) \leq \beta/3$ for any $\varepsilon \leq \varepsilon_0$. On the other hand,

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,T]} |\mathbb{1}_{[0,\tau^\varepsilon\wedge T]}(\tilde\eta^\varepsilon - \tilde\eta)|_{L^2_r} > \delta\right) &\leq \mathbb{P}\left(\sup_{t\in[0,T]} |\mathbb{1}_{[0,\bar\tau^\varepsilon_N\wedge T]}(\tilde\eta^\varepsilon - \tilde\eta)|_{L^2_r} > \delta\right)\\ &+ \mathbb{P}(\bar\tau^\varepsilon_N \wedge T < T).\end{aligned}$$

Concerning the second term, we first show that for any $\beta > 0$ there exist N_0 and ε_0 such that for any $\varepsilon \leq \varepsilon_0$ and $N \geq N_0$,

$$\mathbb{P}\left(\sup_{t\in[0,\tau^\varepsilon\wedge T]} |\varepsilon\eta^\varepsilon|_{\Sigma^2} \geq N\right) \leq \beta/6,$$

and then use the fact that

$$\mathbb{P}(\bar\tau^\varepsilon_N \wedge T < T) \leq \mathbb{P}\left(\sup_{t\in[0,\tau^\varepsilon\wedge T]} |\varepsilon\eta^\varepsilon|_{\Sigma^2} \geq N\right) + \mathbb{P}(\tau^\varepsilon \wedge T < T).$$

Remarking $\varepsilon\eta^\varepsilon(t,r) = e^{i\xi^\varepsilon(t)}u^\varepsilon(t,r) - \phi_{\mu_0,m}$, it suffices to show that

$$\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\in[0,\tau^\varepsilon\wedge T]} |u^\varepsilon(t)|_{\Sigma^2} \geq N\right) = 0, \tag{4.42}$$

uniformly for $\varepsilon \leq 1$. Let $\varepsilon \leq 1$ and $w^\varepsilon(t)$ be the solution of (4.9) with $w^\varepsilon(0) = \phi_{\mu_0,m} \in \Sigma^2$; let moreover $\hat\tau^\varepsilon_M := \inf\{t \geq 0, |w^\varepsilon(t)|_{\Sigma^2} \geq M\}$ for $M > 0$. By Proposition 4.7 (and Remark 4.19 in the Appendix),

$$\lim_{M\to\infty} \hat\tau^\varepsilon_M = +\infty, \quad a.s, \text{ uniformly in } \varepsilon \leq 1,$$

that is, for any finite time $T \in (0, +\infty)$,

$$0 = \lim_{M\to\infty} \mathbb{P}(\hat\tau^\varepsilon_M < T) = \lim_{M\to\infty} \mathbb{P}(\sup_{t\in[0,T]} |w^\varepsilon(t)|_{\Sigma^2} \geq M),$$

uniformly for $\varepsilon \leq 1$. The gauge transformation (4.8) keeps the equivalence of the Σ^2 norm between $u^\varepsilon(t)$ and associated solution $w^\varepsilon(t)$ uniformly for $\varepsilon \leq 1$, thus (4.42) follows.

On the other hand, using Lemma 4.15, we have for any $\beta > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} |\mathbb{1}_{[0,\bar\tau^\varepsilon_{N_0}\wedge T]}(\tilde\eta^\varepsilon - \tilde\eta)|_{L^2_r} > \delta\right) \leq \frac{\beta}{6},$$

for $\varepsilon \leq \varepsilon_0$ sufficiently small, and we deduce that (4.41) holds for $\varepsilon \leq \varepsilon_0$. □

4.6 Numerical observations

The main goal of this section is to confirm numerically the result of Theorem 4.5 and to give numerical evidence that, to some extent, the estimate of Theorem 4.5 is optimal. We also compute the numerical evolution of vortices in the presence of noise. We first present the numerical scheme we use to solve equation (4.4) under the assumption of m-equivariance symmetry. We also briefly explain how we can compute the radial profile $\psi_{\mu,m}$ given by equation (4.17). Then we use a classical Monte Carlo method to estimate the left-hand side of equation (4.24). Eventually, we use two Monte Carlo methods dedicated to computing this quantity when it becomes too small to be evaluated with a direct method. They are called Interacting Particle System (IPS) algorithm and Parallel One-Path (POP) algorithm (see [87]).

4.6.1 Numerical integration

We use a relaxed Crank-Nicolson scheme. We limit the resolution domain to $[0, R]$, where R needs to be chosen large enough to avoid reflections. We set a spatial discretization step $\Delta r = \dfrac{R}{N_r}$ where $N_r \in \mathbb{N}$. We set $r_j = j\Delta r$ for $0 \leq j \leq N_r$, and $r_{j+1/2} = (j+1/2)\Delta r$ for $0 \leq j \leq N_r$. We fix a final time $T > 0$ and a time discretization step $\Delta t = \dfrac{T}{N_t}$ where $N_t \in \mathbb{N}$.

Because of the m-equivariance symmetry we solve equation (4.4) for $f(r,t)$ where $u(x,t) = e^{-im\theta} r^m f(r,t)$ (with the notation of the previous sections). This decomposition is motivated by Proposition 4.4. The numerical scheme, adapted from a scheme used in [12] is the following:

$$\begin{cases} \dfrac{\varphi_{j+\frac{1}{2}}^{n+\frac{1}{2}} + \varphi_{j+\frac{1}{2}}^{n-\frac{1}{2}}}{2} = |f_{j+\frac{1}{2}}^n|^{2\sigma}, \\ r_{j+\frac{1}{2}}^m (f_{j+\frac{1}{2}}^{n+1} - f_{j+\frac{1}{2}}^n) \\ \qquad -\mathrm{i}\left[\dfrac{1}{r_{j+\frac{1}{2}}\Delta r^2}\left(r_{j+1} r_{j+\frac{3}{2}}^m f_{j+\frac{3}{2}}^{n+\frac{1}{2}} - 2 r_{j+\frac{1}{2}}^{m+1} f_{j+\frac{1}{2}}^{n+\frac{1}{2}} + r_j r_{j-\frac{1}{2}}^m f_{j-\frac{1}{2}}^{n+\frac{1}{2}}\right) - m^2 r_{j+\frac{1}{2}}^{m-2} f_{j+\frac{1}{2}}^{n+\frac{1}{2}}\right]\Delta t \\ \qquad +\mathrm{i}\, r_{j+\frac{1}{2}}^{m+2} f_{j+\frac{1}{2}}^{n+\frac{1}{2}} \Delta t + \mathrm{i}\varepsilon r_{j+\frac{1}{2}}^{m+2} f_{j+\frac{1}{2}}^{n+\frac{1}{2}} \Delta W^n + \mathrm{i} r_{j+\frac{1}{2}}^{3m} \varphi^{n+\frac{1}{2}} f_{j+\frac{1}{2}}^{n+\frac{1}{2}} \Delta t = 0, \\ f_{-\frac{1}{2}}^n = f_{\frac{1}{2}}^n, \\ f_{N_r+\frac{1}{2}}^n = 0 \quad \text{or} \quad f_{N_r+\frac{1}{2}}^n = f_{N_r-\frac{1}{2}}^n, \\ \varphi^{-\frac{1}{2}} = |f^0|^{2\sigma}, \end{cases} \tag{4.43}$$

for all $0 \leq n \leq N_t$, where $f_{j+\frac{1}{2}}^n$ is an approximation of $f((j+\frac{1}{2})\Delta r, n\Delta t)$, $f^{n+\frac{1}{2}} = \frac{f^{n+1}+f^n}{2}$, $\Delta W^n = \sqrt{\delta t}\mathcal{N}^n$ with $(\mathcal{N}^n)_{0\leq n\leq N_t}$ a family of independent standard Gaussian random

variables.

The boundary condition at $r = R$ can be either homogeneous Dirichlet condition or homogeneous Neumann condition. Numerical experiments show that the impact is not significant. In this scheme, $\varphi^{n+\frac{1}{2}}$ is an approximation of $|f^{n+\frac{1}{2}}|^{2\sigma}$. This scheme does not conserve exactly the L^2 norm, but with a sufficiently good spatial discretization, the fluctuations are negligible.

In order to compute the radial profile of the vortex $\psi_{\mu,m}$, which is the initial state in our numerical simulations, we use a shooting method inspired by [65]. In our case, setting $\psi_{\mu,m}(r) = r^m f_{\mu,m}(r)$, we look for a $C^2(0,\infty) \cap C([0,\infty))$ real non-negative function $f_{\mu,m}$ satisfying the following equation

$$-f''_{\mu,m}(r) - \frac{2m+1}{r} f'_{\mu,m}(r) - \mu f_{\mu,m}(r) + r^{2\sigma m} f_{\mu,m}(r)^{2\sigma+1} + r^2 f_{\mu,m}(r) = 0, \quad r > 0 \tag{4.44}$$

with

$$\lim_{r\to\infty} |f_{\mu,m}(r)| = 0.$$

This equation can be written as a first degree ODE. We set $X(r) = (f(r), f'(r))$, and equation (4.44) becomes:

$$\begin{cases} X'(r) = F(r, X(r)), & r > 0 \\ X(0) = (\alpha, \beta), \end{cases} \tag{4.45}$$

with

$$F(r,(x,y)) := \left(y, -\frac{2m+1}{r} y - \mu x + r^{2\sigma m} x^{2\sigma+1} + r^2 x \right), \quad \forall r > 0, \quad \forall x, y \in \mathbb{R}. \tag{4.46}$$

In order to solve this equation numerically, we limit the domain to $[0, R]$ and we use a Runge-Kutta scheme of order four. Since the nonlinear term $f_{\mu,m}^{2\sigma}(r)$ vanishes at $r = 0$, it is easily seen by the argument of [99], Section 2 that we must set $\beta = 0$. The idea is to find α thanks to a dichotomy method, looking at the behavior of $f_{\mu,m}$ for large r. We thus search a value of α that enables $f_{\mu,m}$ to be positive on $(0, R)$ and that minimizes $f_{\mu,m}(r)$ for large r.

We show in Figures 4.1, 4.2 and 4.3 some profiles $\psi_{\mu,m}$, computed from equation (4.45) for different m and μ.

We show in Figures 4.4, 4.5 and 4.6 some trajectories (computed from the same realization of a Brownian motion) of $|u^\varepsilon(r,t)|$ for different values of m and $\mu = 2m + 3$. Here, t is represented along the ordinate axis, and r along the abscissa.

These figures have been plotted with ε of order 10^{-1}. For this order of magnitude, we can observe that the wave function keeps the same structure over time and that it oscillates with an almost periodic rhythm. The simulations that will be presented in the sequel of the chapter have been computed for ε of order 10^{-2}. In this context, the oscillations of

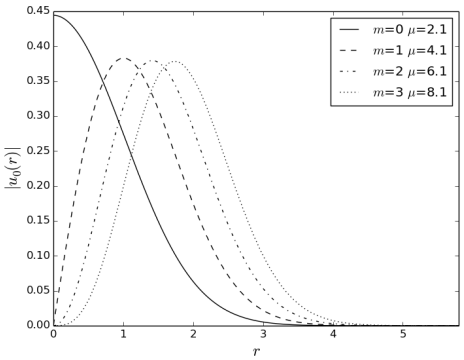

Figure 4.1 – Amplitudes of vortices $|\psi_{\mu,m}|$ for μ close to λ_m, for different values of m

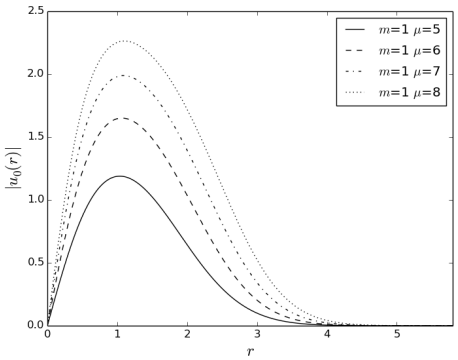

Figure 4.2 – Amplitudes of vortices $|\psi_{\mu,m}|$ for $m = 1$ for different values of μ

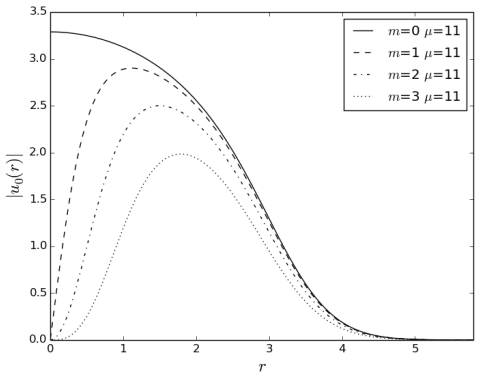

Figure 4.3 – Amplitudes of vortices $|\psi_{\mu,m}|$ for $\mu = 11$ for different values of m

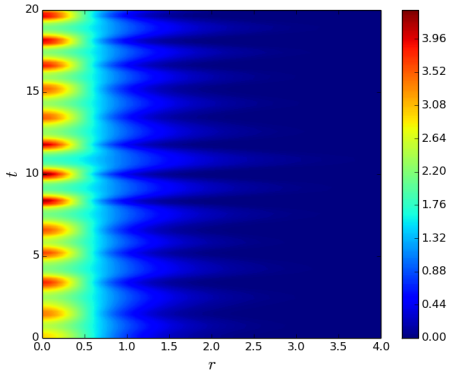

Figure 4.4 – A trajectory of $|u^\varepsilon(r,t)|$ for $m = 0$ and $\sigma = 1$

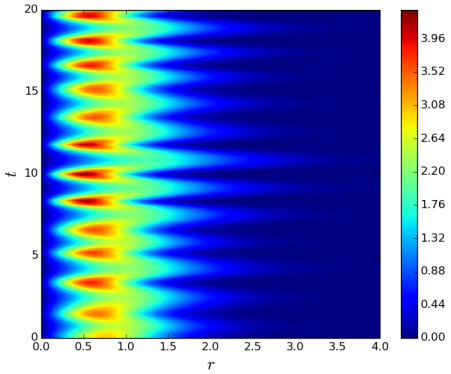

Figure 4.5 – A trajectory of $|u^\varepsilon(r,t)|$ for $m = 1$ and $\sigma = 1$

$|u^\varepsilon|$ are of course much smaller.

The last step is to compute $\xi^\varepsilon(t)$ and $\varepsilon\eta^\varepsilon(x,t)$ from the simulated trajectories. The

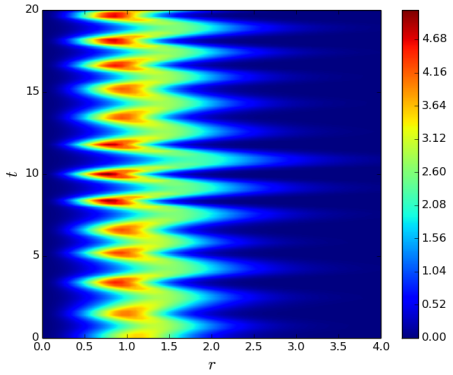

Figure 4.6 – A trajectory of $|u^\varepsilon(r,t)|$ for $m=2$ and $\sigma=1$

computation of $\xi^\varepsilon(t)$ for $t \leq \tau^\varepsilon$ is given by

$$\xi^\varepsilon(t) = -\arg\left(\int_{\mathbb{R}^2} u^\varepsilon(t)\overline{\phi_{\mu_0,m}}\right),$$

thanks to the orthogonality condition (4.22). Then, the computation of $\varepsilon\eta^\varepsilon$ is immediately given by $\varepsilon\eta^\varepsilon(t,x) = u^\varepsilon(t,x)e^{i\xi^\varepsilon(t)} - \phi_{\mu_0,m}(x)$. We represent in Figure 4.7 a trajectory of $\xi^\varepsilon(t) - \mu_0 t$ and $|\varepsilon\eta^\varepsilon(t)|_\Sigma$.

4.6.2 A naive Monte Carlo method

We now estimate the left-hand side of inequality (4.24). We use a classical Monte Carlo method. We set $\sigma \geq 1/2$, $m \in \mathbb{N}$, $\mu_0 > \lambda_m$, and $\varepsilon > 0$. According to Theorem 4.5, we choose $\alpha > 0$ and assume $\alpha \leq \alpha_0$. Taking $T > 0$, $N \in \mathbb{N}$, $\Delta t > 0$ and $\Delta r > 0$, we estimate $\mathbb{P}(\tau_\alpha^\varepsilon \leq T)$ by the following estimator :

$$\widehat{Y}^{\alpha,\varepsilon}_{\Delta r,\Delta t,N} = \frac{1}{N}\sum_{k=1}^{N} Y^{\alpha,\varepsilon,(k)}_{\Delta r,\Delta t} \tag{4.47}$$

with

$$Y^{\alpha,\varepsilon,(k)}_{\Delta r,\Delta t} = \mathbf{1}_{\left\{|(\varepsilon\eta^\varepsilon_{\Delta r,\Delta t})^{(k)}|_{L^\infty((0,T);\Sigma)} > \alpha\right\}} \tag{4.48}$$

where $(\varepsilon\eta^\varepsilon_{\Delta r,\Delta t})^{(k)}$ for $1 \leq k \leq N$ are remainders in the decomposition of independent solutions of the numerical scheme with the discretization steps Δr and Δt. Thus the $Y^{\alpha,\varepsilon,(k)}_{\Delta r,\Delta t}$ for $1 \leq k \leq N$ are iid. This estimator is meaningful if

$$\lim_{\Delta r,\Delta t,N^{-1}\to 0} \mathbb{E}\left(\widehat{Y}^{\alpha,\varepsilon}_{\Delta r,\Delta t,N} - \mathbb{P}(\tau_\alpha^\varepsilon \leq T)\right)^2 = 0.$$

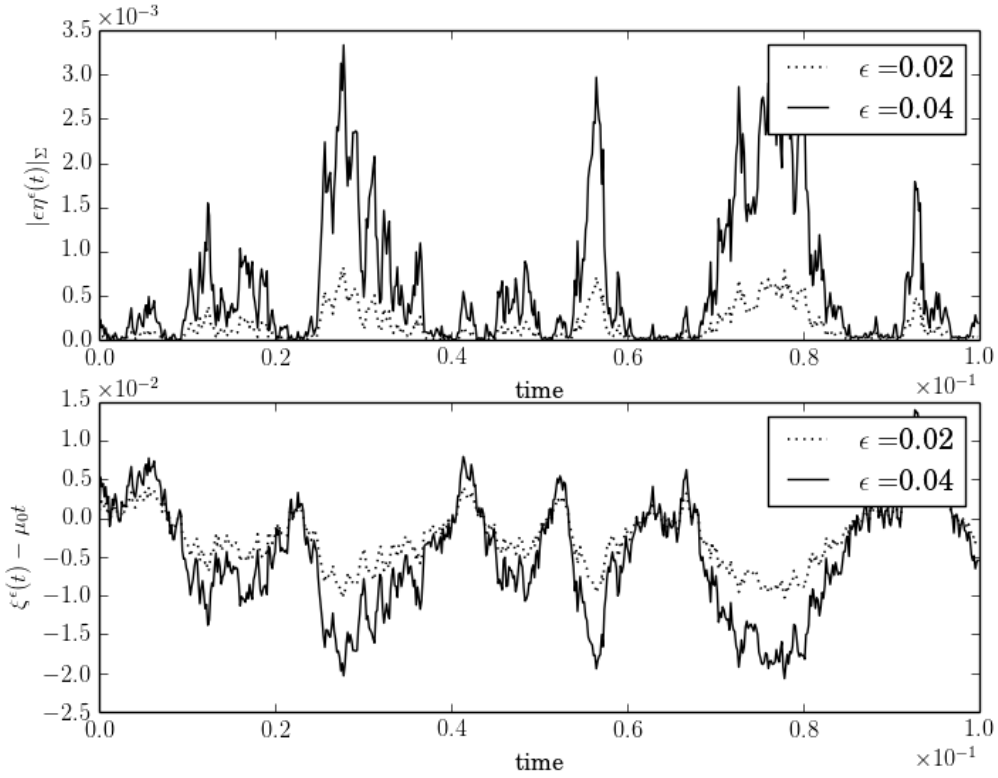

Figure 4.7 – An evolution of $\xi^\varepsilon(t) - \mu_0 t$ and $|\varepsilon\eta^\varepsilon(t)|_\Sigma$, for $m = 2$ and $\mu_0 = 7$

This mean-square error can be written:

$$\mathbb{E}\Big(\widehat{Y}^{\alpha,\varepsilon}_{\Delta r,\Delta t,N} - \mathbb{P}(\tau^\varepsilon_\alpha \leq T)\Big)^2 = \mathrm{Var}\left[\widehat{Y}^{\alpha,\varepsilon}_{\Delta r,\Delta t,N}\right] + [\mathbb{E}Y^{\alpha,\varepsilon,(1)}_{\Delta r,\Delta t} - \mathbb{P}(\tau^\varepsilon_\alpha \leq T)]^2.$$

The first term of the right-hand side converges to zero as N tends to infinity at speed N^{-1}. The convergence to zero of the second term of the right-hand side is an expected property of the numerical scheme. This latter property of the scheme is not theoretically proved, but we will assume that taking a discretisation small enough enables this term to be as small as we wish.

We ran the simulations taking $\alpha = 2.5 \cdot 10^{-3}$ and $\sigma = 1$. The results are given in Figure 4.8 and 4.9. We can observe in Figure 4.8 that for ε small enough the curve becomes a straight line of slope -2, which agrees with the upper bound given by (4.24). For larger ε, the curve goes below the straight line given by the lowest ε. Thus for larger ε, estimate (4.24) is not valid anymore. This shows the existence of the ε_0 given by Theorem 4.5.

In Figure 4.9 we fix $\varepsilon = 9 \cdot 10^{-3}$ and we estimate $\mathbb{P}(\tau^\varepsilon_\alpha \leq t)$ for various time $0 \leq t \leq T = 0.1$. Figure 4.8 ensures that $\varepsilon \leq \varepsilon_0$ for this choice of α and T. We observe that the curve is a straight line of slope -1, which means that the upper bound given by Equation (4.24)

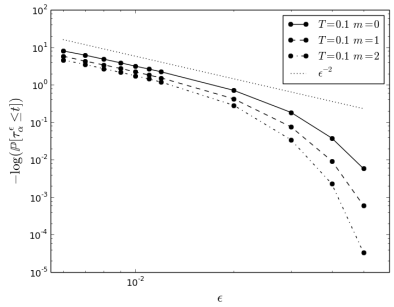

Figure 4.8 – Estimation of $\mathbb{P}(\tau_\alpha^\varepsilon \leq t)$ with respect to ε, by a Monte Carlo method

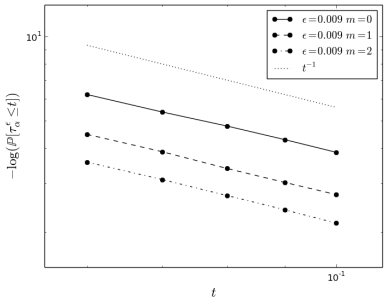

Figure 4.9 – Estimation of $\mathbb{P}(\tau_\alpha^\varepsilon \leq t)$ with respect to t, by a Monte Carlo method

is sharp.

4.6.3 Rare event estimation

The classical Monte Carlo approach we use in Section 4.6.2 here becomes unreasonable to evaluate $\mathbb{P}(\tau_\alpha^\varepsilon \leq T)$ when this probability becomes very small. To overcome this problem, we use a Monte Carlo method suited for the estimation of this kind of rare events. More precisely we relied on the Interacting Particule System (IPS) algorithm and the Parallel One-Path (POP) algorithm constructed with *shaking transformations* as presented in [87].

Short presentation of IPS and POP

We only briefly recall these two algorithms, in our specific context. For a general presentation, we refer to [87]. We recall that we aim at computing the probability of the rare event $\{\tau_\alpha^\varepsilon \leq T\}$. Because of the numerical approximations, we estimate in practice the probability of the event A defined by,

$$A_\alpha^\varepsilon = \left\{\tau_{\alpha,\Delta r,\Delta t}^\varepsilon \leq T\right\}, \text{ where } \tau_{\alpha,\Delta r,\Delta t}^\varepsilon = \inf\left\{t \geq 0; \left\|\varepsilon\eta_{\Delta r,\Delta t}^\varepsilon(t)\right\|_\Sigma \geq \alpha\right\},$$

(instead of $\{\tau_\alpha^\varepsilon \leq T\}$) by a Monte Carlo method, where $\varepsilon\eta_{\Delta r,\Delta t}^\varepsilon$ is defined by,

$$\varepsilon\eta_{\Delta r,\Delta t}^\varepsilon = e^{i\xi}\phi_{\mu_0,m} - u_{\Delta r,\Delta t}^\varepsilon, \text{ with } \xi = \underset{\xi'\in[0,2\pi)}{\operatorname{argmin}} \left\|e^{i\xi'}\phi_{\mu_0,m} - u_{\Delta r,\Delta t}^\varepsilon\right\|_{L^2},$$

where $u_{\Delta r,\Delta t}^\varepsilon$ the solution of the numerical scheme (4.43). In this section, we put aside the question of the bias that arises from the numerical approximation, and we focus on the Monte Carlo methods to estimate the probability of A_α^ε. The approach of the POP and IPS algorithms relies on a splitting approach that involves nested events. We define now a decreasing sequence (for the inclusion) of events $(A_k)_{0\leq k\leq n}$ such that,

$$A_\alpha^\varepsilon =: A_n \subset \ldots \subset A_k \subset \ldots \subset A_0 =: \Omega.$$

This way, the probability $\mathbb{P}(A_n)$ is given by,

$$\mathbb{P}(A_n) = \prod_{k=1}^{n} \mathbb{P}(A_k|A_{k-1}).$$

Both the IPS and the POP algorithms estimate the left hand-side as a product of estimators of the conditional probabilities,

$$\mathbb{P}(A_k|A_{k-1}), \tag{4.49}$$

which are constructed as empirical averages of M replicas for IPS, and as ergodic averages of N iterations of a Markov chain for POP.

To implement these two methods, we need to be able to define reversible Markov chains on the state space of paths (which is in our case $C([0,T],X_m)$ if we linearly interpolate the discrete solution). In practice, to readily use the framework developed in [87], we express this rare event A_α^ε in terms of a set of trajectories for the Brownian motion driving the Equation (4.4). Since our numerical scheme relies on a fixed point iteration, the application $\Psi_{\Delta r,\Delta t}$ that associates to a Brownian motion the solution $u_{\Delta r,\Delta t}$ of the numerical scheme (4.43) is measurable, which justifies the method. More precisely, we obtain,

$$\mathbb{P}(\tau_\alpha^\varepsilon(u_{\Delta t,\Delta r}) \leq T) = \mathbb{P}(\tau_\alpha^\varepsilon \circ \Psi_{\Delta r,\Delta t}(W) \leq T) = \mathbb{P}(W \in (\tau_\alpha^\varepsilon \circ \Psi_{\Delta r,\Delta t})^{-1}([0,T])).$$

We introduce $\bar{A}_k$ the sets of Brownian trajectories

$$\bar{A}_k = (\tau_\alpha^\varepsilon \circ \Psi_{\Delta r,\Delta t})^{-1}([0,T]).$$

This way, the conditional probability (4.49) can be expressed in the following way,

$$\mathbb{P}(A_k|A_{k-1}) = \mathbb{P}(W \in \bar{A}_k|W \in \bar{A}_{k-1}). \tag{4.50}$$

This formulation is especially interesting since it enables us to readily apply the POP and the IPS algorithms on the Brownian trajectories, using the shaking transformations on the Brownian motion W defined in Section 3.1 [87].

We present now this shaking transformations on the Brownian motion. We define the measurable mapping K by,

$$K(X,Y) = \rho X + \sqrt{1-\rho^2}Y,\ \forall (X,Y) \in C(\mathbb{R}^+,\mathbb{R})^2,$$

for some $\rho \in (0,1)$. More general kernels K could be used, but this simple form yields goods results in our case. These mappings are used as shaking transformations in the following. This terminology is justified by the fact that if X and Y are two independent Brownian motions, then $K(X,Y)$ is another Brownian motion, supposed to be close to X when ρ is close to 0. The idea is that the mapping K enables to slightly "shake" the

Brownian motion X. We also define the shaking transformations with rejections M_k^K, for $k = 0, \dots, n$, by,

$$M_k^K : C(\mathbb{R}^+, \mathbb{R})^2 \to C(\mathbb{R}^+, \mathbb{R}),\ (X, Y) \mapsto K(X, Y)\mathbb{1}_{K(X,Y)\in \bar{A}_k} + X\mathbb{1}_{K(X,Y)\notin \bar{A}_k}.$$

In practice the mappings M_k^K enable to construct some Markov chains whose marginal laws are those of a Brownian motion conditioned to take values in the sets $\bar{A}_k$.

We present now the IPS algorithm. We set $\gamma \in [0, 1]$, a parameter of this algorithm, and we introduce a set $(U^{(k,m)})_{(k,m)\in[\![0,n-2]\!]\times[\![1,m]\!]}$ of independent random variables uniformly distributed over $[0, 1]$. We suppose in addition that they are independent with every other random variables that appears in the algorithm.

Algorithm 4.17 (The Interacting Particule System (IPS) algorithm)**.**

1. *Initialisation:*
   (a) *Draw* $(W^{(0,m)})_{1\le m\le M}$ *independent Brownian motions,*
   (b) *Set* $p^{(0)} = \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_{\bar{A}_1}(X^{(0,m)})$.
2. *Iterate for* $k = 0$ *until* $n - 2$*:*
   (a) *Set* $I_k = \{m \in \{1, \dots, M\}$ *s.t.* $X^{(k,m)} \in \bar{A}_{k+1}\}$
   (b) *Iterate for* $m = 1$ *until* M
      i. *Selection step: if* $U^{(k,m)} < \gamma$ *and* $X^{(k,m)} \in \bar{A}_{k+1}$ *then set* $\hat{X}^{(k,m)} = X^{(k,m)}$*, otherwise set* $\hat{X}^{(k,m)} = X^{(k,\hat{m})}$ *where* $\hat{m}$ *is drawn uniformly and independently of everything else in the set* I_k.
      ii. *Mutation step: set* $X^{(k+1,m)} = M_{k+1}^K(\hat{X}^{(k,m)}, Y^{(k,m)})$*, where* $Y^{(k,m)}$ *is a Brownian motion independent of everything else.*
   (c) *Set* $p^{(k+1)} = \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_{\bar{A}_{k+2}}(X^{(k+1,m)})$
3. *Return* $p = \prod_{k=0}^{n-1} p^{(k)}$.

As we said previously, the $p^{(k)}$ are the empirical estimators of the conditional probabilities (4.49). The selection with probability γ enables to increase the independent resampling effect (for low γ). For convergence results about this algorithm, we refer to Theorem 2.6 [87].

We present now the POP algorithm. As we stated previously, the idea is to estimate the conditional probabilities (4.49) as ergodic averages of a Markov chain of Brownian trajectories conditionned to be in $\bar{A}_k$. In the following description, we denote by N the number of steps in these ergodic averages, which is equal for every levels.

Algorithm 4.18 (Parallel One-Path (POP) algorithm)**.**

1. *Initialisation: Sample* $X_{0,0}$ *a Brownian motion.*
2. *Iterate for* $k = 0$ *until* $n - 1$*:*
   (a) *For* $i = 1$ *until* $N-1$*, set* $X_{k,i} = M_k^K(X_{k,i-1}, Y_{k,i-1})$*, where* $Y_{k,i-1}$ *is a Brownian motion independent of everything else.*

(b) *Set* $p^{(k)} = \frac{1}{N}\sum_{i=0}^{N-1} \mathbb{1}_{\bar{A}_{k+1}}(X_{k,i})$.

(c) *Set* $i_k = \text{argmin}\{j : X_{k,j} \in \bar{A}_{k+1}\}$.

(d) *Set* $X_{k+1,0} = X_{k,i_k}$.

3. *Return* $p = \prod_{k=0}^{n-1} p^{(k)}$.

In fact, the loops in k can be parallelised as soon as an element of the next level is sampled.

Numerical results

We present in this section some numerical results for the POP and the IPS algorithms. The numerical parameters for the dynamics are given by, $\alpha = 2.5 \cdot 10^{-3}$, $\sigma = 1$, $\varepsilon = 5 \cdot 10^{-3}$ and $T = 5$. The simulation is run with $\delta r = 5.86 \cdot 10^{-2}$ and $\delta t = 0.4\delta r^2$. We chose to introduce $n = 4$ nested events and we defined the events $(\bar{A}_k)_{0\leq k\leq n}$ by,

$$\bar{A}_k = \{W = (W_t)_{t\leq T}) \in C([0,T];\mathbb{R}); \tau^{\varepsilon}_{\alpha_k,\Delta r,\Delta t} \circ \Psi_{\Delta r,\Delta t}(W) \leq T\}, \text{ with } \alpha_k = k\alpha/n.$$

We ran the POP algorithm with $N = 10^4$ and the IPS algorithm with $M = N$, for a fair comparison. We compare these estimators for $\rho \in \{0.9, 0.85, 0.8\}$ (the shaking intensity), and for $\gamma \in \{1, 0.8, 0.6\}$ (the resampling parameter) for IPS. A comparison of the variances is given in Table 4.1 and 4.2 respectively for the IPS and the POP algorithm. We can observe similar orders of magnitude.

Table 4.1 – Estimation of the variance for the IPS algorithm

IPS	mean	std	std/mean
$\rho = 0.9, \gamma = 1$	$1.12 \cdot 10^{-3}$	$8.20 \cdot 10^{-5}$	$7.16 \cdot 10^{-2}$
$\rho = 0.9, \gamma = 0.8$	$1.16 \cdot 10^{-3}$	$9.56 \cdot 10^{-5}$	$8.24 \cdot 10^{-2}$
$\rho = 0.9, \gamma = 0.6$	$1.13 \cdot 10^{-3}$	$10.40 \cdot 10^{-5}$	$9.23 \cdot 10^{-2}$
$\rho = 0.85, \gamma = 1$	$1.12 \cdot 10^{-3}$	$7.48 \cdot 10^{-5}$	$6.69 \cdot 10^{-2}$
$\rho = 0.85, \gamma = 0.8$	$1.16 \cdot 10^{-3}$	$7.02 \cdot 10^{-5}$	$6.03 \cdot 10^{-2}$
$\rho = 0.85, \gamma = 0.6$	$1.11 \cdot 10^{-3}$	$9.27 \cdot 10^{-5}$	$8.33 \cdot 10^{-2}$
$\rho = 0.80, \gamma = 1$	$1.14 \cdot 10^{-3}$	$5.71 \cdot 10^{-5}$	$5.02 \cdot 10^{-2}$
$\rho = 0.80, \gamma = 0.8$	$1.14 \cdot 10^{-3}$	$7.68 \cdot 10^{-5}$	$6.72 \cdot 10^{-2}$
$\rho = 0.80, \gamma = 0.6$	$1.13 \cdot 10^{-3}$	$8.07 \cdot 10^{-5}$	$7.17 \cdot 10^{-2}$

Table 4.2 – Estimation of the variance for the POP algorithm

POP	mean	std	std/mean
$\rho = 0.9$	$1.09 \cdot 10^{-3}$	$9.84 \cdot 10^{-5}$	$9.03 \cdot 10^{-2}$
$\rho = 0.85$	$1.11 \cdot 10^{-3}$	$9.97 \cdot 10^{-5}$	$8.94 \cdot 10^{-2}$
$\rho = 0.80$	$1.14 \cdot 10^{-3}$	$8.26 \cdot 10^{-5}$	$7.23 \cdot 10^{-2}$

We present in Table 4.3 the mean rejection ratio for both IPS and POP algorithms. This mean ratio is supposed to be the same for the two algorithms, and it does not depend

Table 4.3 – Estimation of the mean rejection ratio for POP and IPS algorithms

Mean rejection ratio	level $k=1$	level $k=2$	level $k=3$	level $k=4$
$\rho = 0.9$	0.24	0.39	0.48	0.55
$\rho = 0.85$	0.30	0.47	0.57	0.65
$\rho = 0.8$	0.34	0.53	0.64	0.72

on the choice of γ. We can observe that this ratio increases with k. This can be understood by the fact that as k increases, the sets of trajectories $\bar{A}_k$ become "smaller", and thus the shaking transformation is more likely to push the Brownian trajectory out of these sets. To avoid this behaviour, the shaking intensity ρ could be chosen decreasingly with respect to k. One good practice could be to pre-run the algorithm to tune both this parameter and the probabilities (4.49). Moreover, an adaptive version for the POP method has been proposed in [4].

4.7 Appendix

Proof of Proposition 4.4 (i): Since $v \in X_m$, we write $v(x) = e^{im\theta} f(r)$ with $f(r) = f_R(r) + i f_I(r)$ where $f_R = \operatorname{Re} f$, and $f_I = \operatorname{Im} f$. With this notation, we may describe $S''_\mu(\phi_{\mu,m})$ as follows.

$$\langle S''_\mu(\phi_{\mu,m})v, v\rangle = 2\pi\Big[(L^{1,r}_{\mu,m} f_R, f_R)_{L^2_r} + (L^{2,r}_{\mu,m} f_I, f_I)_{L^2_r}\Big], \tag{4.51}$$

where $L^{1,r}_{\mu,m}$ and $L^{2,r}_{\mu,m}$ are defined by (4.19) and (4.20). First, the self-adjointness of $L^{1,r}_{\mu,m}$ and $L^{2,r}_{\mu,m}$ for $m \geq 1$ with domain D follows from similar arguments as in Appendix X-1 of [149] via the use of the unitary transform

$$\begin{aligned} U : L^2((0,+\infty), r dr) &\to L^2((0,+\infty), dr) \\ \varphi &\mapsto U\varphi = r^{1/2}\varphi, \end{aligned}$$

and the spectrum of both operators is purely discrete since

$$\frac{m^2}{r^2} + r^2 - \mu + |\psi_{\mu,m}(r)|^{2\sigma} \to +\infty, \text{ as } r \to +\infty$$

(see [21, Chapter 2] and [148]). Remark that $\psi_{\mu,m}(r)$ is positive for $r > 0$ and satisfies $L^{2,r}_{\mu,m}\psi_{\mu,m} = 0$, thus $\psi_{\mu,m}$ is the simple eigenfunction corresponding to the eigenvalue 0 (see Chapter 3-3 of [21] for details). This concludes that there exists $\delta > 0$ such that

$$(L^{2,r}_{\mu,m} h, h)_{L^2_r} \geq \delta |h|^2_{L^2_r}$$

for any $h \in \Sigma_r$ satisfying $(h, \psi_{\mu,m})_{L^2_r} = 0$. Here, recall that the norm Σ_r is defined in (4.40). Note that for any $h \in \Sigma_r$,

$$(L^{1,r}_{\mu,m}h, h)_{L^2_r} = (L^{2,r}_{\mu,m}h, h)_{L^2_r} + 2\sigma \int_0^\infty |\psi_{\mu,m}(r)|^{2\sigma}|h(r)|^2 r dr. \tag{4.52}$$

Therefore, if we denote the first eigenvalues of both operators by

$$\mu_1^{(j)} := \inf\{(L^{j,r}_{\mu,m}h, h)_{L^2_r},\ h \in \Sigma_r,\ |h|_{L^2_r} = 1\}, \quad j = 1, 2,$$

it turns out from the relation (4.52) that

$$\mu_1^{(1)} \geq \mu_1^{(2)} + 2\sigma \inf\left\{\int_0^\infty |\psi_{\mu,m}(r)|^{2\sigma}|h(r)|^2 r dr,\ h \in \Sigma_r,\ |h|_{L^2_r} = 1\right\}.$$

Since $\psi_{\mu,m}(r)$ is strictly positive for $r > 0$, we see that $\mu_1^{(1)} > \mu_1^{(2)} = 0$, and we get

$$(L^{1,r}_{\mu,m}h, h)_{L^2_r} \geq \mu_1^{(1)}|h|^2_{L^2_r}$$

for any $h \in \Sigma_r$. Finally going back to (4.51), we may see that there exists $\nu > 0$ such that

$$\langle S''_\mu(\phi_{\mu,m})v, v\rangle \geq 2\pi \int_0^\infty \nu(|f_I(r)|^2 + |f_R(r)|^2) r dr$$

for any $f_I \in \Sigma_r$ satisfying $(f_I, \psi_{\mu,m})_{L^2_r} = 0$, i.e.,

$$\langle S''_\mu(\phi_{\mu,m})v, v\rangle \geq \nu|v|^2_{L^2},$$

for any $v \in X_m$ with $\mathrm{Re}(v, i\phi_{\mu,m})_{L^2(\mathbb{R}^2,dx)} = 0$. This implies the statement (i). □

Proof of Lemma 4.3: First of all, we note that the formula (4.10) is well defined as an oscillatory integral untill the time $T_0 \wedge \tilde{T}$, and using Proposition 6 in [56], we see that if the initial data $u_0 \in \Sigma$, then $w(t) \in \Sigma$. Also, if u_0 is written in the form $u_0(x) = e^{im\theta}h(r)$ for some radial function $h(r)$, $w(t, x)$ defined by (4.10) is in such form too; indeed, for any $x \in \mathbb{R}^2$, define for any $g \in L^1(\mathbb{R}^2)$,

$$\tilde{g}(x) := \int_{\mathbb{R}^2} e^{i\beta(t)x\cdot y} g(y) dy. \tag{4.53}$$

With the following argument, it suffices to show that for any phase $\phi \in \mathbb{R}$, and any $f \in L^1(\mathbb{R}^2)$,

$$\tilde{f}(e^{i\phi}x) = \int_{\mathbb{R}^2} e^{i\beta(t)x\cdot y} f(e^{i\phi}y) dy \tag{4.54}$$

holds. Suppose $f(x) = e^{im\theta}h(r)$ where $m \in \mathbb{Z}$, $\theta \in \mathbb{R}$ and $x = re^{i\theta}$ ($r = |x|$). Then,

$$f(e^{i\phi}x) = f(x)e^{im\phi}$$

for any $x \in \mathbb{R}^2$ and $\phi \in \mathbb{R}$. We operate the transformation (4.53) on both sides and we get by (4.54), $\tilde{f}(e^{i\phi}x) = e^{im\phi}\tilde{f}(x)$. In particular, taking $x = r \geq 0$, we have $\tilde{f}(e^{i\phi}r) = e^{im\phi}\tilde{f}(r)$. Since r and ϕ are arbitrary, putting $z = re^{i\phi}$, for any $z \in \mathbb{R}^2$, the relation $\tilde{f}(z) = e^{im\phi}\tilde{f}(r)$ is satisfied. This means that any function in the form $e^{im\theta}h(r)$ is preserved in the same form by the transformation (4.53).

Now we verify the equality (4.54). Noting that

$$U = e^{i\phi} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

is unitary and its determinant equals 1,

$$\begin{aligned} \int_{\mathbb{R}^2} e^{i\beta(t)x\cdot y} f(e^{i\phi}y)dy &= \int_{\mathbb{R}^2} e^{i\beta(t)x\cdot y} f(Uy)dy = \int_{\mathbb{R}^2} e^{i\beta(t)x\cdot U^{-1}w} f(w)dw \\ &= \int_{\mathbb{R}^2} e^{i\beta(t)(U^{-1})^* x\cdot w} f(w)dw \\ &= \int_{\mathbb{R}^2} e^{i\beta(t)Ux\cdot w} f(w)dw = \tilde{f}(Ux) = \tilde{f}(e^{i\phi}x). \end{aligned}$$

We recall Theorem 1 of [56]. For fixed $T_0 > 0$ and $\alpha \in (0, 1/2)$, and for $\omega \in \Omega$ such that $W(\cdot, \omega) \in C^\alpha([0, T_0])$, there exist $T_\omega > 0$ and a propagator $\{U^\omega(t, s), t, s \in [0, T_0], |t-s| \leq T_\omega\}$ corresponding to Eq.(4.7). By the uniqueness of solutions, the solution of (4.7), constructed in (3.21) of [56], with the kernel of this propagator $U^\omega(t, 0)$ is the same as $w(t, x)$ defined by (4.10) for a small time interval $[0, T]$ with $T \leq T_0 \wedge T_\omega \wedge \tilde{T}$.

On the other hand, to consider the solution of the nonlinear equation (4.9), let $u_0 \in X_m$. Using the integral form for (4.9) with $w(0) = u_0$, for $t \in [0, T]$,

$$w(t) = U^\omega(t, 0)u_0 - i\lambda \int_0^t U^\omega(t, s)|w|^{2\sigma}w(s)ds. \tag{4.55}$$

We see that if $w \in X_m$, then $|w|^{2\sigma}w \in X_m$, and so, by the above argument we get $U^\omega(t, s)|w|^{2\sigma}w(s) \in X_m$, too.

In particular, the initial data $u_0 \in X_m$ belongs to Σ and thus it follows from Proposition 4.1 that there is a unique solution $w(\cdot) \in C([0, \tau^*_{u_0,\omega}), \Sigma)$ of (4.9). Since the space X_m is conserved by the equation (4.9), this solution is in fact in $C([0, \tau^*_{u_0,\omega}), X_m)$ almost surely. Since we are in the case $\lambda = 1$, it follows $\tau^*_{u_0,\omega} = +\infty$ from Proposition 4.1. □

Proof of Proposition 4.7: Local existence follows from the arguments of Theorem 4.10.1 in [38]. Let $u_0 \in \Sigma^2$ and fix $M, T > 0$. Set $I = [0, T]$ with $0 < T \leq T_0 \wedge \tilde{T} \wedge T_\omega$. The local existence in Σ^2 is proved by a fixed point method applied to the map

$$(\mathcal{T}^\omega w)(t) = U^\omega(t, 0)u_0 - i\lambda \int_0^t U^\omega(t, s)|w|^{2\sigma}w(s)ds,$$

where $U^\omega(t, 0)u_0$ is the solution of equation (4.7), with initial data u_0. Note that we drop

the ε in the notation for simplicity here, see however Remark 4.19 for the ε dependence of the estimates. Setting

$$B_M := \{v \in L^\infty(I;\Sigma^2), |v|_{L^\infty(I,\Sigma^2)} \leq M\}$$

with the metric $d(u,v) := |u-v|_{L^\infty(I,L^2(\mathbb{R}^2))}$, and following the arguments of Theorem 4.10.1 in [38], one may prove that $\mathcal{T}^\omega$ is a contraction mapping on (B_M, d) for $M = 2|u_0|_{\Sigma^2}$, provided that T is small enough, depending on $\omega, T_0, \tilde{T}$ and M. This allows us to show the local existence and blow up alternative in Σ^2.

This solution exists in fact globally. To see this, we follow the argument in [86]. Remind that $u_0 \in \Sigma^2$, then in particular $u_0 \in \Sigma$. Thus there exists a unique solution $w(t) \in C([0,\tau^*_{u_0,\omega}),\Sigma)$ of (4.9) with the maximal time $\tau^*_{u_0,\omega}$. Moreover, since $\lambda = 1$, $\tau^*_{u_0,\omega} = +\infty$ a.s. We may suppose that there exists a uniform constant $K_{\omega,T_0} > 0$ such that

$$\sup_{0\leq t\leq T_0} |w(t)|_\Sigma \leq K_{\omega,T_0} < \infty, \quad a.s. \tag{4.56}$$

This uniform bound in Σ implies that for any $q > 2$ there exists a constant $\tilde{K}_{\omega,T_0,q} > 0$ such that

$$\sup_{0\leq t\leq T_0} \{|\nabla w(t)|_{L^q(\mathbb{R}^2)} + |xw(t)|_{L^q(\mathbb{R}^2)}\} \leq \tilde{K}_{\omega,T_0,q} < \infty, \quad a.s. \tag{4.57}$$

Indeed, let $u_0 \in \Sigma^2$; Using (2) of Proposition 6, and Lemma 4.1 in [56], we obtain the existence of bounded real-valued functions $a_{jk,lm}(t,s)$ for $j,k,l,m \in \{1,2\}$ such that, for $t,s \in [0,T]$ with $|t-s| \leq T_\omega$, where T_ω is given by Lemma 4.1 in [56],

$$\begin{aligned} x_j U^\omega(t,s) &= U^\omega(t,s)x_j - (t-s)U^\omega(t,s)(i\partial_{x_j}) \\ &\quad +(t-s)\sum_{k=1}^{2}\{I(t,s,a_{jk,11})x_k + I(t,s,a_{jk,12})(i\partial_{x_k})\}, \\ i\partial_{x_j} U^\omega(t,s) &= U^\omega(t,s)i\partial_{x_j} + \sum_{k=1}^{2}\{I(t,s,a_{jk,21})x_k + (t-s)I(t,s,a_{jk,22})(i\partial_{x_k})\}, \end{aligned}$$

where we have set

$$I(t,s,a)f(x) = (2\pi i(t-s))^{-1}a(t,s)\int_{\mathbb{R}^2} e^{iS(t,s,x,y)}f(y)dy, \quad \text{for } f \in C_0^\infty(\mathbb{R}^2),$$

and $S(t,s,x,y)$ is a real valued continuous function of all its arguments (see [56]). Using then the integral equation (4.55), we easily deduce that for $t \in [0,T]$, $q \in (2,\infty)$, $q' =$

$q/(q-1)$, and $j=1,2$,

$$\begin{aligned}|\partial_{x_j} w(t)|_{L^q(\mathbb{R}^2)} + |x_j w(t)|_{L^q(\mathbb{R}^2)} &\leq (1+T)|U^\omega(t,0)\nabla u_0|_{L^q(\mathbb{R}^2)} + 2|U^\omega(t,0)xu_0|_{L^q(\mathbb{R}^2)} \\ &+ \sum_{k=1}^{2}\Big\{|I(t,0,a_{jk,21})x_k u_0|_{L^q(\mathbb{R}^2)} + t|I(t,0,a_{jk,22})\partial_{x_k}u_0)|_{L^q(\mathbb{R}^2)} \\ &+ t|I(t,0,a_{jk,11})x_k u_0|_{L^q(\mathbb{R}^2)} + t|I(t,0,a_{jk,12})\partial_{x_k}u_0|_{L^q(\mathbb{R}^2)}\Big\} \\ +C_{\omega,T_0}\int_0^t \Big(|t-s|^{-(1-2/q)} + |t-s|^{2/q}\Big)\Big(|\partial_{x_j}(|w|^{2\sigma}w)|_{L^{q'}(\mathbb{R}^2)} &+ |x_j|w|^{2\sigma}w|_{L^{q'}(\mathbb{R}^2)}\Big)\,ds.\end{aligned}$$

Thanks to the Sobolev embedding $\Sigma \subset L^q(\mathbb{R}^2)$ and the continuity in Σ of $U^\omega(t,s)$ and $I(t,s,a)$, the first three lines in the right hand side above are estimated by $C_{T_0,\omega}|u_0|_{\Sigma^2}$. On the other hand, for any $l<\infty$ with $\frac{1}{l}=\frac{1}{2}-\frac{1}{q}$, by the Sobolev embeddings in $\mathbb{R}^2$, for $j=1,2$

$$\begin{aligned}|\partial_{x_j}(|w|^{2\sigma}w)|_{L^{q'}(\mathbb{R}^2)} &\leq C|w|^{2\sigma}_{L^{\sigma l}(\mathbb{R}^2)}|\nabla w|_{L^q(\mathbb{R}^2)} \leq C|w|^{2\sigma}_{\Sigma}|\nabla w|_{L^q(\mathbb{R}^2)}, \\ |x_j|w|^{2\sigma}w(s)|_{L^{q'}(\mathbb{R}^2)} &\leq C|w|^{2\sigma}_{\Sigma}|xw|_{L^q(\mathbb{R}^2)}.\end{aligned}$$

In summary, we obtain by (4.56), for $t\in[0,T]$,

$$\begin{aligned}|\nabla w(t)|_{L^q} + |xw(t)|_{L^q} &\leq C_{\omega,T_0}|u_0|_{\Sigma^2} \\ &+ C_{\omega,T_0}\int_0^t \Big(|t-s|^{-(1-2/q)} + |t-s|^{2/q}\Big)\left(|\nabla w(s)|_{L^q} + |xw(s)|_{L^q}\right)ds.\end{aligned}$$

Note that $|t|^{-(1-2/q)}$ is integrable near $t=0$, and then by Gronwall's inequality, together with an iteration argument on $[T,2T]$, ... we obtain (4.57). Once (4.57) is proved, by the Sobolev embedding, we conclude that $w\in L^\infty(\mathbb{R}^2)$ and the L^∞ norm is also uniformly bounded in time since $q>2=d$.

We finally estimate the solution of (4.55) in Σ^2 norm, using the fact that for $\sigma\geq 1/2$,

$$\begin{aligned}|w(t)|_{\Sigma^2} &\leq |U^\omega(t,0)u_0|_{\Sigma^2} + \int_0^t |U^\omega(t,s)|w|^{2\sigma}w(s)|_{\Sigma^2}ds \\ &\leq |u_0|_{\Sigma^2} + C_{\omega,T_0}\int_0^t ||w|^{2\sigma}w(s)|_{\Sigma^2}ds \\ &\leq |u_0|_{\Sigma^2} + C_{\omega,T_0}\int_0^t \Big\{|w|^{2\sigma}_{L^\infty}(|w|_\Sigma + |x\nabla w|_{L^2} + |x^2 w|_{L^2}) \\ &\quad + |w|^{2\sigma-1}_{L^\infty}|\nabla w|^2_{L^4} + |w|_{L^\infty}|\nabla^2 w|_{L^2}\Big\}ds \\ &\leq C_{\omega,T_0}\int_0^t |w(s)|_{\Sigma^2}ds,\end{aligned}$$

where we have used (4.57). The uniform bound in Σ^2 follows a.s. ω on any interval $[0,T_0]$ again from the Gronwall inequality. □

Remark 4.19. *It follows from the above computations, together with the fact that all the constants appearing in Proposition 6 and Lemma 4.1 in [56] depend only on* $|W|_{C^\alpha([0,T_0])}$*, for some* $\alpha > 0$*, that if we replace* W *by* εW*, and if the constant in* (4.56) *is uniform for* $\varepsilon \leq 1$*, then the bound on* $\sup_{t\in[0,T_0]} |u^\varepsilon(t)|_{\Sigma^2}$ *is also uniform for* $\varepsilon \leq 1$*. Now, since* $\sup_{t\in[0,\tau^\varepsilon\wedge T_0]} |u^\varepsilon(t)|_\Sigma$ *is uniformly bounded for* $\varepsilon \leq 1$ *by the definition of* τ^ε*, we deduce that*

$$\lim_{N\to+\infty} \mathbb{P}\left(\sup_{t\in[0,\tau^\varepsilon\wedge T_0]} |u^\varepsilon(t)|_{\Sigma^2} \geq N \right) = 0,$$

uniformly for $\varepsilon \leq 1$*.*

Deuxième partie

Méthodes numériques pour la modélisation d'un condensat de Bose-Einstein à température non-nulle

Chapitre 5

Generalized and hybrid Metropolis-Hastings overdamped Langevin algorithms

It has been shown that the nonreversible overdamped Langevin dynamics enjoy better convergence properties in terms of spectral gap and asymptotic variance than the reversible one ([96, 97, 113, 170, 150, 151, 72]). In this chapter we propose a variance reduction method for the Metropolis-adjusted Langevin Algorithm (MALA) that makes use of the good behaviour of these nonreversible dynamics. It consists in constructing a nonreversible Markov chain (with respect to the target invariant measure) by using a Generalized Metropolis-Hastings adjustment on a lifted state space. We present two variations of this method and we discuss the importance of a well-chosen proposal distribution in terms of average rejection probability. We conclude with numerical experiments to compare our algorithms with MALA, and show variance reductions of several orders of magnitude in some favourable toy cases.

This chapter corresponds to the preprint [141] "Generalized and hybrid Metropolis-Hastings overdamped Langevin algorithms".

5.1 Prerequisites

This chapter deals with numerical methods to sample a random variable distributed according to some target distribution. Since these methods are not specific to Bose-Einstein condensation, we present them in a general setting, as it has been done in the preprint [141]. Nevertheless, we take advantage of this manuscript to explain, in the first section, in which way these methods are related to Bose-Einstein condensation, and we recall some general results about the overdamped Langevin equation.

The overdamped Langevin equations The Langevin equations are a class of diffusions of the form,

$$dX_t = -\nabla U(X_t)\,\mathrm{d}t + \gamma(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{5.1}$$

where the process (X_t) takes values in $\mathbb{R}^d$, and where U is called a *potential*, which is a real valued function that we suppose to be continuously differentiable and such that $e^{-U} \in L^1(\mathbb{R}^d)$. We assume in addition that γ is divergence-free, by which we means that the following condition holds,

$$\nabla \cdot (\gamma e^{-U}) = 0.$$

An easy way to construct such a vector field is to notice that for any skew-symmetric matrix J, the vector field $J\nabla U$ is divergence-free. Moreover, under the divergence-free condition, the following equation,

$$dX_t = \gamma(X_t)\,\mathrm{d}t, \tag{5.2}$$

conserves the energy U, which justifies the Hamiltonian denomination of the term γ.

We suppose from now on that Equation (5.1) possesses a unique, non-explosive, strong solution. A condition of non-explosion is given in Theorem 5.1. This dynamics is of special interest since, under some regularity conditions, it is ergodic with unique invariant Gibbs measure π given (up to a multiplicative constant) by,

$$\pi(dx) \propto e^{-U(x)}dx, \quad \forall x \in \mathbb{R}^d. \tag{5.3}$$

We begin by recalling the definition of an invariant measure and Birkhoff's Ergodic Theorem which is of particular interest in this chapter. Suppose that the solution $(X_t)_{t\geq 0}$ of Equation (5.1) is initialized with the measure π. Then π is an invariant measure for Equation (5.1) if and only if for all positive time t, $\mathrm{Law}(X_t) = \pi$, that is, if and only if the dynamics (5.1) leaves *invariant* the measure π. In this case, and for every observable f such that $f \in L^1(\pi)$, Birkhoff's Ergodic theorem states that,

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T f(X_s)ds = \int_{R^d} f(x)\pi(dx), \quad \text{a.s. and a.e. } X_0.$$

This theorem is of practical interest since it enables to replace a high-dimensional integration by a one-dimensional one.

Theorem 5.1 ([152, Theorem 2.1])**.** *Suppose that the potential U is continuously differentiable, and that, for some positive constants $R, a, b < +\infty$, for all $x \in \mathbb{R}^d$,*

$$\|x\| > R \Rightarrow (-\nabla U(x) + \gamma(x)) \cdot x \leq a\,\|x\|^2 + b.$$

Then, the solution of Equation (5.1) is non-explosive.

When the solution of Equation (5.1) is non-explosive, the measure π, given by Equation (5.3) is invariant for the solution of Equation (5.1), which is ergodic with respect to π.

The original theorem is stated in [152] in the case $\gamma = 0$. Its extension to the case $\gamma \neq 0$ is not complicated. The proof of this theorem relies on the fact that π is invariant for the Langevin dynamics, which itself relies on the very important characterization of an invariant measure of the dynamics (5.1) in terms of its generator. This characterization states that a probability measure π is invariant for Equation (5.1) if and only if for any C^∞ and compactly supported test function ϕ,

$$\int_{\mathbb{R}^d} \mathcal{L}\phi \, d\pi = 0, \tag{5.4}$$

where $\mathcal{L}$ is the generator of the diffusion (5.1) given by,

$$\mathcal{L}\phi = -\nabla U \cdot \nabla\phi + \gamma \cdot \nabla + \Delta\phi. \tag{5.5}$$

Equation (5.4) follows from a simple integration by part. The rest of the proof of this theorem follows from [125, Theorem 6.1] and requires to prove furthermore that the process is Harris recurrent and that some skeleton of the chain is irreducible (for the Lebesgue measure).

In the special case $\gamma = 0$, the solution of the overdamped Langevin equation (5.1) is reversible with respect to π. This means that if Equation (5.1) is initialized with its invariant measure, then for all $n \in \mathbb{N}$ and all $t_0 \leq t_1 \leq \cdots \leq t_n$,

$$\text{Law}\left((X_{t_0}, X_{t_1}, \ldots, X_{t_n})\right) = \text{Law}\left((X_{t_n}, \ldots, X_{t_1}, X_{t_0})\right)$$

A characterization of this property can be given using the generator $\mathcal{L}$ of Equation (5.1) given by Equation (5.5).

Theorem 5.2 (Theorem 4.5 [133])**.** *Let $\mathcal{L}$ be the generator of a stationary diffusion $(X_t)_{t\geq 0}$ of invariant probability measure π. This diffusion is reversible with respect to π if and only if for any test functions ϕ_1 and ϕ_2,*

$$\int_{\mathbb{R}^d} \phi_1 \mathcal{L}(\phi_2)\pi = \int_{\mathbb{R}^d} \phi_2 \mathcal{L}(\phi_1)\pi,$$

that is to say if and only if its generator is self-adjoint in $L^2(\pi)$.

Exponential convergence of the overdamped Langevin equation. It is of particular interest to know at what speed the Langevin Equation (5.1) *converges* towards the Gibbs measure π. This interest partly comes from the fact that this dynamics can be used to design numerical methods to sample random variables distributed with respect to π. They actually are the object of the current chapter. We will present in the following some

classical results about a large class of Langevin equations that converge exponentially fast towards their invariant measures.

One strong criterion to measure the convergence of the law of the solution of Equation (5.1) towards π is the distance in total variation given for any couple of probability measures (ν_1, ν_2) by,

$$d_{\mathrm{TV}}(\nu_1, \nu_2) = \sup_{\|f\|_\infty \leq 1} \left| \int_{\mathbb{R}^d} f(x)(\nu_1 - \nu_2)(dx) \right|$$

For quite a general class of potentials U, the solution of the overdamped Langevin Equation (5.1) converges exponentially fast towards π. A sufficient condition to ensure this rate of convergence is the existence of a Foster-Lyapunov function, for which a drift condition holds.

Definition 5.3. *We say that a drift condition, or a Foster-Lyapunov criterion, holds for a diffusion of generator $\mathcal{L}$ with the Lyapunov function $V : \mathbb{R}^d \to [1, +\infty[$, if there exist $c > 0$ and $b \in \mathbb{R}$ such that for all $x \in \mathbb{R}^d$,*

$$\mathcal{L}V(x) \leq -cV(x) + b\mathbb{1}_C(x), \tag{5.6}$$

where C is a compact.

Theorem 5.4 ([68, Theorem 5.2.c])**.** *Let $\mathcal{L}$ be a generator of a diffusion $(X_t)_{t\geq 0}$ with invariant measure π. Suppose in addition that a drift condition (5.6) holds for this diffusion with a Lyapunov function V. Then, there exist two positive constants c and λ such that,*

$$d_{TV}(p_t(x, \cdot), \pi) \leq cV(x)e^{-\lambda t},$$

where $p_t(x, \cdot)$ is the law of X_t given that $X_0 = x$.

From a practical viewpoint, the following lemma provides a sufficient condition for a drift condition to be satisfied for the specific dynamics (5.1).

Lemma 5.5 ([72, Lemma 1])**.** *Suppose that there exists a skew-symmetric matrix J such that $\gamma = J\nabla U$, and that the density π is bounded and for some $0 < \beta < 1$,*

$$\liminf_{\|x\| \to +\infty} \left((1 - \beta) \|\nabla U(x)\|^2 - \Delta U(x) \right) > 0,$$

then the drift condition (5.6) holds with

$$V(x) = \pi(x)^{-\beta}(x).$$

SPGPE as a nonreversible Langevin equation. We present in this section one way to formulate the SPGPE model presented in Section 1.3 as a nonreversible Langevin

equation. We recall that this model is given by, ((1.17)):

$$d\phi_t = -(i + c_1)\nabla E^K(\phi_t)\,\mathrm{d}t + c_1 c_3 dB_t. \tag{5.7}$$

As explained in the introduction in Section 1.3, the meaning of the gradient of E^K is not clear since the domain of E^K could be considered to be a $\mathbb{C}$-vector space, and its codomain could be considered as a $\mathbb{R}$-vector space. Actually, we consider the domain of E^K to be a vector space with real scalar field to make sense of ∇E^K. This way, $\mathbb{C}$ is identified as $\mathbb{R}^2$, and the multiplication by i is interpreted as an endomorphism of K, seen as a $\mathbb{R}$-vector space. We call J this operator, which can be interpreted as a skew-symmetric matrix of size $2\dim(K)$ (if $\dim(K)$ is the dimension of K seen as a $\mathbb{C}$-vector space). Then, Equation (5.7) can be written,

$$d\phi_t = -(J + c_1 Id)\nabla E^K(\phi_t)\,\mathrm{d}t + c_1 c_3 dB_t, \tag{5.8}$$

with $\nabla E^K(\phi_t) \in \mathbb{R}^{2\dim(K)}$. This last relation makes its interpretation as a nonreversible Langevin equation clear. We also recall that if we consider the mapping E^K to take values in $\mathbb{C}$, and if we consider the domain of E^K to be a vector space with complex scalar field, then E^K would not be differentiable.

5.2 Introduction

This chapter proposes a new class of MCMC algorithms whose objective is to compute expectations

$$\pi(f) := \mathbb{E}_\pi(f) = \int_{\mathbb{R}^d} f(x)\pi(dx), \tag{5.9}$$

for a given observable f, with respect to a probability measure $\pi(dx)$ absolutely continuous, with respect to the Lebesgue measure, with density $\pi(x) = e^{-U(x)}$. We suppose, as it is the case in many practical situations, that π is only known up to a multiplicative constant.

Many techniques have been developed to solve this problem. Deterministic quadratures can be very efficient for low dimensional spaces. Yet, in the high dimensional case, these methods tend to become inefficient or even impossible to apply, and MCMC methods can be used instead. The basic idea is to construct an ergodic Markov chain with respect to π, and to approximate $\pi(f)$ by the time average of this Markov chain. There are infinitely many ways to construct such a discrete time process. The general idea is to use an approximate time discretization of a time-continuous process known to be ergodic with respect to π. Generally, we cannot expect any approximate time discretization to be ergodic with respect to π. Indeed, the discretized chain could be ergodic with respect to a perturbed measure $\pi_{\delta t}$, or even transient [123, 152]. Thus one can use a Metropolis-Hastings

acceptance-rejection step that ensures the detailed balance, and thus makes the chain reversible and ergodic with respect to π. In the case of the Euler-Maruyama discretization of the overdamped Langevin dynamics,

$$dX_t = \nabla \log \pi(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{5.10}$$

with $(W_t)_{t\geq 0}$ a standard Brownian motion in $\mathbb{R}^d$, the method is called, in the computational statistics literature, Metropolis Adjusted Langevin Algorithm (MALA, [152]). Actually this method was already known in in the chemistry literature as Smart Monte-Carlo [154].

Yet, it has been noticed in several contexts that departing from the reversibility can improve the performances of MCMC methods. This chapter aims at proposing a generalization of the standard MALA that can be able to construct nonreversible Markov chains that can outperform classical MALA.

5.2.1 Nonreversible dynamics

On the continuous time setting, studies have been carried out to compare the convergence properties of some time-continuous dynamics that are ergodic with respect to π [72, 96, 97, 113, 150, 151, 170], based on two kinds of optimality criterion: the speed of convergence towards equilibrium, measured in terms of spectral gap in $L^2(\pi)$, and asymptotic variance of the time averages of given observable. Obviously, from a computational viewpoint, an increase in the spectral gap enables to reduce the burn-in, and a reduction of the asymptotic variance leads to a decrease of the computational complexity of the corresponding MCMC method. These analyses compare, for different vector fields γ, the overdamped Langevin dynamics given by Equation (5.1) where the divergence-free condition $\nabla \cdot (\gamma\pi) = 0$ is satisfied. It is well known that among all vector fields γ satisfying the non explosion condition and the divergence-free condition, the dynamics given by (5.10) in the reversible case ($\gamma = 0$) has the worst rate of convergence in terms of spectral gap in $L^2(\pi)$ [96, 97, 113]. Recent work has been done to construct divergence free (with respect to π) perturbations of the drift that achieve optimal convergence properties in the Gaussian case [113, 170]. Recent works also show that breaking the non reversibility with such divergence free perturbations on the drift also leads to improvement on the asymptotic variance. It is shown in [150] that the asymptotic variance decreases under the addition of the irreversible drift. Moreover, is has been shown in [151] that the asymptotic variance is monotonically decreasing with respect to the growth of the drift, and the limiting behavior for infinitely strong drifts is characterized. More recently, in [72] the authors investigate the dependence of the asymptotic variance on the strength of the nonreversible perturbation.

On the discrete time setting, classical methods that depart from reversible sampling consist in hybrid (Hamiltonian) MCMC [70, 114] and generalized hybrid MCMC methods [105]. In the former method, the drift direction is chosen isotropically at each time step and long time Hamiltonian integration is then carried out in this direction. The latter can be seen as a generalization of the former that brings some inertia in the direction of

the Hamiltonian dynamics. Another class of nonreversible samplers is composed by lifting methods. They are designed in the discrete state space case to construct a Markov chain that satisfies some skew detailed balance [43, 66, 94, 165]. They consist in increasing the state space to take into account a privileged drift direction that is explored more efficiently. More recently, Bierkens proposed an extension of the classical Metropolis-Hastings algorithm to generate unbiased nonreversible Markov chain [23]. This is achieved by modifying the acceptance probability to depart from detailed balance. Eventually, a quite different approach has been proposed in [128, 137] in the one dimensional case, and then generalised in [24] in the multidimensional case. The authors construct a continuous time piecewise deterministic Markov process. It is a constant velocity model where the velocity direction switches at random times with a rate depending on the target distribution.

5.2.2 Outline

We propose in this chapter a bias-free algorithm similar to MALA that aims at exploiting the asymptotic variance reduction of the nonreversible time-continuous process. The idea is to construct a Markov chain with invariant measure π, by discretizing an equation of the form (5.1), instead of Equation (5.10), enhanced with an acceptance-rejection step. The main difficulty consists in unbiasing the unadjusted chain. Indeed, it is not worth considering the use of a standard Metropolis-Hastings acceptance probability since it is designed to impose detailed balance with respect to the target distribution π, and thus defines a reversible Markov chain with respect to π. It would lead to a poor average acceptance ratio. Nevertheless, a discussion about such a method is conducted in [131] where the authors consider the special case of a Gaussian target in the limit as the dimension diverges to infinity. An elegant way would be to construct an adequate acceptance probability with respect to this proposal, to ensure a high average acceptance ratio. In the setting of [23], it would consist in finding a good vorticity kernel. Yet, we are not able to exactly do this. Instead, we propose a class of lifted algorithms that rely on these unadjusted chains. More precisely the first algorithm is a generalized Metropolis-Hastings algorithm in an enhanced state space, and the second one can be seen as the analogous of the generalized hybrid Monte Carlo method for the overdamped Langevin equation.

In Section 5.3 we present the first algorithm (generalized MALA). We discuss in 5.3.1 how its performances are closely related to the choice of the transition kernel of the unadjusted chain. In Section 5.3.2 we prove geometric convergence of the Markov chain constructed with this algorithm under some hypotheses, that ensure the existence of a central limit theorem. Then, in Section 5.4 we propose a modification of this algorithm (the generalized hybrid MALA). In Section 5.5 we present numerical comparisons of these algorithms with respect to classical MALA, which is followed by concluding remarks.

5.3 Generalized MALA

In this section, we construct a nonreversible Markov chain, ergodic with respect to a target distribution π known up to a normalizing constant. The algorithm is similar to MALA in the sense that it constructs a Markov chain from the discretization of an overdamped Langevin dynamic, augmented with an acceptance-rejection step that makes it ergodic with respect to π. The difference is that we construct a Markov chain on the discretization of a nonreversible Langevin equation to try to benefit from the smaller asymptotic variance of this kind of Markov processes, than the reversible ones. The main issue is then to choose a right acceptance probability that preserves the good ergodic properties of the underlying Markov process.

To state our algorithm, we slightly modify Equation (5.1). For $\xi \in \mathbb{R}$, we consider the diffusions,

$$dX_t = \nabla \log \pi(X_t)\,\mathrm{d}t + \xi\gamma(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{5.11}$$

with divergence-free condition $\nabla \cdot (\gamma\pi) = 0$. This way, ξ specifies the direction and the intensity of the nonreversibility. We denote now by Q^ξ a proposal kernel that correspond to some discretization of the diffusions (5.11) with parameter ξ. We propose the following algorithm that we call Generalized MALA (GMALA),

Algorithm 5.6 (Generalized MALA)**.** *Let $h > 0$, $(x^0, \xi^0) \in \mathbb{R}^d \times \mathbb{R}$ be an initial point and an initial direction. Iterate on $n \geq 0$.*

1. *Sample y^{n+1} according to $Q^{\xi^n}(x^n, dy)$.*
2. *Accept the move with probability*

$$A^{\xi^n}(x^n, y^{n+1}) = 1 \wedge \frac{\pi(y^{n+1})Q^{-\xi^n}(y^{n+1}, x^n)}{\pi(x^n)Q^{\xi^n}(x^n, y^{n+1})}. \tag{5.12}$$

and set $(x^{n+1}, \xi^{n+1}) = (y^{n+1}, \xi^n)$; otherwise set $(x^{n+1}, \xi^{n+1}) = (x^n, -\xi^n)$.

The important part of this algorithm is that the direction ξ^n of the Hamiltoninan exploration must be inverted at each rejection to ensure its unbiasedness, as in Generalized Hybrid Monte-Carlo methods. A good choice of Q^ξ is given in the next section.

The unbiasedness of Algorithm 5.12 is obvious since this algorithm is actually built as a Generalized Metropolis-Hastings algorithm on the increased state space $E = \mathbb{R}^d \times \{-\xi_0, \xi_0\}$. To simplify notation, we denote by x_ξ all element $(x, \xi) \in E$. We set S the involutive transformation defined for all element $x_\xi \in E$ by $S(x_\xi) = x_{-\xi}$. We extend the definition of π on E by $\pi(dxd\xi) = \frac{\pi(x)}{2}(\delta_{-\xi_0}(d\xi) + \delta_{\xi_0}(d\xi))dx$ for $x_\xi \in E$. Obviously π is unchanged by S. Then, the algorithm constructs a Markov chain with transition kernel

density P given by,

$$P(x_\xi, y_\eta) = Q^\xi(x,y)A^\xi(x,y)\mathbb{1}_\xi(\eta) + \delta_{S(x_\xi)}(y_\eta)\left(1 - \int_{\mathbb{R}^d} Q^\xi(x,z)A^\xi(x,z)dz\right),$$

where $\mathbb{1}_{xi}$ denotes the characteristic function and $\delta_{S(x_\xi)}$ a Dirac delta function, that satisfies the following skew detailed balance,

$$\forall x_\xi, y_\eta \in E, \quad \pi(x_\xi)P(x_\xi, y_\eta) = \pi(y_\eta)P(S(y_\eta), S(x_\xi)). \tag{5.13}$$

Heuristically, we hope that the discretization of the time-continuous process (5.1) specified by Q^ξ benefits from the same good behavior in terms of asymptotic variance reduction. Since the Markov chain is constructed as parts of the discretization of the time-continuous dynamics between the rejections, then the closer to one is the average acceptance probability, the longer are these parts, and the more we can hope to benefit from this good behavior. Thus, to give a hint about the relevance of this algorithm, we compute in Section 5.3.1 these average acceptance probabilities, and we show that they are of the same order of those for MALA for some well-chosen discretization. Moreover, we numerically show in Section 5.5 that it can outperform MALA by several order of magnitude in terms of asymptotic variance.

Yet, before showing these results, we present a heuristic justification about this algorithm. In the reversible case, the time-continous equation satisfies the detailed balance with respect to π,

$$\forall x, y \in \mathbb{R}^d, \forall h > 0, \quad \pi(x)P_h(x,y) = \pi(y)P_h(y,x),$$

where $P_h(x,y)$ denotes the density of the transition kernel to go from x to y after a time h for the dynamics (5.10). Noting $\mathcal{L}$ the generator of the diffusion (5.10), the kernel P_h is given by $P_h = e^{h\mathcal{L}}$. Moreover, MALA imposes that the Markov chain also satisfies the detailed balance with respect to π. In the nonreversible case, the time continuous process (5.11) satisfies a skew detailed balance as stated by the following lemma,

Lemma 5.7. *For all x, $y \in \mathbb{R}^d$, for all $\xi \in \mathbb{R}$ and for all $h > 0$, the following relation holds,*

$$\pi(x)P_h^\xi(x,y) = \pi(y)P_h^{-\xi}(y,x), \tag{5.14}$$

where $P_h^\xi(x,dy)$ is the transition probability measure of the h-skeleton of the process $(X_t^\xi)_{t\geq 0}$, solution of Equation (5.11).

The analogous of the reversible case would be to construct a π-invariant Markov chain that satisfies the same kind of skew detailed balance. Because the skew detailed balance (5.14) can be seen as a detailed balance on the enhanced state space E up to the

transformation S,

$$\forall x, y \in \mathbb{R}^d, \forall \xi \in \mathbb{R}, \forall h > 0 \quad \pi(x) P_h(x_\xi, y_\xi) = \pi(y) P_h(S(y_\xi), S(x_\xi)), \tag{5.15}$$

where $P_h(x_\xi, y_\xi) = P_h^\xi(x, y)$, the Generalized Metropolis-Hastings method given by Algorithm 5.6 is actually the classical way to construct such a Markov chain. Nevertheless, the main difference between the time-continous and the discrete time dynamics is the fact that the latter requires some direction switching of the nonreversible component of the dynamics. As stated previously, the idea of lifting the state space to construct a nonreversible chain has been used in the discrete state space setting. Yet, the idea is quite different. With classical lifting methods, the goal is to switch between several directions with well-chosen probabilities, to quickly explore the state space in all of these directions. In our case, we do not aim to switch between several nonreversible directions. Yet, we are forced to do so at each rejection, and we have no choice but to reverse the current nonreversible directions. That is to say that we control neither the probability of switching nor the direction. On the contrary, we can cite the Bouncy Particle Sampler (BPS) which is a piecewise deterministic Markov process that explores the state space by changing the direction of the nonreversibility [28, 128, 137]. We can also note that the intensity of the non-reversibility could be defined as a random variable to generalize this method.

Proof of Lemma 5.7. We denote by $\mathcal{L}^\xi$ the generator of the diffusion (5.11). One can show that,

$$\mathcal{L}^\xi = \mathcal{S} + \xi \mathcal{A},$$

where $\mathcal{S} = \nabla \log \pi(x) \nabla \cdot$, and $\mathcal{A} = \gamma \cdot \nabla$. Then, one can show that $\mathcal{S}$ and $\mathcal{A}$ are respectively the symmetric and the antisymmetric parts of $\mathcal{L}^1$, with respect to $L^2(\pi)$. Moreover, one can show (following Theorem VIII.3 [147]) that for all $\xi \in \mathbb{R}$, $\mathcal{L}^\xi$ and $\mathcal{L}^{-\xi}$ are essentially adjoint in $L^2(\pi)$ one from another with $D(\mathcal{L}^\xi) = D(\mathcal{L}^{-\xi}) = C_c^2(\mathbb{R}^d)$, the set of twice continuously differentiable functions. Thus, there exists a unique extension of $\mathcal{L}^\xi$ and $\mathcal{L}^{-\xi}$ such that they are adjoint one from another. Considering this extension, it follows that the semigroups $e^{\mathcal{L}^\xi t}$ and $e^{\mathcal{L}^{-\xi} t}$ are adjoint one from another (Corollary 10.6, Chapter 1, [134]). Then, for all bounded functions f and g,

$$\begin{aligned}
\int_{(\mathbb{R}^d)^2} f(x) g(y) P_h^\xi(x, dy) \pi(dx) &= \int_{\mathbb{R}^d} f(x) e^{\mathcal{L}^\xi h} g(x) \pi(dx), \\
&= \int_{\mathbb{R}^d} e^{(\mathcal{L}^\xi)^* h} f(x) g(x) \pi(dx), \\
&= \int_{\mathbb{R}^d} e^{\mathcal{L}^{-\xi} h} f(x) g(x) \pi(dx), \\
&= \int_{(\mathbb{R}^d)^2} f(y) g(x) P_h^{-\xi}(x, dy) \pi(dx).
\end{aligned}$$

□

5.3.1 Choice of the proposition kernel

The method presented above can be tuned with the choice of the proposition kernel Q^ξ. It should be chosen such that it approximates the law of the transition probability of the h-skeleton of the process $(X_t^\xi)_{t\geq 0}$ solution of equation (5.11). The basic idea is to define Q^ξ as the density of the law of a discretization of Equation (5.11). Doing so, we can hope that the skew-detailed balance for the unadjusted chain would be almost satisfied (since it is satisfied for the time-continous dynamics), that is to say we can hope to benefit from an acceptance ratio close to 1. Moreover, we recall that the choice of the proposal kernel is the only degree of freedom that enables to tune the rate of the direction switching. Formally, the closer the proposal kernel is from the transition kernel of the h-skeleton of the continuous process, the larger the acceptance ratio is (and thus the less frequent is the direction switching).

The basic idea would be to propose according to the Maruyama-Euler approximation of Equation (5.11), as it is done for MALA. We denote by Q_1^ξ this kernel. That is to say $Q_1^\xi(x, dy)$ is given by the law of y, solution of

$$y = x - h\nabla U(x) - h\xi\gamma(x) + \sqrt{2h}\chi, \tag{5.16}$$

where χ a standard normal deviate. Then, Q_1^ξ is given by

$$Q_1^\xi(x, dy) = \frac{1}{(4\pi h)^{d/2}} \exp\left(\frac{1}{4h}\|y - (x - h\nabla U(x) - h\xi\gamma(x))\|^2\right) dy. \tag{5.17}$$

Sadly, even though the simplicity of this proposal is appealing, this proposition kernel leads to an average rejection rate of order h when $\xi \neq 0$ and γ is non-linear as stated in Proposition 5.12. It is significantly worse than MALA that enjoys an average rejection rate of order $h^{3/2}$ (see [27]). As shown later by the numerical simulations, this bad rejection rate is not a pure theoretical problem: it forbids to use large discretization steps h, and thus the method only generates highly correlated samples.

To overcome this problem, we propose to implicit the resolution of the nonreversible term in Equation (5.11) with a centered point discretization. More precisely, we propose a move y_ξ from x_ξ, such that y would be distributed according to the solution of

$$y = x - h\nabla U(x) - h\xi\gamma\left(\frac{x+y}{2}\right) + \sqrt{2h}\chi, \tag{5.18}$$

where χ a standard normal deviate. We define the function $\Phi_x^{h\xi}$ by,

$$\Phi_x^{h\xi} : \quad \mathbb{R}^d \to \mathbb{R}^d, \quad y \mapsto y + h\xi\gamma\left(\frac{x+y}{2}\right). \tag{5.19}$$

Then, Equation (5.18) can be written as,

$$\Phi_x^{h\xi}(y) = x - h\nabla U(x) + \sqrt{2h}\chi.$$

The existence of such a proposal kernel, denoted by Q_2^ξ, is ensured under the hypothesis that γ is globally Lipschitz, for h sufficiently small.

Assumption 5.8. *We suppose that γ is globally Lipschitz.*

Proposition 5.9. *Under Assumption 5.8, there exists $h_0 > 0$, such that for all $h < h_0$, for all $x_\xi \in E$, for all $\chi \in \mathbb{R}^d$, there is a unique $y \in \mathbb{R}^d$ solution of Equation (5.18). Moreover, when χ is a standard normal deviate, the function Y_ξ defined almost surely as the solution of Equation (5.18) is a well-defined random variable, with law $Q_2^\xi(x, dy)$, given by,*

$$Q_2^\xi(x, dy) = \frac{1}{(4\pi h)^{d/2}} |\operatorname{Jac} \Phi_x^{\xi h}(y)| \exp\left(\frac{1}{4h}\|\Phi_x^{h\xi}(y) - (x - h\nabla \log \pi(x))\|^2\right) dy. \quad (5.20)$$

Finally

$$x - y = \sqrt{2h}\chi + O\left(h\|\nabla U(x)\| + h\|\gamma(x)\| + h^{3/2}\|\chi\|\right),$$

and there exists $C > 0$, independent of x, χ, and h, such that,

$$\|\nabla U(y)\| - \|\nabla U(x)\| \leq Ch(\|\nabla U(x)\| + \|\gamma(x)\|) + C\sqrt{2h}\|\chi\|.$$

Proof of Proposition 5.9. The first point is a corollary of Lemma 5.10 below. The second point can be proven by making use of the fact that ∇U is Lipschitz. □

The following lemma is used to prove Proposition 5.9. We can note that the maximal time step h_0 given by this lemma depends on the Lipschitz constant of ∇U.

Lemma 5.10. *Under Assumption 5.8, there exists $h_0 > 0$, such that for all $h < h_0$, for all $x_\xi \in E$, the function $\Phi_x^{h\xi} : \mathbb{R}^d \to \mathbb{R}^d$, defined by (5.19) is a C^1-diffeomorphism.*

Proof of Lemma 5.10. The proof that $\Phi_x^{h\xi}$ is bijective relies on Picard fixed point theorem. For any fixed $z \in \mathbb{R}^d$, we define Ψ on $\mathbb{R}^d$ by $\Psi(y) = z - h\xi\gamma\left(\frac{x+y}{2}\right)$. Then, since γ is supposed to be globally Lipschitz, then for small enough h this application is a contraction mapping from $\mathbb{R}^d$ to $\mathbb{R}^d$. The fact that $\Phi_x^{h\xi}$ is everywhere differentiable is clear, and the fact that its inverse is everywhere differentiable as well comes from the fact that the determinant of the Jacobian of $\Phi_x^{h\xi}$ is strictly positive for sufficiently small h. □

Sampling from the this proposal kernel can be done by a fixed point method when no analytical solution is available since the proof of Lemma 5.10 uses a Picard fixed point argument. This transition kernel leads to a rejection rate of order $h^{3/2}$ (see Proposition 5.12), which is an improvement from Q_1^ξ. The main advantage of this kernel is that the computation of $|\operatorname{Jac} \phi_x^{\xi h}(y)|$ can be avoided in the case where γ is defined by $\gamma(x) = J\nabla \log \pi(x)$,

with J a skew symmetric matrix. Indeed, only the ratio $|\operatorname{Jac}\phi_x^{\xi h}(y)|/|\operatorname{Jac}\phi_y^{-\xi h}(x)|$ is required to compute the acceptance probability $A(x,y)$, and in this case it is equal to 1, as stated by the following lemma.

Lemma 5.11. *For all $x, y \in \mathbb{R}^d$, and for all $h > 0$,*

$$|\operatorname{Jac}\phi_x^{\xi h}(y)| = |\operatorname{Jac}\phi_y^{-\xi h}(x)|$$

Proof of Lemma 5.11. This result uses the more general fact that for any skew-symmetric matrix A and any symmetric matrix S, the matrices $Id + AS$ and $Id - AS$ have same determinant. This statement is equivalent to say that $\chi_{AS} = \chi_{-AS}$, where we denote by χ_M the characteristic polynomial of any square matrix M. This last statement is true since for any square matrices A and S (of the same size) $\chi_{AS} = \chi_{SA}$. Then using the transposition, $\chi_{AS} = \chi_{A^t S^t}$. Eventually using the fact that $A^t = -A$ and $S^t = S$, it comes $\chi_{AS} = \chi_{-AS}$. The result follows from the fact that the matrix $\operatorname{Jac}\Phi_x^{\xi h}(y)$ is of the form $Id + \xi AS$ with $A = J$ and $S = \nabla^2 U$. □

We denote by $\alpha_{h,\xi}^1$ and $\alpha_{h,\xi}^2$ respectively the acceptance probability for proposal kernels Q_1^ξ and Q_2^ξ (defined respectively by (5.17) and (5.20)). The following proposition provides an upper bound on the moments of the rejection probability.

Proposition 5.12. *Suppose that U is three times differentiable with bounded second and third derivatives. Then for all $l \geq 1$, there exists $C(l) > 0$ and $h_l > 0$ such that for all positive $h < h_l$, and for all $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left[(1 - \alpha_{h,\xi}^1(x, Y_{h,\xi}^1))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{2l}$$
$$\mathbb{E}\left[(1 - \alpha_{h,\xi}^2(x, Y_{h,\xi}^2))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{3l}$$

Proof of Proposition 5.12. To deal with both results at once, we define for all $x \in \mathbb{R}^d$, for all $\xi \in \{-1, 1\}$, and for all $\theta \in \{0, 1\}$, the random variable $Y_{x_\xi,\theta}$ that satisfies the following implicit equation almost surely,

$$Y_{x_\xi,\theta} = x - h\nabla U(x) - h\xi J\left(\theta\nabla U\left(\frac{x + Y_{x_\xi,\theta}}{2}\right) + (1-\theta)\nabla U(x)\right) + \sqrt{2h}\chi,$$

where χ is a standard normal deviate in $\mathbb{R}^d$. Well-posedness of $Y_{x_\xi,\theta}$ is given by Proposition 5.9. We set R_θ^ξ the associated proposal kernel. We get $R_0^\xi = Q_1^\xi$ and $R_1^\xi = Q_2^\xi$. We define the Metropolis-Hastings ratio $r(x_\xi, y_\xi)$ for proposing y_ξ from x_ξ with kernel R_θ^ξ by,

$$r(x_\xi, y_\xi) = \frac{\pi(y_{-\xi})R_\theta(y_{-\xi}, x_{-\xi})}{\pi(x_\xi)R_\theta(x_\xi, y_\xi)},$$

where we set $\chi \in \mathbb{R}^d$ such that,

$$y = x - h\nabla U(x) - h\xi J\left(\theta\nabla U\left(\frac{x+y}{2}\right) + (1-\theta)\nabla U(x)\right) + \sqrt{2h}\chi,$$

Then, a straightforward computation gives,

$$\begin{aligned}
\log(r(x_\xi, y_\xi)) =& U(x) - U(y) + \langle y - x, \nabla U(x)\rangle \\
&+ \frac{1}{2}\langle y - x, \nabla U(y) - \nabla U(x)\rangle \\
&- \frac{\xi}{2}(1-\theta)\langle y - x, J(\nabla U(y) - \nabla U(x))\rangle \\
&+ \frac{h}{4}\left(\|\nabla U(x)\|^2 - \|\nabla U(y)\|^2\right) \\
&+ \frac{h}{2}\xi\theta\langle J\nabla U\left(\frac{x+y}{2}\right), \nabla U(x) + \nabla U(y)\rangle \\
&+ \frac{h}{2}\xi^2\theta(1-\theta)\langle J\nabla U\left(\frac{x+y}{2}\right), J(\nabla U(x) - \nabla U(y))\rangle \\
&+ \frac{h}{4}(1-\theta)^2\xi^2\left(\|J\nabla U(x)\|^2 - \|J\nabla U(y)\|^2\right).
\end{aligned}$$

A Taylor expansion of these terms, making use of Proposition 5.9 and noticing that $\theta(1-\theta) = 0$ for $\theta \in \{0, 1\}$ leads to

$$\begin{aligned}
\log(r(x_\xi, y_\xi)) =& -\xi h(1-\theta)\langle \chi, JD^2U(x)\cdot\chi\rangle \\
&+ \mathrm{O}(h^{3/2}(\|\nabla U(x)\|^2 + \|\chi\|^2 + \|\chi\|^3)) \\
=& \mathrm{O}(h(1-\theta)\|\chi\|^2 + h^{3/2}(\|\nabla U(x)\|^2 + \|\chi\|^2 + \|\chi\|^3)).
\end{aligned}$$

Moreover,

$$\begin{aligned}
(1 - 1 \wedge r(x_\xi, y_\xi))^{2l} =& \mathrm{O}(\log(r(x_\xi, y_\xi))^{2l}) \\
=& \mathrm{O}(h^{2l}(1-\theta)^{2l}\|\chi\|^{4l} + h^{3l}(\|\nabla U(x)\|^{4l} + \|\chi\|^{4l} + \|\chi\|^{6l})),
\end{aligned}$$

and thus there exists $C(l) > 0$ such that,

$$\mathbb{E}\left[(1 - 1 \wedge r(x_\xi, Y_{x_\xi,\theta}))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{3l} + C(l)h^{2l}(1-\theta)^{2l}$$

□

The previous method is quite efficient since no computation of the Hessian of $\log\pi$ is required. Nevertheless, the method requires ∇U to be globally Lipschitz. We propose a last kernel, denoted by Q_3^ξ, that does not require this hypothesis but still requires γ to be of the form $J\nabla U$ with J a skew-symmetric matrix, and also needs the computation of the

Hessian of U. More precisely, $Q_3^\xi(x, dy)$ is the law of y, solution of

$$\left(Id + \frac{h\xi}{2} J \operatorname{Hess}(U)(x)\right)(y - x) = -h(Id + \xi J)\nabla U(x) + \sqrt{2h}\chi, \tag{5.21}$$

where χ a standard normal deviate. Then, Q_3^ξ is given by

$$Q_3^\xi(x, dy) = \frac{1}{(4\pi h)^{d/2}} \det M^\xi(x) \exp\left(\frac{1}{4h}\|M^\xi(x)(y - x) + h(Id + \xi J)\nabla U(x)\|^2\right) dy, \tag{5.22}$$

with $M^\xi(x) = Id + \dfrac{h\xi}{2} J \operatorname{Hess} U(x)$. This transition kernel offers also a rejection rate $a^3_{h,\xi}$ of order $h^{3/2}$.

Proposition 5.13. *Suppose that U is three times differentiable with bounded second and third derivatives. Then for all $l \geq 1$, there exists $C(l) > 0$ and $h_0 > 0$ such that for all positive $h < h_0$, and for all $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left[(1 - \alpha^3_{h,\xi}(x, Y^2_{h,\xi}))^{2l}\right] \leq C(l)(1 + \|\nabla U(x)\|^{4l})h^{3l}$$

Proof of Proposition 5.13. The proof involves the same arguments as in Proposition 5.12 and is left to the reader. □

Even though the kernel Q_2^ξ enables to circumvent the bad acceptance ratio of the explicit proposal given by Q_1^ξ, it raises some difficulties. First, it requires the potential U to be globally Lipschitz, which is quite restrictive. To use the GMALA method in the non globally Lipschitz case, one can resort to importance sampling (to get back to a globally Lipschitz setting), or use a globally Lipschitz truncation of the potential to build the proposition kernel. The same idea is used in MALTA [152]. This last trick might lead to high average rejection ratio in the truncated regions of the potential. Moreover, the computation of the proposed move requires the resolution of a nonlinearly implicit equation, that can be solved by a fixed point method, which is more costly than the Euler-Maruyama discretization used by MALA. Specific methods of preconditioning should be considered to accelerate the convergence of the fixed point.

It would be interesting to suggest a proposition kernel based on an explicit scheme, that would achieve a better acceptance ratio than Q_1^ξ, with a lower computational cost than Q_2^ξ, and whose asymptotic variance is closer to the one of the continuous process. The question of decreasing the Metropolis-Hastings rejection rate has been recently studied in [75]. The authors give a proposition kernel constructed from an explicit scheme, which is a correction (at order $h^{3/2}$) of the standard Euler-Maruyama proposal. Nevertheless, in our case this approach would not enable us to construct a proposition kernel based on an explicit scheme with a high average acceptance ratio. Work has also been done in this direction with the Metropolis–Hastings algorithm with delayed rejection, proposed in

[162], and more recently [9]. Yet, it is unclear whether this approach could be use efficiently in our case.

5.3.2 Convergence of GMALA

In this section, we only treat the case in which GMALA is used with Q_2^ξ. To obtain a central limit theorem for the Markov chain built with GMALA, it is convenient to prove the geometric convergence in total variation norm towards the target measure π. Because we require ∇U to be globally Lipschitz to ensure well-posedness of GMALA with transition kernel Q_2^ξ, MALA would be likely to benefit from geometric convergence in this case. Indeed, it is well-known that MALA can be geometrically ergodic when the tails of π are heavier than Gaussian, under suitable hypotheses (see Theorem 4.1 [152]). For strictly lighter tails, we cannot hope for geometric convergence (see Theorem 4.2 [152]). This part is devoted to showing that under some hypotheses, GMALA can benefit from geometric convergence.

Classically for MALA, the convergence follows from the aperiodicity and the irreducibility of the chain, since by construction the chain is positive with invariant measure π. Under these conditions, the geometric convergence can be proven by exhibiting a suitable Foster-Lyapunov function V such that,

$$\lim_{|x|\to\infty} \frac{PV(x)}{V(x)} < 1. \tag{5.23}$$

In our case, the situation is slightly different. The Markov chain built with GMALA is still aperiodic and phi-irreducible. This is a simple consequence of the surjectivity of the application $\Phi_x^{h\xi}$ in Lemma 5.10. To be able to prove a drift condition such as (5.23), it is usually required to be able to show that the acceptance probability for a proposed move starting from x_ξ does not vanish in expectation for large x, which is not always true in our setting. More precisely, we can only say that the maximum of the two average acceptance probabilities for the proposed moves starting from x_ξ and $x_{-\xi}$ does not vanishes for large x. Then one strategy could be to choose a Foster-Lyapunov function V that decreases in expectation when ξ is changed to $-\xi$ after a rejection. We present in the following an other strategy that seems to be more natural and more easily generalizable to a potential U that does not satisfy the specific hypotheses we use in this section. We show that the odd and the even subsequences of the Markov chain converge geometrically quickly to the target measure π by showing a drift condition on P^2: the transition probability kernel of the two-steps Markov chain defined by

$$P^2(x, dy) = \int_E P(x, dz)P(z, dy), \quad \forall x \in E,$$

under the following assumption.

Assumption 5.14. *We suppose that U is three times differentiable, and that,*

1. *eigenvalues of* $D^2U(x)$ *are uniformly upper bounded and lower bounded away from* 0*, for* x *outside of a ball centered in* 0*,*
2. *the product* $\left\|D^3U(x)\right\| \|\nabla U(x)\|$ *is bounded uniformly in* x*.*

Proposition 5.15. *Suppose Assumption 5.14. There exists* $s > 0$ *and* $h_0 > 0$ *such that for all* $h \leq h_0$*, for* $\xi \in \{-1, 1\}$*,*

$$\lim_{\|x\| \to +\infty} \frac{P^2 V_s(x_\xi)}{V_s(x_\xi)} = 0, \tag{5.24}$$

where the Foster-Lyapunov function V *is defined by* $V_s(x_\xi) = \exp(sU(x))$*.*

Remark 5.16. *In order to prove the drift condition (5.23), one can use the Foster-Lyapunov* $\widetilde{V}$ *defined by,*

$$\widetilde{V}_s(x_\xi) = \exp\left(sU(x) + s\frac{\xi h^2 \langle \nabla U(x), D^2U(x) J \nabla U(x)\rangle}{|\xi h^2 \langle \nabla U(x), D^2U(x) J \nabla U(x)\rangle|}\right),$$

for s *small enough. Yet, this proof is left to the reader, but uses the arguments developed in the following.*

Assumption 5.14 is not meant to be sharp. We are not striving for optimality here, but instead we aim to propose a simple criterion which is likely to be satisfied for smooth potentials.

The first step to prove Equation (5.24), is to show that the proposed move decreases the Lyapunov function V_s in expectation.

Lemma 5.17. *Under Assumption 5.8, there exists* $s_0 > 0$ *such that for all* $s < s_0$*, there exists* $C_1, C_2 > 0$ *such that for all* $h < h_0$ *given by Proposition 5.9, and for all* $x \in \mathbb{R}^d$ *large enough (depending on* h*),*

$$\mathbb{E}\left[V_s(Y_{h,\xi})\right] \leq V_s(x) e^{-s\frac{3h(1-C_1 h)}{4}\|\nabla U(x)\|^2 + sC_2},$$

where $Y_{h,\xi}$ *is defined in Proposition 5.9. Thus, for* h *small enough,*

$$\mathbb{E}\left[V_s(Y_{h,\xi})\right] \leq V_s(x) e^{-s\frac{h}{2}\|\nabla U(x)\|^2 + sC_2},$$

and for x *large enough,*

$$\mathbb{E}\left[V_s(Y_{h,\xi})\right] \leq V_s(x) e^{-s\frac{h}{4}\|\nabla U(x)\|^2}.$$

Proof of Lemma 5.17. For all $y \in \mathbb{R}^d$, we set $\chi \in \mathbb{R}^d$ such that,

$$y = x - h\nabla U(x) - h\xi J\left(\theta \nabla U\left(\frac{x+y}{2}\right) + (1-\theta)\nabla U(x)\right) + \sqrt{2h}\chi,$$

A Taylor expansion of y in h yields,

$$\begin{aligned} U(y) &= U(x) + \nabla U(x) \cdot (y-x) + \mathrm{O}(\|y-x\|^2) \\ &= U(x) - h\|\nabla U(x)\|^2 + \sqrt{2h}\nabla U(x)\cdot\chi + h\xi\nabla U(x)\cdot J\nabla U(\frac{x+y}{2}) + \mathrm{O}(\|x-y\|^2). \end{aligned}$$

Noticing that

$$h\xi\nabla U(x)\cdot J\nabla U\left(\frac{x+y}{2}\right) = h\xi\nabla U(x)\cdot J\left(\nabla U\left(\frac{x+y}{2}\right) - \nabla U(x)\right),$$

the Cauchy-Schwarz, the triangular inequalities and the fact that ∇U is globally Lipschitz yield,

$$\left|h\xi\nabla U(x)\cdot J\nabla U(\frac{x+y}{2})\right| \leq Ch\|\nabla U(x)\|\|x-y\|.$$

Thus, by Young inequality and Proposition 5.9, for $h \leq 1$,

$$U(y) \leq U(x) - \frac{3h(1-Ch)}{4}\|\nabla U(x)\|^2 + \mathrm{O}(\|\chi\|^2),$$

The conclusion holds for small enough s such that $e^{s\mathrm{O}(\|G\|^2)}$ is integrable, where G is a standard normal deviate in $\mathbb{R}^d$. □

Classically ([152]), proofs of geometric convergence of MALA require to show that the average acceptance ratio of the proposed move from x, does not vanishes when x is large. Namely that there exists $\varepsilon > 0$ such that

$$I(x) = \{y : \alpha(x,y) \leq 0\}$$

asymptotically has q-measure 0, where α is the acceptance probability. Our case is sightly different since it is possible that the average acceptance ratio of the proposed move from $x_\xi \in E$ vanishes when $\|x\| \to +\infty$, which leads to a rejection and to the switching $\xi \leftarrow -\xi$ with probability close to one. Yet the average acceptance probability of the next proposed move from $x_{-\xi}$ is close to one. This behavior is described by the following lemma.

Lemma 5.18. *Under Assumption 5.14, there exists $h_0 > 0$ such that for all $h < h_0$ and for all $\varepsilon > 0$, there exists $C(h,\varepsilon) > 0$ such that for all $x_\xi \in E$ satisfying $\|x\| \geq C(h,\varepsilon)$,*

$$\xi\langle\nabla U(x), D^2U(x)\cdot J\nabla U(x)\rangle \geq 0 \Longrightarrow \mathbb{P}\left(\alpha^2_{h,\xi}(x_\xi, y_\xi) = 1\right) \geq 1-\varepsilon,$$

where y is defined by Equation (5.18).

Proof of Lemma 5.18. We recall that $\alpha^2_{h,\xi}$ is defined for all $x_\xi, y_\xi \in E$ by $\alpha^2_{h,\xi}(x_\xi, y_\xi) =$

$1 \wedge e^{r(x_\xi, y_\xi)}$, with,

$$r(x_\xi, y_\xi) = U(x) - U(y) + \langle \sqrt{2h}\chi, \nabla U(x)\rangle + \frac{1}{2}\langle \sqrt{2h}\chi, \nabla U(y) - \nabla U(x)\rangle \\ - h\left\|\nabla U(x)\right\|^2 - h\langle \nabla U(y) - \nabla U(x), \nabla U(x)\rangle - \frac{h}{4}\left\|\nabla U(x) - \nabla U(y)\right\|^2.$$

The proof is done by computing a Taylor expansion in h of this quantity to evaluate its sign in the asymptotic case $\|x\| \to +\infty$. We denote by $\cdot$ the matrix vector product, and by $:$ and $\vdots$ respectively the double and triple dot products.

$$\begin{aligned} &U(x) - U(y) + \langle \sqrt{2h}\chi, \nabla U(x)\rangle - h\left\|\nabla U(x)\right\|^2 \\ &\quad = h\xi\nabla U(x)\cdot J\left(\nabla U\left(\frac{x+y}{2}\right) - \nabla U(x)\right) - \frac{1}{2}D^2U(x) : (y-x)^2 + \mathrm{O}(\|D^3U(x)\vdots(y-x)^3\|) \\ &\quad = \frac{h}{2}\xi\nabla U(x)\cdot J\left(D^2U(x)\cdot(y-x) + \mathrm{O}(D^3U(x):(y-x)^2)\right) \\ &\quad - \frac{1}{2}D^2U(x):(y-x)^2 + \mathrm{O}(D^3U(x)\vdots(y-x)^3), \end{aligned}$$

$$\begin{aligned} \frac{1}{2}\langle \sqrt{2h}\chi, \nabla U(y) - \nabla U(x)\rangle &= \frac{1}{2}\langle \sqrt{2h}\chi, D^2U(x)\cdot(y-x) + \mathrm{O}(D^3U(x):(y-x)^2)\rangle, \\ -h\langle \nabla U(y) - \nabla U(x), \nabla U(x)\rangle &= -h\langle D^2U(x)\cdot(y-x) + \mathrm{O}(D^3U(x):(y-x)^2), \nabla U(x)\rangle, \\ -\frac{h}{4}\left\|\nabla U(x) - \nabla U(y)\right\|^2 &= h\mathrm{O}(\|x-y\|^2). \end{aligned}$$

The rests above are independent of h. Collecting these terms, and using Proposition 5.9 and Assumption 5.14, it comes for $h \leq 1$

$$r(x_\xi, y_\xi) = -\frac{h}{2}\langle \nabla U(x), D^2U(x)\cdot(y-x)\rangle + \mathrm{O}(h^3\left\|\nabla U(x)\right\|^2) + \mathrm{O}(\|\chi\|^2 + \|\chi\|^3),$$

and eventually,

$$r(x_\xi, y_\xi) = \frac{h^2}{2}\left(\langle \nabla U(x), D^2U(x)\cdot\nabla U(x)\rangle + \xi\langle \nabla U(x), D^2U(x)\cdot J\nabla U(x)\rangle\right) \\ + \mathrm{O}(h^3\left\|\nabla U(x)\right\|^2) + \mathrm{O}(\|\chi\|^2 + \|\chi\|^3),$$

Then, using the definite positiveness of $D^2U(x)$ and the fact that $\|\nabla U(x)\|$ is coercive, for all h small enough and for all $\varepsilon > 0$, there exists $C(h,\varepsilon) > 0$ such that for all $x \in \mathbb{R}^d$ satisfying $\|x\| \geq C(h,\varepsilon)$,

$$\xi\langle \nabla U(x), D^2U(x)\cdot J\nabla U(x)\rangle \geq 0 \Longrightarrow \mathbb{P}\left(r(x_\xi, y_\xi) \geq 0\right) \geq 1 - \varepsilon.$$

□

Proposition 5.15 can now be obtained as a consequence of Lemmas 5.18 and 5.17.

Proof of Proposition 5.15. Because of the acceptance-rejection step, we get for all $x_\xi \in E$,

$$\begin{aligned}
P^2V(x) = &\int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi) \int_{\mathbb{R}^d} q(y_\xi, z_\xi)\alpha^2_{h,\xi}(y_\xi, z_\xi)V(z_\xi)dzdy \\
&+ \int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi)V(y_{-\xi}) \left(1 - \int_{\mathbb{R}^d} q(y_\xi, z_\xi)\alpha^2_{h,\xi}(y_\xi, z_\xi)dz\right) dy \\
&+ \left(1 - \int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi)dy\right) \left(\int_{\mathbb{R}^d} q(x_{-\xi}, y_{-\xi})\alpha^2_{h,\xi}(x_{-\xi}, y_{-\xi})V(y_{-\xi})dy\right) \\
&+ V(x_{-\xi}) \left(1 - \int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi)dy\right) \left(1 - \int_{\mathbb{R}^d} q(x_{-\xi}, y_{-\xi})\alpha^2_{h,\xi}(x_{-\xi}, y_{-\xi})dy\right).
\end{aligned}$$

Simply using $\alpha^2_{h,\xi} \leq 1$ leads to,

$$\begin{aligned}
P^2V(x) \leq &\int_{\mathbb{R}^d} q(x_\xi, y_\xi) \int_{\mathbb{R}^d} q(y_\xi, z_\xi)V(z_\xi)dzdy \\
&+ \int_{\mathbb{R}^d} q(x_\xi, y_\xi)V(y_{-\xi}) \left(1 - \int_{\mathbb{R}^d} q(y_\xi, z_\xi)\alpha^2_{h,\xi}(y_\xi, z_\xi)dz\right) dy \\
&+ \left(1 - \int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi)dy\right) \left(\int_{\mathbb{R}^d} q(x_{-\xi}, y_{-\xi})V(y_{-\xi})dy\right) \\
&+ V(x_{-\xi}) \left(1 - \int_{\mathbb{R}^d} q(x_\xi, y_\xi)\alpha^2_{h,\xi}(x_\xi, y_\xi)dy\right) \left(1 - \int_{\mathbb{R}^d} q(x_{-\xi}, y_{-\xi})\alpha^2_{h,\xi}(x_{-\xi}, y_{-\xi})dy\right).
\end{aligned}$$

Lemmas 5.17 and 5.18 give,

$$\begin{aligned}
\frac{P^2V(x)}{V(x)} \leq \quad & e^{-s\frac{h}{2}\|\nabla U(x)\|^2 + 2sC_2} + e^{-s\frac{h}{2}\|\nabla U(x)\|^2 + sC_2} \\
&+ (1 - \mathbb{E}\left[\alpha^2_{h,\xi}(x_\xi, Y_{h,\xi})\right])e^{-s\frac{h}{2}\|\nabla U(x)\|^2 + sC_2} \\
&+ (1 - \mathbb{E}\left[\alpha^2_{h,\xi}(x_\xi, Y_{h,\xi})\right])(1 - \mathbb{E}\left[\alpha^2_{h,\xi}(x_{-\xi}, Y_{h,-\xi})\right]).
\end{aligned}$$

Thus,

$$\lim_{\|x\| \to +\infty} \frac{P^2V(x)}{V(x)} = 0$$

□

5.4 Generalized hybrid MALA

We propose in this part a second algorithm. It is based on a splitting method where the reversible and the non reversible parts of Equation (5.11) are alternatively solved with measure preserving schemes. The idea is then to integrate the purely reversible equation

$$dx = -\nabla U(x)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{5.25}$$

by using MALA, and to integrate the conservative equation

$$dx = -\xi J \nabla U(x)\, \mathrm{d}t, \tag{5.26}$$

by an hybrid Monte-Carlo method based on a suitable integrator. This method presents some theoretical advantages on Generalized MALA. Before presenting them, we begin by explaining the algorithm.

We denote by Ψ_t^ξ the flow of the conservative Equation (5.26) on the time interval $[0, t]$, and by Φ_h^ξ a numerical integrator for (5.26) on a time step h. We precise later necessary conditions on this integrator that ensure unbiasedness of the algorithm.

Algorithm 5.19 (**Generalized Hybrid MALA**). *Let x_0 be an initial point. Set $\xi = \pm 1$. Let $h > 0$. Iterate on $n \geq 0$,*

1. *Integration of the reversible part (5.25):*
   MALA is used to sample $x_{n+1/2}$ from x_n, with time-step h.
2. *Integration of the non reversible part (5.26):*
   (a) *Compute $\widetilde{x}_{n+1} = \Phi_h^\xi(x_{n+1/2})$.*
   (b) *Set $x_{n+1} = \widetilde{x}_{n+1}$ with probability*

$$\begin{aligned}\beta_{h,\xi}(x_{n+1/2}) &= \min\left(1, \exp(U(x_{n+1/2}) - U(\widetilde{x}_{n+1}))\right) \\ &= \min\left(1, \exp(U(x_{n+1/2}) - U(\Phi_h^\xi(x_{n+1/2})))\right).\end{aligned}$$

Otherwise set $x_{n+1} = x_{n+1/2}$ and $\xi \leftarrow -\xi$.

Similarly to the Hybrid Monte-Carlo algorithm, the first step enables to explore the state space across the iso-potential lines, whereas the second step enables to explore it along the iso-potential lines.

To ensure unbiasedness of the second step, the integrator Φ_h^ξ must satisfy the following properties,

$$\Phi_h^\xi = (\Phi_h^{-\xi})^{-1}, \tag{5.27}$$

$$\det \nabla \Phi_h^\xi = 1. \tag{5.28}$$

These properties are classical for hybrid Monte-Carlo methods (see Chapter 2. [114]).

Lemma 5.20. *Under conditions (5.27) and (5.28), the second step leaves the measure π invariant.*

Proof of Lemma 5.20. The proof can be found in [114]. It consists in seeing this step as a generalized Metropolis-Hastings step, with proposal $Q(x_\xi, y_\eta) = \delta_{\Phi_h^\xi(x)}\delta_\xi(\eta)$, and symmetric operator $S(x_\xi) = x_{-\xi}$. Then, it is enough to show that the Metropolis-Hastings

ratio r defined by the following Radon-Nikodym derivative

$$r(x_\xi, y_\eta) = \frac{Q(S(y_\eta), S(dx_\xi))\pi(dy_\eta)}{Q(x_\xi, dy_\eta)\pi(dx_\xi)},$$

is equal to $\exp(\log\pi(y) - \log\pi(x))$. □

For example, for all $x_\xi \in E$ the centered point integrator $\Phi_h^\xi(x)$ defined by the solution y of the following equation

$$y = x - h\xi J\nabla \log\pi\left(\frac{x+y}{2}\right), \tag{5.29}$$

satisfies these properties, under some assumptions that ensure well-posedness.

Lemma 5.21. *Under Assumption 5.8, there exists $h_0 > 0$ such that for all positive $h < h_0$, there exists a unique solution to Equation 5.29, and the integrator Φ_h^ξ is well-defined and satisfies equations (5.27) and (5.28).*

Proof of Lemma 5.21. The proof follows from Lemma 5.10. □

The main benefit of this algorithm with respect to GMALA, is the better average rejection rate of the Hybrid step with the centered point integrator, which is of order $\mathrm{O}(h^3)$ instead of $\mathrm{O}(h^{3/2})$ for GMALA, which may enable to reduce the rate of switching directions.

Lemma 5.22. *Suppose Assumption 5.14. Then, for all $l \geq 1$, there exists $h_0 > 0$ such that for all positive $h < h_0$, and for all $x \in \mathbb{R}^d$,*

$$(1 - \beta_{h,\xi}(x))^{2l} \leq C\,\|\nabla U(x)\|^{2l}\,h^{6l}.$$

Proof of Lemma 5.22. For $x \in \mathbb{R}^d$, we set $y = \Phi_h^\xi(x)$. A Taylor expansion of $U(y)$ yields,

$$|U(y) - U(x)| = \mathrm{O}(h^3\left\|D^3U(x)\right\|\|\nabla U(x)\|^3).$$

The conclusion follows from the fact that

$$(1 - \beta_{h,\xi}(x))^{2l} = \mathrm{O}\left(U(y) - U(x)\right),$$

where the domination holds uniformly in h and ξ. □

The centered point integrator is an example of integrator for the Hamiltonian dynamics, that can be used as soon as the potential U is globally Lipschitz. Actually, similarly to GMALA in the non globally Lispchitz case, one can construct an approximate integrator Φ_h^ξ by using a truncation of ∇U to ensure its global Lipschitzness. Yet, this trick may lead to high rejection rates in the areas where ∇U is truncated. Again, other strategies can be used like importance sampling or even a change of variable in the integrand (5.9).

The GHMALA algorithm is especially interesting when we are able to integrate efficiently the conservative dynamics. We show later on the numerical experiments two examples of specific integrators that enable to improve significantly the performance of GHMALA with respect to GMALA with proposal kernel Q_2^ξ. We can recall for example the case of Hamiltonian dynamics with separable Hamiltonian energy that can be integrated with explicit volume preserving schemes which are time-reversible and symmetric ([144]).

The authors propose in [72] a similar splitted scheme where the hybrid step is replaced by fourth-order Runge-Kutta method. Even though this choice leads to a biased estimator, it allows to get rid of the centered point integrator, that may be more costly than the Runge-Kutta method in terms of computational time.

5.5 Numerical experiments

In this section we illustrate the theoretical results of Proposition 5.12 and Lemma 5.22 discussed above, and we compare the asymptotic variance of MALA, GMALA and GHMALA on three toy cases. All the estimations in this paragraph are computed as the average of 10^3 independent realizations, by taking the time average of 10^5 points sampled by MALA, GMALA and GHMALA. The underlying numerical scheme for GMALA and GHMALA is based on a mid-point discretisation. Because of the nonlinearity, the proposals in GMALA and in the second step of GHMALA are computed using a fixed point method. The convergence is supposed to be attained when the relative error between two iterates becomes smaller than 10^{-15}, which is just above the number of significant digits of a "double".

5.5.1 Anisotropic distribution

In this section, we test our nonreversible MALA algorithm by computing the average of an observable f with respect to a two dimensional anisotropic distribution. More precisely we want to estimate $\mathbb{E}[f(X)]$ with $f(x_1, x_2) = x_1^2 1_{x_1>15}$, and $X \sim \pi(x)$ with $\pi(x) \propto e^{-V_1(x)}$, where

$$V_1((x_1, x_2)) = \frac{x_1^2}{\sqrt{1 + 50x_1^2}} + x_2^2.$$

Such a distribution is more stretched out in the x_1 direction rather than the x_2. A plot of this potential is given in the left subplot of Figure 5.1. We expect our lifted algorithm to be more favorable than MALA when the anisotropy is strong. Indeed, MALA tends to perform a slower exploration of the state space in the x_1 direction with respect to x_2 the direction. The lifted algorithm is supposed to correct this problem since the conservative dynamics should lead to a fast exploration of the iso-potential lines, which are stretched

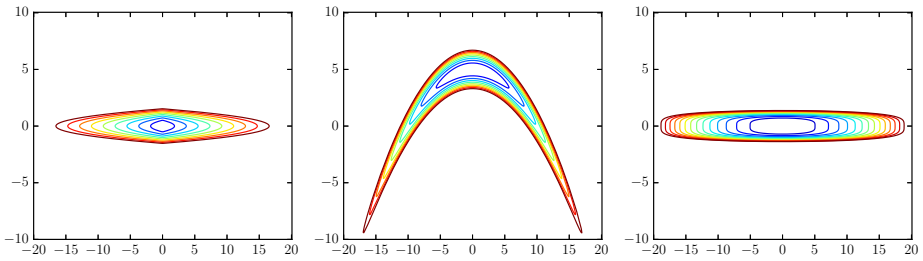

Figure 5.1 – Contour plot, from left to right, of the potentials V_1, V_2 and V_3

in the direction x_1. We choose as the skew-symmetric matrix

$$J = \alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{5.30}$$

with α a real parameter.

To compare MALA with GMALA and GHMALA, different optimal time steps parameters should be used. There is a tradeoff between achieving high average acceptance ratios (obtained with small time steps) and small correlations between the successive samples (obtained with large time-steps). More precisely MALA usually gives its best results with a large h that ensures a significant average rejection ratio. On the contrary, GMALA and GHMALA require much smaller time steps to avoid the regress that happens with the direction switching at each rejection. Thus, MALA should be used with a larger time step than the two others. Moreover, as GMALA with proposal kernel Q_1^ξ leads to a worse average acceptance ratio than Q_2^ξ, the former requires a smaller time-step than the latter. Figure 5.2 shows an estimation of the average acceptance ratio with respect to the time step h for GMALA with proposal kernels Q_1^ξ and Q_2^ξ. We can verify that the average rejection ratio of GMALA with Q_1^ξ is indeed much worse than MALA and is of order h. In practice, this limitation leads to very correlated successive samples, and thus bad asymptotic variances for the estimators based on such Markov chain. In the following, we always consider GMALA with proposal kernel Q_2^ξ instead of Q_1^ξ.

Figure 5.3 shows the rejection ratio with respect to the time step h for the three algorithms. GHMALA is composed by two steps: the first one consists of MALA and the second one consists of a Hybrid iteration. Both steps are corrected by an acceptance/rejection step. The average acceptance rate of the first step does not depend on α, and is denoted by *GHMALA MALA* in the legend. Moreover, it is the same as a plain MALA. The average acceptance rate of the second step is denoted by *GHMALA Hybrid*. We also observe that the rejection ratio for GMALA with Q_2^ξ is close to MALA's and scales as $h^{3/2}$. This is due to the fact that the non reversibility contributes at order h^2 in the expression of the average rejection ratio. We can also verify the upper bound on the rejection probability of the hybrid step, given by Lemma 5.22.

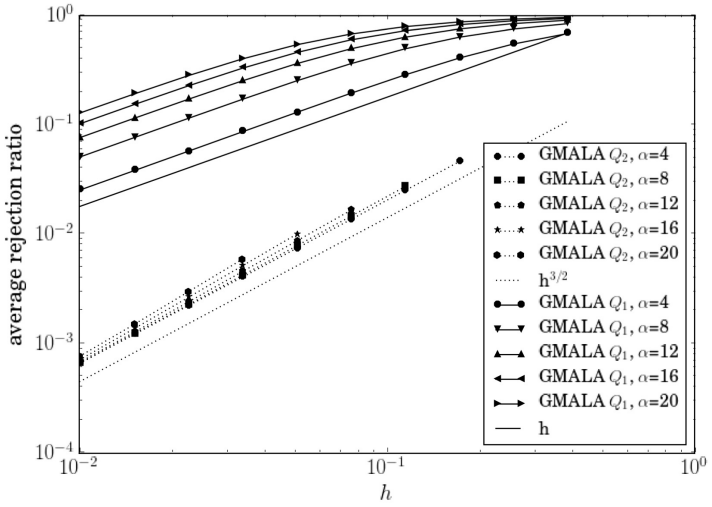

Figure 5.2 – Comparison of average rejection ratio for GMALA with proposal kernels Q_1^ξ and Q_2^ξ.

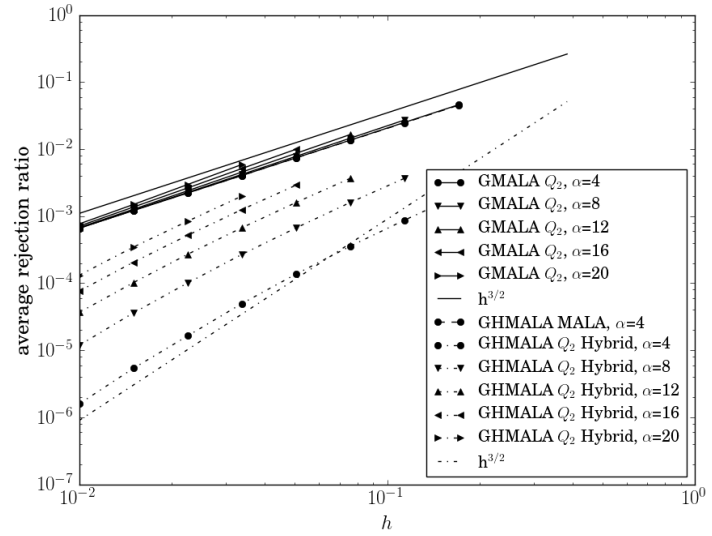

Figure 5.3 – Comparison of average rejection ratio for MALA, GMALA (with proposal kernel Q_2^ξ) and GHMALA.

We now compare the asymptotic variance of the estimators build with MALA, GMALA and GHMALA. We plot in Figure 5.4 the empirical relative variance of these estimators, We observe that GMALA performs much better than MALA by a factor 20. We should make precise that a reduction in the variance does not necessarily mean a reduction in computing time since one iteration of GMALA is more costly than MALA as it requires a fixed point iteration.

We present in Figure 5.5 the average number of Picard iterations for GMALA (with

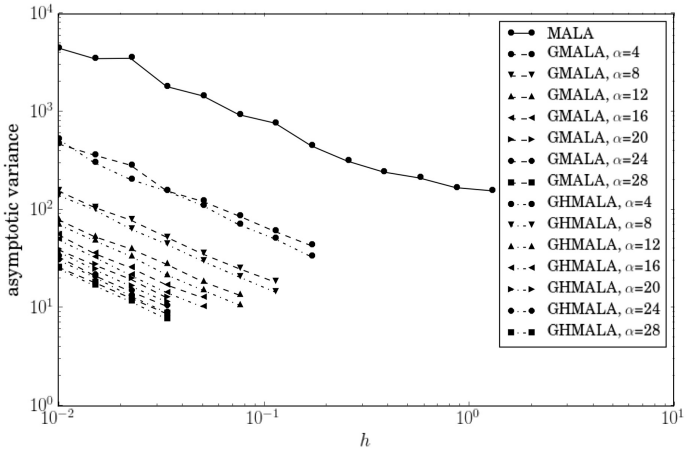

Figure 5.4 – Variance comparison of MALA, GMALA (Q_2^ξ) and GHMALA on the anisotropic distribution

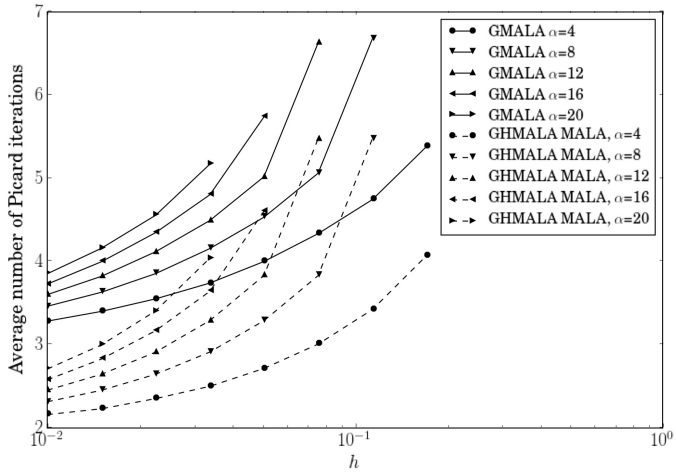

Figure 5.5 – Comparison of average number of Picard iterations for GMALA (with proposal kernel Q_2^ξ) and the second step of GHMALA.

proposal kernel Q_2^ξ) and the second step of GHMALA. We can observe that the number of Picard iterations seems lower for GHMALA that GMALA. This is no surprising since the proposals for the hybrid step of GHMALA are at a distance of order δt, instead of $\sqrt{\delta t}$ for GMALA, from the current point. Thus, we expect somehow smaller variations of the gradient of V_1 between the proposal and the current point, which could explain a faster convergence.

The question of choosing α is quite natural. In the case of the time-continuous process, it is known that the asymptotic variance decreases as the intensity α increases, which suggests to use algorithms with large α [150, 72, 98]. Yet, the well-posedness of the proposal kernel Q_2^ξ requires the product αh to be small enough to ensure the convergence of the

fixed point. From a computational point of view, the cost of the method is proportional to the number of Picard iterations, which scales as $-\log \rho$, where ρ denotes the contraction ratio (that scales as αh). Then, the choice of parameters α and h should take into account these two effects.

5.5.2 Warped Gaussian distribution

This example deals with a potential growing faster than quadratically. Both GMALA and GHMALA can be adapted to this case. The simple idea is to build a proposal kernel with a truncation of ∇U, to make it globally Lipschitz. This is the same idea as the one used for MALTA ([152]). Moreover, it is also possible to choose a more efficient integrator than the centered point integrator defined by Equation (5.29).

To illustrate these two methods, we test now our algorithms in the case of a two-dimensional warped Gaussian distribution. This toy case has been introduced in [90] and used as a benchmark for variance reduction methods based on nonreversible Langevin samplers in [72]. More precisely, we aim at estimating $\mathbb{E}[f(X)]$ with the observable f and X distributed according to the measure π, defined by,

$$f(x) = \mathbb{1}_{\{x_1 \geq 0\}} x_1^2, \quad \pi(x) \propto e^{-V_2(x)}, \quad \text{with} \quad V_2(x) = \frac{x_1^2}{100} + \left(x_2 + \frac{x_1^2}{20} - 5\right)^2.$$

A plot of this potential is given in the middle subplot of Figure 5.1. From this figure, we can expect a slow convergence of MALA since the potential is quite elongated alongside the two branches of the warped Gaussian distribution. Indeed, we expect the GMALA and the GHMALA algorithms to perform better than MALA by switching more frequently between the two branches of the potential. The choice of the observable f should reflect this behaviour by leading to a much lower asymptotic variance for the nonreversible algorithms. We define the skew symmetric matrix J by (5.30). It appeared in [72] that the nonreversible Langevin dynamics enables to reduce the asymptotic variance by several orders of magnitude.

We propose to implement GMALA using a truncated drift to make it globally Lipschitz to ensure the well-posedness of the method, and to implement GHMALA using a specific integrator of the conservative dynamics. We show that GHMALA performs better than GMALA and MALA in this case. More precisely, the integrator is defined as the centered point integrator, after the symplectic change of variable ψ defined for all $(x_1, x_2) \in \mathbb{R}^2$ by

$$\psi(x_1, x_2) = \left(x_1, x_2 + \frac{x_1^2}{20} - 5\right).$$

Typically, this integrator enables to solve this dynamics for time steps larger than GMALA with proposal kernel Q_2^ξ, and thus reduces the asymptotic variance. Figure 5.6 displays the asymptotic variance for the estimators build with these algorithms. A plot of this potential is given in Figure 5.1.

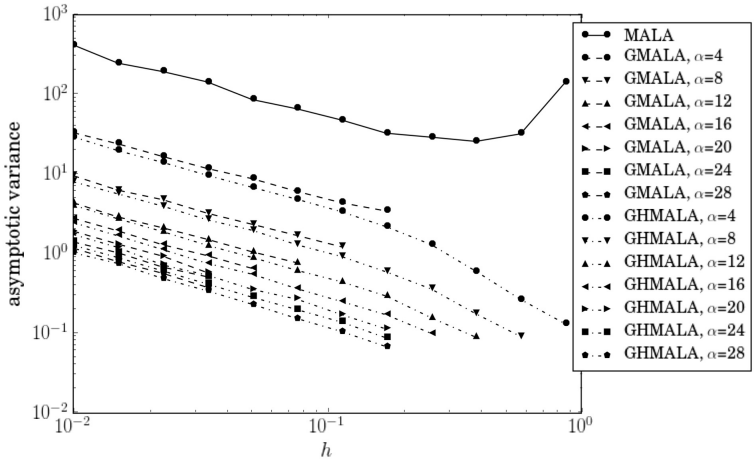

Figure 5.6 – Variance comparison of MALA, GMALA and GHMALA on the warped Gaussian distribution

We can observe that GMALA and GHMALA performs similarly for small time steps h. Yet, for larger time-steps, it is not possible to define the proposal kernel for GMALA, whereas it is still the case for GHMALA. Eventually, we achieve a variance reduction of about a factor 500 with GHMALA and 60 with GMALA, compared to classical MALA.

5.5.3 Quartic Gaussian distribution

This toy case aims to present a particular case where GHMALA can be used without implicit integrator, and in a non globally Lipschitz case, which may enable to reduce the computational cost by avoiding the fixed point iteration. Again, we aim at estimating $\mathbb{E}[f(X)]$ where X is distributed according to the measure π, and where,

$$f(x) = x_1^2 + x_2^2, \quad \pi(x) \propto e^{-V_3(x)}, \quad \text{with} \quad V_3(x) = \frac{x_1^2}{100} + x_2^4. \tag{5.31}$$

A contour plot of this potential is given in the right subplot of Figure 5.1. We define the skew-symmetric matrix J by (5.30). In this case, the Hamiltonian dynamics defines then a separable system and volume-preserving explicit methods can be used (see [144]). More precisely, we define Φ_h^ξ for all $x = (x_1, x_2) \in \mathbb{R}^2$ by $\Phi_h^\xi(x) = (y_1, y_2)$, where,

$$\begin{aligned} y_1^{1/2} &= x_1 - \frac{h}{2}\alpha\xi\frac{\partial V_3}{\partial x_1}(x) \\ y_2 &= x_2 + h\alpha\xi\frac{\partial V_3}{\partial x_2}\left(y_1^{1/2}, x_2\right) \\ y_1 &= y_1^{1/2} - \frac{h}{2}\alpha\xi\frac{\partial V_3}{\partial x_1}\left(y_1^{1/2}, y_2\right) \end{aligned}$$

Thus, the non Lipschitz nonlinearity of ∇V is not an obstacle to the well-posedness of this integrator. We plot in Figure 5.7 the asymptotic variance for the time average estimators

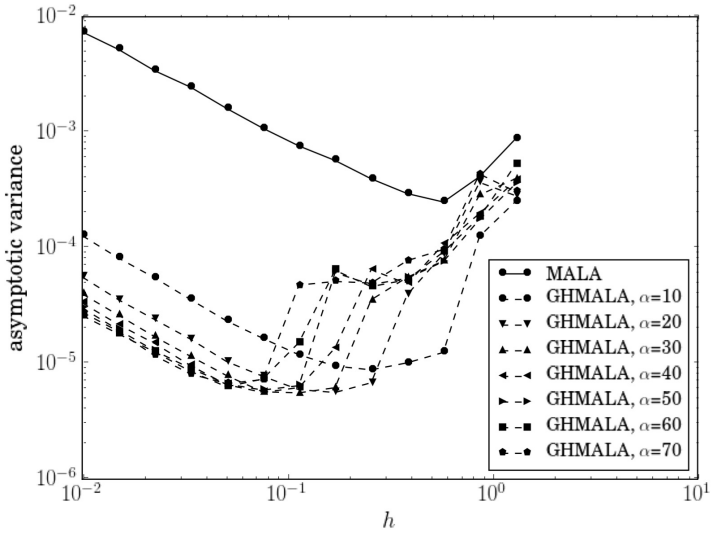

Figure 5.7 – Variance comparison of MALA and GHMALA on the quartic Gaussian distribution

of $\mathbb{E}[f(X)]$ build with GHMALA (where f is given by Equation (5.31)). We can observe that the decrease in variance between MALA and GHMALA for small h is around 280. Nevertheless, for larger h, the explicit integration is not accurate enough, which leads to an increase in the asymptotic variance for larger time steps. Eventually, the smallest asymptotic variance of the time average estimator of GHMALA is around 50 times lower than the smallest asymptotic variance for MALA.

5.6 Conclusion

We presented a class of unbiased algorithm that enables us to benefit from the variance reduction of the nonreversible Langevin equations (5.1) with respect to the reversible dynamics (5.10). More precisely, we presented two variations of these algorithms. The first one (GMALA) can be viewed as a lifting method, and more specifically as a generalized Metropolis Hastings methods on a lifted state space. The second one (GHMALA), similar to the first one, can be viewed as a Generalized Hybrid Monte-Carlo method.

Numerical experiments show that variance reductions (compared with classical MALA) of several orders of magnitude can be achieved for potentials concentrated on a lower dimensional submanifold. We also expect these algorithms to perform better in the case of entropic barriers. The main difficulty is, in the case of GMALA, to use a proposal that allows to achieve a sufficiently high average acceptance ratio (to compete with MALA). For example this can be done by using a mid-point discretization. Even though this scheme is implicit, the computation of the Metropolis-Hastings acceptance probability does not require the computation of the Hessian of $\log \pi$. In the case of GHMALA, numerical experiments show that the choice of a suitable integrator for the conservative dynamics

may lead to large improvements and computational cost reduction compared with the mid-point method.

Chapitre 6

Numerical analysis of metastable dynamics in rotating BEC

6.1 Introduction and motivations

This chapter aims at providing a numerical method to analyse the metastable behaviour of a slowly rotating Bose-Einstein condensate at finite temperature, modelled by the Stochastic Projected Gross-Pitaevskii Equation.

6.1.1 The physical setting

It has been experimentally observed that rotating Bose-Einstein condensates take the form of vortex lattices, with a number of vortices that depends on the rotating speed of the thermal cloud [3, 30, 34, 122]. We are interested in the situation where several vortex lattices configurations, characterised by different numbers of vortices, are locally stable for a same fixed rotation speed. We aim at investigating metastable dynamics between these vortex configurations in this case. We make precise this notion of metastability in Section 6.1.2.

In the following, we consider the case of a rotating dilute Bose gas trapped into an isotropic harmonic trap centred at the origin O. Moreover, we suppose that the thermal cloud rotates around the $O\overrightarrow{z}$ axis. We also assume that the intensity of the trap in the z direction is strong, so that the condensate is flattened enough to be described by a two dimensional model. We refer to the Section 2.4.2 in Chapter 2 for precisions about this simplification. Thus, we limit the present study to the two dimensional case. Because of the isotropy of the system, its energy is invariant by rotations around the $O\overrightarrow{z}$ axis. Moreover, the energy is also invariant by any uniform change of phase. In the following we will say that two states of the system are *equivalent* is they are equal up to a rotation around the $O\overrightarrow{z}$ axis and a uniform shift of the phase.

In order to study this kind of dynamics in the isotropic case, we will model the system

by the Stochastic Projected Gross-Pitaevskii Equation (SPGPE). See [30, 80, 81] for a derivation of this model, and [25] for a review of similar techniques and models. This model aims at describing the dynamics of (possibly rotating) Bose-Einstein condensate at finite temperature. It has mainly been built to provide a tool to analyse the growth dynamics of a condensate during the phase transition and especially the spontaneous symmetry breaking and the nucleation of vortices during this process [111, 166, 167]. Yet, when the Gross-Pitaevskii energy at zero temperature (defined later in (6.5)) has several different local minima which are not equivalent, the solution of the SPGPE exhibits a metastable behaviour. From a mathematical point of view, the vortices correspond to singularity points of the phase field of the wave function, but not of the complex field itself. Thus the density of the condensate vanishes at these points. They create holes inside the condensate and thus they can be observed during experiments. Note that it is also possible to observe some phase differences in practice. This kind of methods can provide experimental vortex localisation methods.

To our knowledge, there is no experimental evidence of such a metastable behaviour. We recall that during experimentations the condensate is generally destroyed, during an expansion procedure, to make the vortices observable. Thus, most of the experimentations only give access to the law of an observable at a given time, but not to trajectorial observables. In this chapter we do not pretend to highlight a metastable behaviour of the considered system. Indeed, we only propose a numerical method to analyse the metastable dynamics of the solutions of the SPGPE with rotating thermal cloud, without judging if this model is valid and relevant to predict such a physical phenomenon. Nevertheless we note that in [119] the authors try to form condensates, tuning the speed of rotation to prescribe the number of vortices. These configurations are shown in Figure 6.1. They observe that a prescribed rotation velocity may lead to different numbers of vortices in a condensate. They cannot identify if these fluctuations come from a lack of experimental reproducibility, or if they are intrinsic to the system considered. Our work proposes an explanation, by means of a metastable behaviour, for this observation. More recently in [171] the authors study experimentally Bose-Fermi superfluid mixtures. The number of counted vortices fluctuates quite a lot with respect to the rotation frequency. These phenomena could be caused by metastable dynamics between vortex configurations. Metastable behaviours should also occur in rotating fermionic superfluids. Zwierlein and co-authors report observations of vortex lattices in such a superfluid in [173]. Numerical experiments in the context of rotating BEC at finite temperature have been carried out in [100]. The model is not based on the SPGPE, but on the Zaremba, Nikuni, and Griffin (ZNG) approach. The authors study the dynamics of an off-centred vortex in a harmonically trapped pancake-shaped condensate. The vortex is known to decay by spiraling out toward the edge of the condensate. They quantify the dependence of this decay on temperature, atomic collisions, and thermal cloud rotation. This experiment has been done using a numerical scheme proposed in [101].

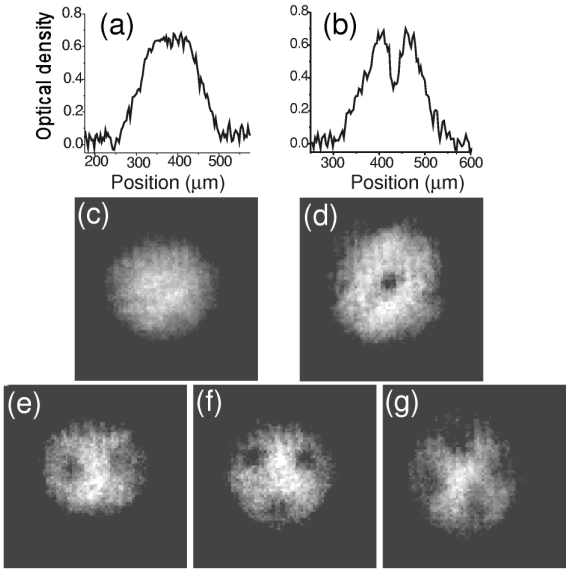

Figure 6.1 – Transverse absorption images of a Bose-Einstein condensate stirred with a laser beam for various rotation frequencies (from [119]).

6.1.2 Basics about metastability

We usually say that an ergodic dynamics is *metastable* when it can be split up between a slow and a fast time scale and that the fast dynamics consists of fluctuations inside a *metastable region* inside which the dynamics stays trapped for large times (compared with the fast time scale). Let us focus on a typical kind of metastable dynamics given by the reversible overdamped Langevin equation,

$$dX_t^\varepsilon = -\nabla E(X_t^\varepsilon)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t, \tag{6.1}$$

where the energy landscape E presents some *potential* or some *entropic* barriers. A potential barrier appears when the energy E has several local minima separated by a saddle-point of higher energy than both local minima. It means that any continuous trajectory that links two local minima needs to cross some level sets of higher energy than this saddle point. The scarcity of a change of metastable region comes from the fact that crossing the barrier requires some unlikely moves of the stochastic fluctuations that enable to provide enough energy to the system to cross the potential barrier. An entropic barrier appears when several regions of the energy are separated by a narrow continuous path of constant energy. In this case, the rare event comes from the fact that the system must find the good path to cross metastable regions. We refer to Figure 1.1 [115] for an illustration for these two kinds of energy landscape. In our application case, the metastability of the condensate dynamics is caused by a potential barrier.

In the simplifying case of a dynamics given by (6.1), we can formally define a metastable state as a state such that the typical exit time from it is much larger than the local

equilibration time inside the state. One convenient way to give a rigorous sense to this idea is to use the notion of quasi-stationary distribution which provides a way to define a metastable state with respect to the first eigenvalues of the generator of the diffusion (6.1). We refer to [46, 64, 115] for further preferences.

The interesting long time behaviour of a metastable dynamics relies on the exit events of the metastable regions, defined by the exit times and possibly the exit points. Since the system remains trapped a long time inside the metastable region, it is natural to hope that the exit event can be modeled as a jump Markov process as explained in [64]. These exit times highly depend on the kind of barrier as stated above. In the small temperature limit ($\varepsilon \to 0$), they grow linearly with respect to ε^{-1} in the case of an entropic barrier, and exponentially with respect to ε^{-1} in the case of a potential barrier. This last result is known as the Arrhenius law, and can be formulated and proven with the theory of large deviations [62, 78] under appropriate assumptions on E. In the case of potential barriers for the reversible overdamped Langevin dynamics (6.1), some works have been carried out to estimate the prefactors of the average exit time of a metastable state. See for instance [93]. This kind of results are known as Eyring-Kramers formulas. Suppose that that E describes a double-well energy in an arbitrary space dimension. Let x_- and x_+ be the two minima of each well. Let x^* be the only saddle point. Then, we denote by D a neighbourhood of x_+ of size δ, chosen small enough, and by τ_D the hitting time of the set D. Then, the Eyring-Kramers formula is given by,

$$\mathbb{E}_{x_-}(\tau_D) = \frac{2\pi}{\lambda_1(x^*)}\sqrt{\frac{\left|\det(\nabla^2 E(x^*))\right|}{\det(\nabla^2 E(x_-))}}e^{(E(x_-)-E(x^*))/\varepsilon}\left[1+\mathcal{O}\left(\varepsilon^{1/2}\left|\log\varepsilon\right|^{3/2}\right)\right], \tag{6.2}$$

where $\lambda_1(x^*)$ is the unique negative eigenvalue of $\nabla^2 E(x^*)$, and $\mathbb{E}_{x_-}(\tau_D)$ is the expectation of τ_D for a Markov process given by (6.1) and starting from x_-.

The SPGPE model is given by a nonreversible overdamped Langevin equation, of the form,

$$dX_t^\varepsilon = -(\varepsilon + i)\nabla E(X_t^\varepsilon)\,\mathrm{d}t + \sqrt{2}\varepsilon\,\mathrm{d}W_t. \tag{6.3}$$

This dynamics exhibits a metastable behaviour similar to the one of (6.1). The additional term corresponds to a purely Hamiltonian dynamics that conserves the energy E, and thus only the reversible overdamped Langevin component of this dynamics allows to cross saddle points. We refer to [29] for a generalization of the Eyring-Kramers formula to the nonreversible case.

In practice, the direct computation of the law of the exit times is a difficult problem. The Eyring-Kramers formula is only asymptotic, and requires to look for the order one saddle points of the energy (see Section 2.1 of [61] for a presentation of such methods). Efficient Monte Carlo methods have been developed to compute these exit times, without using the Eyring-Kramers formula. They consist in sampling the ensemble of paths joining

two metastable regions, without necessarily requiring any guess on them. See Chapter 5 of [61] for a review of transition path sampling algorithm in the context of molecular dynamics. In this chapter, we make use of such a method: the Adaptive Multilevel Splitting (AMS), presented in Section 6.4. See [33, 39, 40] for a mathematical background.

6.1.3 Objectives

We are interested in this chapter in providing numerical methods allowing to numerically analyse the transitions between the metastable states obtained in the experiment of [119], which are characterized by the small number of vortices located inside the condensate, as shown in Figure 6.1. To achieve this goal, we first propose in Section 6.3 a numerical scheme to solve the SPGPE model. Then we use in Section 6.4 the Adaptive Multilevel Splitting (AMS) algorithm to sample *reactive trajectories* and estimate some corresponding quantities such as the average exit time out of a basin of attraction. We refer to Section 6.4 for a precise definition of a reactive trajectory, but it should be understood as a trajectory for which the system undergoes a rare transition. In our case it corresponds to the entrance or the exit of a vortex inside the condensate. This algorithm is a Monte Carlo method to efficiently sample these trajectories and provides unbiased estimators that are functions of them. The algorithm depends on a *reaction coordinate* parametrizing the progress of the rare transitions by a real number. We propose to build this parametrization based on the positions of the vortices only. Eventually, we make use of the AMS algorithm to estimate the average transition times between two configurations. In other words, we estimate the average duration that the system stays in a configuration before changing to another one.

6.2 The Stochastic Projected Gross-Pitaevskii Equation

In this chapter, we model the dynamics of a Bose-Einstein condensate at finite temperature using the Stochastic Projected Gross-Pitaevskii Equation. This model is described in detail in Section 1.3 of the introduction. We briefly recall first the physical motivation of this model, and then its formulation.

Diluted ultra cold Bose gases are well described by the Gross-Pitaevskii Equation (GPE) (see [63]). This model is based on the assumption that the system is constituted of a large number of bosons, and that they are well-represented by a single condensate wave function. Nevertheless, the experiments are often conducted at temperatures not low enough to prevent spontaneous and incoherent processes to occur. Under those conditions the GPE may not accurately describe the system. This is especially true when the temperature approaches the condensation temperature T_c, which leads to the presence of a sizeable thermal cloud that interacts with the condensate. It has been experimentally shown that the GPE, that neglecting these interactions, may not accurately describe the condensate when the temperature becomes larger than $0.6T_c$ [95, 103]. To overcome this limitation, theoretical works have been devoted to describe these effects [80, 81, 159]. In this work,

we focus on the model developed in [81], and later enriched with [30] to take into account a rotation of the system. This model, called the Stochastic Projected Gross-Pitaevskii Equation, describes the time evolution of the condensed phase at finite temperature. This condensed phase behaves like an open system in interaction with a non-condensed phase supposed to be at thermal equilibrium.

The condensed phase is constituted (by definition) by the bosons on the lowest energy levels. Furthermore, we suppose that the condensate is flattened enough to use a two dimensional model (see the dimension reduction procedure in Section 2.4.2). Thus, the wave function of this system belongs to a finite-dimensional linear subspace of $L^2(\mathbb{R}^2, \mathbb{C})$, spanned by these lowest modes, and denoted by K. We make its definition precise later on. Eventually, the model is given by,

$$d\phi_t = -(\varepsilon + i)\nabla E^K(\phi_t)\,\mathrm{d}t + \varepsilon\sqrt{2\gamma}\,\mathrm{d}W_t, \tag{6.4}$$

where the process $(\phi_t)_{t\geq 0}$ is defined on a probability space $(\Gamma, \mathcal{F}, \mathbb{P})$ and takes values in K. The process $(W_t)_{t\geq 0}$ is a Brownian motion taking values in K. Before making precise the meaning of ∇E^K, we introduce the Gross-Pitaevskii energy E of the system at zero temperature.

$$E(\phi) = \frac{1}{2}\int_{\mathbb{R}^2} \overline{\phi}\cdot\left(\frac{1}{2}(-\Delta + |x|^2) - \mu - \Omega L_z + \frac{g}{2}|\phi|^2\right)\phi\,\mathrm{d}x \tag{6.5}$$

Then, we define E^K as the restriction of the functional E to K. We denote by ∇E^K the gradient of the E^K, and refer to Section 1.3 for precisions about the scalar fields used to define this gradient. The computation of this gradient leads to, for all $\phi \in K$,

$$\nabla E^K(\phi) = P_K\left(\frac{1}{2}(-\Delta + |x|^2) - \mu - \Omega L_z + g|\phi|^2\right)\phi,$$

where P_K denotes the orthogonal projection on K for the $L^2(\mathbb{R}^2)$-norm. To define the linear subspace K, we introduce A, the linear operator given by,

$$A = \frac{1}{2}(-\Delta + |x|^2) - \Omega L_z. \tag{6.6}$$

The eigenvectors of this operator are the Gauss-Laguerre modes given by,

$$Y_{n,l}(r,\theta) = \sqrt{\frac{n!}{\pi(n+|l|)!}}e^{il\theta}r^{|l|}e^{-r^2/2}L_n^{|l|}(r^2), \tag{6.7}$$

where $L_n^m(x)$ are the generalized Laguerre polynomials [14, 157]. The associated eigenvalues $\omega_{n,l}$ are given by,

$$AY_{n,l} = \omega_{n,l}Y_{n,l}, \quad \omega_{n,l} = 2n + |l| - \Omega l + 1.$$

The mode set K is then defined as the vector space spanned by the modes $Y_{n,l}$ such that $\omega_{n,l} \leq \omega_{\max} + 1$:

$$K = \mathrm{Vect}\left\{Y_{n,l}, (n,l) \in K_{\mathrm{indices}}\right\} \quad \text{with} \quad K_{\mathrm{indices}} = \left\{(n,l), \omega_{n,l} \leq \omega_{\max} + 1\right\}.$$

We refer to Section 2.4.1 for a dimensional definition of Equation (6.5), and further precisions about the choice of the set K.

6.3 Numerical scheme for the SPGPE

This section is devoted to the construction of a numerical scheme for Equation (6.4) that will be used to analyse its metastable properties in the small temperature limit. In this limit, characteristic times of occurrence of metastable effects can by extremely large, and numerically out of reach. Nevertheless, we present in Section 6.4 some dedicated algorithms to sample these rare occurrences.

In the small temperature limit, metastable behaviours such as an exit of some attractive region may occur when the system's energy becomes greater than the energy of some saddle point that separate this attractive region from another one. Since the nonreversible part of the Langevin equation is Hamiltonian, and thus conserves the energy, the metastable effects are thus related to the reversible Langevin part of Equation (6.4).

Many works have been devoted to the numerical integration of the deterministic Gross-Pitaevskii equation in real and imaginary times. See [7] for a review and comparisons of many techniques. In the stochastic context, numerical simulations of the SPGPE have been used (see [153] in the case without rotation, and [30] in the rotating case) by the physics community. Yet, up to our knowledge, no numerical analysis has been provided for these schemes. Moreover, being able to solve precisely the SPGPE model for dissipation intensities as low as those predicted by the model is indeed a challenge. For this reason, the numerical experiments that can be found in the physics literature use a larger dissipation intensity than what is prescribed by the model [30].

The main difference in our setting with the deterministic Gross-Pitaevskii equation comes from the projection operator P_K. Indeed, to exactly take into account this projector we propose a numerical scheme for which it can be interpreted as a spectral discretisation in space of the formal unprojected Stochastic Partial Differential Equation (SPDE),

$$d\phi_t = -(\varepsilon + i)\nabla E(\phi_t))\,\mathrm{d}t + \varepsilon\sqrt{2\gamma}\,\mathrm{d}W_t,$$

where $(W_t)_{t\geq 0}$ is a cylindrical Wiener process in $L^2(\mathbb{R}^d, \mathbb{C})$. This method amounts to solving Equation (6.4) as an ordinary SDE, whose components are the projection of $\phi(t)$ onto the elements of a well-chosen basis of the finite-dimensional vector space K. This basis is chosen to be composed by the Gauss-Laguerre modes given by Equation (6.7). Note that this discretisation has been used in [14] to solve the (purely Hamiltonian) Gross-

Pitaevskii equation with the angular momentum rotation term. The scheme presented therein uses a splitting method for the time discretisation. With such an approach, the noise discretisation relies on from classical spectral Galerkin discretisations used in the context of SPDEs approximation.

6.3.1 Definition of the numerical scheme

As we previously stated, we aim at solving Equation (6.4) in the small temperature limit, which means from a physical point of view to choose ε of order 10^{-3} to 10^{-5} (see the end of Section 2.4.3 for a justification of this order of magnitude). Thus, the dynamical system defined by (6.4) contains two time scales: a quick Hamiltonian dynamics and a slow stochastic fluctuation/dissipation dynamics. Moreover, we are interested in the long time effects of the slow dynamics on the Hamiltonian one. Thus, special care must be taken to design a numerical scheme taking into accounts both dynamics. To do so, we propose to split the system into the purely Hamiltonian part,

$$d\phi_t = -i\nabla E^K(\phi_t)\,\mathrm{d}t, \tag{6.8}$$

and the reversible stochastic part,

$$d\phi_t = -\varepsilon\nabla E^K(\phi_t)\,\mathrm{d}t + \varepsilon\sqrt{2\gamma}\,\mathrm{d}W_t. \tag{6.9}$$

Note that using the change of variable $t \leftarrow t\varepsilon$, that is to say setting $\psi(t) := \phi(\varepsilon^{-1}t)$, the process $(\psi(t))_{t\geq 0}$ is equal in law to the solution of the following dynamics,

$$d\psi_t = -\nabla E^K(\psi_t)\,\mathrm{d}t + \sqrt{2\gamma\varepsilon}\,\mathrm{d}W_t, \tag{6.10}$$

for any Brownian motion $(W_t)_{t\geq 0}$ taking values in K, since $(\varepsilon^{1/2}W_{\varepsilon^{-1}t})_{t\geq 0}$ is also a Brownian motion.

It appears that the drift over a time step h for Equation (6.9) is ε times smaller than the drift for Equation (6.8). Then, the idea is to alternate one step of resolution of Equation (6.10) with ε^{-1} steps of resolution of Equation (6.8). More precisely, for $s, t \in \mathbb{R}_+$ and $t \geq s$, we denote by $\Phi_H(t,s)$ the flow of Equation (6.8), and by $\Phi(t,s)$, $\Phi_L(t,s)$ and $\Phi_{L,\varepsilon}(t,s)$ respectively the stochastic flows of equations (6.4), (6.9) and (6.10) between times s and t [110]. Then for any time step h and any $n \in \mathbb{N}^*$, we approximate the law of $\Phi(nh,0)$ by

$$\Phi(nh,0) \approx \prod_{k=0}^{(n-1)\varepsilon} \Phi_{L,\varepsilon}((k+1)h,kh)\Phi_H(\varepsilon^{-1}(k+1)h,\varepsilon^{-1}kh). \tag{6.11}$$

The approximation (6.11) makes clear that the integration time of Equation (6.8) is ε^{-1} times longer than the integration time of Equation (6.10). This remark motivates our choice to use a more precise integration method for Equation (6.8) than for (6.10). We

propose to use a simple explicit exponential Euler scheme to approximate the flow $\Phi_{L,\varepsilon}$. Because of the additive noise, we classically expect it to be of strong order one, which is the best order we can reach without computing iterated Itô integrals. We refer to Section 6.3 [102] for a general presentation of exponential Euler schemes for SPDEs. We propose to use a symplectic Lawson method of high order to approximate the Hamiltonian dynamics. Furthermore, since it conserves the energy E, and only the Langevin dynamics enables to cross through the levels of energy, we are keen on having a Hamiltonian integrator to conserve the energy well enough between all the numerical integrations of the Langevin dynamics.

We now present the two numerical integrators to approximate the flows $\Phi_{L,\varepsilon}$ and Φ_H. We denote them respectively by $\Phi_{L,\varepsilon,h}$ and $\Phi_{H,h}$, where h stands for the time step. In each case, let T be the time horizon of integration. Let $(t_n)_{0\leq n\leq N}$ be a uniform subdivision of $[0,T]$, that is to say for all $n\in\mathbb{N}$, $n\leq N$, $t_n = nh$ with $h = T/N$ (which denotes the time step). Both of these integrators suppose that the nonlinearity $P_K\left|\phi\right|^2\phi$ can be computed exactly for all $\phi\in K$. We explain in Appendix 6.6.1 how this can be done in practice.

Numerical integration of the Langevin dynamics

We begin by presenting the numerical integrator for the Langevin dynamics (6.10). The drift $\nabla E(\psi(t))$ can be split into a linear part $-P_KA$ and a nonlinear part $\mathcal{N}(\psi(t)) = P_K\left(\mu - g\left|\psi(t)\right|^2\right)\psi(t)$. Thus, a mild formulation of Equation (6.10) is given for all $s,t>0$ with $t\geq s$ by,

$$\begin{aligned}\psi(t) &= \Phi_{L,\varepsilon}(t,s)\psi(s)\\ &= S(t-s)\psi(s) + \int_s^t S(t-\sigma)\mathcal{N}(\psi(\sigma))d\sigma + \sqrt{(2\gamma\varepsilon)}\int_s^t S(t-\sigma)P_K dW_\sigma,\end{aligned}$$

where $S(\cdot)$ denotes the semigroup of generator $-A$. The explicit exponential Euler scheme can be seen as a discretisation of this formulation. It consists in approximating the flow $\Phi_{L,\varepsilon}$ between times t_n and t_{n+1} by the discrete flow $\phi_{L,\varepsilon,h}$ defined for $n\leq N$ and any $\psi\in L^2(\mathbb{R}^2)$ by

$$\begin{aligned}\Phi_{L,\varepsilon,h}(t_{n+1},t_n)\psi =& S(h)\psi + A^{-1}(S(h)-\mathrm{Id})\mathcal{N}(\psi)\\ &+\sqrt{(2\gamma\varepsilon)}\int_{t_n}^{t_{n+1}} S(t_{n+1}-\sigma)P_K dW_\sigma.\end{aligned}$$

The first term can be trivially computed and is diagonal in the Gauss-Laguerre basis. The third term, called stochastic convolution is equal in law to a Gaussian random variable with a diagonal covariance matrix, and can be trivially simulated as well. The main difficulty consists in computing the non-linearity $P_K\left|\psi\right|^2\psi$ in the Gauss-Laguerre basis. The same difficulty appears in the numerical integration of the Hamiltonian dynamics, and is postponed to Appendix 6.6.1.

This scheme is expected to converge strongly at order 1, which is the best order we can obtain without discretising the iterated Itô integrals. The use of an exponential Euler scheme rather than a classical Euler scheme comes from classical results from the SPDE literature. In infinite dimension the exponential Euler scheme involves only bounded operators, and thus it is usually more stable and does not require a CFL condition to hold. This property is especially interesting for us since even though the problem is finite dimensional, its dimension may be large.

Integration of the Hamiltonian dynamics

The numerical integrator for the Hamiltonian dynamics (6.8) uses techniques similar to the ones used for the Langevin dynamics. The idea follows from [22], where the authors propose several numerical high order integrators for nonlinear Schrödinger equations. We propose to use a Lawson method, as described in [22, Section 3]. Yet, our scheme is slightly different, since it amounts to considering a different *spatial discretisation* that coincides with the projection P_K as explained above. Thus the main difference comes from the way we split the linear and nonlinear parts of Equation (6.8). We include the harmonic potential and the rotation operator inside the linear part, contrarily to [22]. Moreover, this discretisation seems to be of practical interest thanks to its good approximation properties with few degrees of freedom (see Remark 6.9).

The idea consists in approximating Equation (6.8) after a change of unknown that transforms this equation into a nonstiff one. More precisely, suppose that

$$u(t,\mathbf{r}) = S(t,0)^{-1}\phi(t,\mathbf{r}), \tag{6.12}$$

where S is the semigroup of generator $-iA$. Then for any $\phi_0 \in L^2(\mathbb{R}^2)$, $\phi(t,\mathbf{r})$ is solution of Equation (6.8) with initial condition $P_K\phi_0$ if and only if $u(t,\mathbf{r})$ is the solution of

$$\begin{cases} du(t) = S(t,0)^{-1}\mathcal{N}(S(t,0)u(t))\,\mathrm{d}t, \\ u(0) = P_K\phi_0, \end{cases} \tag{6.13}$$

with this time $\mathcal{N}(v) = i\left(\mu - gP_K\left|v\right|^2\right)v$ for all $v \in L^2(\mathbb{R}^2)$.

The Lawson method consists now in applying a Runge-Kutta scheme to Equation (6.13). For the sake of completeness, we reproduce the clear explanations given in [22]. We consider a s-stage Runge-Kutta method with Butcher tableau given by,

$$\begin{array}{c|ccc} c_1 & a_{1,1} & \cdots & a_{1,s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s,1} & \cdots & a_{s,s} \\ \hline & b_1 & \cdots & b_s \end{array} \tag{6.14}$$

We approximate the flow Φ_H between times t_n and t_{n+1}, *i.e.* approximate $\phi_{n+1} \in L^2(\mathbb{R}^2)$

defined by $\phi_{n+1} = \Phi_{H,h}\phi_n$ for any $\phi_n \in L^2(\mathbb{R}^2)$. To this end, we consider the sequence $(u_n)_{n\in\mathbb{N}}$, approximating the solution u of Equation (6.13), defined iteratively by,

$$u_{n+1} = u_n + \sum_{k=1}^{s} b_k h S(t_n + c_k h)^{-1}\mathcal{N}(S(t_n + c_k h)u_{n,k}),$$

where $(u_{n,k})_{1\leq k\leq s}$ is given for all $k \in \mathbb{N}^*$, $k \leq s$ by,

$$u_{n,k} = u_n + \sum_{l=1}^{s} a_{k,l} h S(t_n + c_l h)^{-1}\mathcal{N}(S(t_n + c_l h)u_{n,l}).$$

This last system of nonlinearly implicit equations can be solved by a fixed-point method under sufficient conditions given in Theorem 6.2 below. Using the Lawson change of unknown given by (6.12), this method can be written as,

$$\phi_{n+1} = S(h)\phi_n + \sum_{k=1}^{s} b_k h S((1 - c_k)h)\mathcal{N}(\phi_{n,k}), \tag{6.15}$$

where $(\phi_{n,k})_{1\leq k\leq s}$ is given for all $k \in \mathbb{N}^*$, $k \leq s$ by,

$$\phi_{n,k} = S(c_k h)\phi_n + \sum_{l=1}^{s} a_{k,l} h S((c_k - c_l)h)\mathcal{N}(\phi_{n,l}). \tag{6.16}$$

For all $\phi_n \in K$, we denote by $\Phi_{H,h}$ the approximation of the flow $\Phi_H(t + h, t)$ of Equation (6.8), for all $t \geq 0$, given by equations (6.15) and (6.16), if they are well-posed. That is to say, for all $\phi_n \in K$, we set $\Phi_{H,h}\phi_n = \phi_{n+1}$ given by Equation (6.15).

Since this numerical scheme is nonlinearly implicit, the well-posedness of $\Phi_{H,h}$ is not clear. Actually, since the non-linearity is only locally Lipschitz, as stated in Lemma 6.1, we are not able to prove its well-posedness for all time steps h, and we need them to be small enough with respect to ϕ_n. This is sufficient to show local existence and uniqueness in K by a fixed point method, for a time step h small enough, as stated in Theorem 6.2. Then, since the $L^2(\mathbb{R}^2)$-norm is conserved by the scheme, and that is a finite dimensional space, it is clear that the numerical scheme is actually globally well-posed in this case. The classical local Lipschitz property is given by the following lemma.

Lemma 6.1. *For all $\phi_1, \phi_2 \in K$, for all $k > 1(= d/2)$ there exists $C(k) > 0$ (independent of the cutoff K) such that,*

$$\|\mathcal{N}(\phi_1) - \mathcal{N}(\phi_2)\|_{\Sigma^k} \leq C(k)\,\|\phi_1 - \phi_2\|_{\Sigma^k}\left(1 + \left\|\phi^1\right\|_{\Sigma^k}^2 + \left\|\phi^2\right\|_{\Sigma^k}^2\right),$$

with for all $\phi \in K$,

$$\|\phi\|_{\Sigma^k} = \langle A^k\phi, \phi\rangle_{L^2},$$

where the operator A is given by Equation (6.6), and $\langle \cdot, \cdot \rangle_{L^2}$ denotes the $L^2(\mathbb{R}^d)$-scalar product.

Theorem 6.2. *Let $k > d/2$. For all $M > 0$, there exists a maximal time step $h_0 > 0$ (uniform with respect to K) such that for all $h \leq h_0$ and for all $\phi_n \in K$ satisfying $\|\phi_0\|_{\Sigma^k} \leq M$, the system composed by Equations (6.15) and (6.16) has a unique solution.*

Proof of Theorem 6.2. The local existence follows from a Picard argument in any space $L^\infty([0,T], \Sigma^k)$ for $k \in \mathbb{N}$, and Lemma 6.1. □

In the following, we use a special kind of Lawson method, called Gauss-Lawson method. It is actually a special kind of collocation methods. Let us begin by recalling the definition of a collocation method. Suppose that we want to approximate an ODE of first order, given by,

$$\frac{d}{dt}u(t) = F(t, u(t)). \tag{6.17}$$

That is to say, we are looking for an approximation of $u(t_0+h)$, knowing $u(t_0)$. A collocation method of degree s consists in finding a polynomial function y of degree s, taking values in K, and satisfying Equation (6.17) at least in s prescribed points.

Definition 6.3. *Let $(c_i)_{1\leq i\leq s} \in [0,1]^s$ be distinct real numbers. We define the collocation polynomial $y(t)$ as the polynomial of degree at most s such that,*

$$\begin{aligned} y(t_0) &= y_0, \\ \dot{y}(t_0 + c_i h) &= F(t_0 + c_i h, y(t_0 + c_i h)), \quad \text{for} \quad i = 1, \dots, s. \end{aligned} \tag{6.18}$$

Then, the numerical solution $y_1 = y(t_0 + h)$ is called a collocation method for Equation (6.17).

Such a method corresponds to a special kind of Butcher tableau where the coefficients $(a_{i,j})_{1\leq i,j\leq s}$ and $(b_j)_{A\leq j\leq s}$ are given by the following theorem.

Theorem 6.4 (Theorem 1.4 [91])**.** *The collocation method given by Definition 6.3, for a given set $(c_i)_{1\leq i\leq s} \in [0,1]^s$ of distinct real numbers, corresponds to a special kind of Runge-Kutta method, with Butcher tableau (6.14) given by,*

$$a_{i,j} = \int_0^{c_i} l_j(\tau) d\tau, \quad b_i = \int_0^1 l_i(\tau) d\tau,$$

where l_i is the Lagrange polynomial given by $l_i(\tau) = \prod_{l\neq i}(\tau - c_l)/(c_i - c_l)$.

The Gauss collocation method consists in a special choice of the nodes $(c_i)_{1\leq i\leq s}$. They are chosen to be the zeros of the sth shifted Legendre Polynomial, which is given by,

$$\frac{d}{dx^s}\left(x^s(x-1)^s\right).$$

This choice of nodes is especially interesting since it ensures that such a method conserves quadratic invariants (such as the L^2-norm in our case) [92, Theorem 2.2], converges at order $2s$ (under sufficient regularity assumptions) and is symplectic [92, Theorem 4.2]. We recall that symplectic schemes ensure good long time conservation of invariants (in particular the energy) by reproducing, at the discrete level, some structure of the continuous equation. To clearly define this property, we consider the isomorphism $\mathbb{C} \to \mathbb{R}^2, x \mapsto (\Re x, \Im x)$, and replace $\mathbb{C}$ by $\mathbb{R}^2$ using this isomorphism. The multiplication by i of an element of K can be represented by a product with the square matrix J of size $2 \cdot \dim(K)$ given by $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. A numerical scheme is said to be symplectic if its discrete flow $\tilde{\Phi}_H$ (taking values in $\mathbb{R}^{2 \cdot dim(K)}$) is symplectic, *i.e.*,

$$(Jac(\tilde{\Phi}_H)(\phi))^T J Jac(\tilde{\Phi})_H(\phi) = J, \quad \forall \phi \in K,$$

where $Jac(\tilde{\Phi}_H)$ denotes the Jacobian of the flow $\tilde{\Phi}_H$.

Theorem 6.5. *Let $k > d/2$. The Gauss-Lawson method given by the discrete flow $\Phi_{H,h}$ conserves the $L^2(\mathbb{R}^2)$-norm as long as it is well posed in the sens of Theorem 6.2. As a consequence, the Gauss-Lawson method is globally well-posed, whenever it is locally well-posed. More precisely, for all $M > 0$, there exists $h_0 > 0$ such that for all $h \leq h_0$ and for all $\phi_0 \in K$ satisfying $\left\|\phi^0\right\|_{\Sigma^k} \leq M$, the following equation has a unique solution: $(\phi_{n+1})_{n\in\mathbb{N}} \in K^{\mathbb{N}}$.*

$$\phi_0 = \phi^0, \quad \textit{and} \quad \phi_{n+1} = \Phi_{H,h}\phi_n, \forall n \in \mathbb{N}. \tag{6.19}$$

Proof of Theorem 6.5. Global existence follows from the conservation of the $L^2(\mathbb{R}^2)$-norm. It is classical that the Gauss collocation method conserves the quadratic first integrals, and in particular the $L^2(\mathbb{R}^2)$-norm in this case. See Corollary 12 [22]. Suppose that the ODE defined by Equation (6.17) is such that the quadratic form I defined by $I(u) = u^T C u$ with C a symmetric matrix, is a first integral of this dynamics. Then, for all u, and for all $t \geq 0$ $u^T C F(t, u) = 0$. Let $y(t)$ be the collocation polynomial of the Gauss method. Then it holds,

$$I(y(t_0 + h)) - I(y(t_0)) = 2 \int_{t_0}^{t_0+h} y(t)^T C \dot{y}(t)\, \mathrm{d}t.$$

Since $y(t)^T C \dot{y}(t)$ is a polynomial of degree $2s - 1$, it is integrated without error by the Gaussian quadrature formula, which vanishes since for all $i = 1, \ldots, s$,

$$y(t_0 + c_i h)^T C \dot{y}(t_0 + c_i h) = y(t_0 + c_i h)^T C F(t_0 + c_i h, y(t_0 + c_i h)) = 0.$$

□

The symplecticity follows from the fact that every Runge-Kutta scheme that preserve quadratic first integrals, is a symplectic method (see section VI.4 from [91]).

Proposition 6.6. *The Gauss-Lawson method defined by the discrete flow $\Phi_{H,h}$ is symplectic whenever it is well-posed (by Theorem 6.5).*

As said previously, the Gauss-Lawson method converges at order $2s$. This result is the analogous of Theorem 14 [22].

Theorem 6.7. *Let $\phi^0 \in K$ and $T > 0$. Then there exists $h_0 > 0$ such that for all $h \leq h_0$, the s-stage Gauss-Lawson method is well-posed in the sense of Theorem 6.5. Let $(\phi_{n+1})_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ be the solution of Equation (6.19) with initial condition ϕ^0. Let also $(\phi(t))_{t\geq 0}$ be the solution of (6.8) with initial condition ϕ^0. Then, there exists $C > 0$ such that,*

$$\forall h \leq h_0, \forall n \in \mathbb{N} s.t. 0 \leq nh \leq T, \quad \|\phi(t_n) - \phi_n\|_{\Sigma^k} \leq Ch^{2s}. \tag{6.20}$$

Remark 6.8. *Let $k > d/2$. In our stochastic setting, Theorem 6.2 is quite weak. Since the global scheme to solve Equation (6.4) consists in alternating the discrete flows $\Phi_{L,\varepsilon,h}$ and $\Phi_{H,h}$, the initial condition for the flow $\Phi_{H,h}$ becomes random. Thus the maximal time steps h_0 that ensures well-posedness (in Theorem 6.5) is also random. If it becomes smaller than the initial choice of h, then we will not be able to solve the Hamiltonian dynamics after this time without refining the time step. This maximal time step h_0 may become very small if the $L^2(\mathbb{R}^2)$-norm of the solution of the flow $\Phi_{L,\varepsilon,h}$ becomes large. Because of the additive noise, the $L^2(\mathbb{R}^2)$-norm of the scheme is not conserved, and worse, it can take arbitrarily large values with positive probability. Then, we cannot conclude to the almost sure well-posedness of the global scheme for any time step small enough. One way to rigorously define a well-posed version of this scheme would be to introduce a stopping time when the solution of the global numerical scheme becomes larger (for the $L^2(\mathbb{R}^2)$ norm) than a prescribed threshold. Then, it is enough to stop the dynamics at this stopping time. We believe that such a procedure would at least enable to prove a convergence in probability. This kind of argument has been used in [20, 142]. In practice we do not observe this problem of definition for sensible time steps h. This is related to the fact that the solution of Equation (6.4) takes large values only with small probability thanks to the dissipative term.*

6.3.2 Numerical results for the Hamiltonian integration

We present now some numerical results for the Gauss-Lawson integrator of the Hamiltonian dynamics (6.8), since it is the key point in the numerical integration of (6.4). We numerically verify the order $2s$ of convergence and the good conservation of the energy. To do so, we compare the solutions of the Gauss-Lawson scheme for various time steps and number s of collocation points. We choose for initial condition $\phi^0 = Y_{0,0} + Y_{0,1}$, where $Y_{0,0}$ and $Y_{0,1}$ are given by (6.7), and where the numerical parameters are given later in Table 6.1. The results are displayed in Figure 6.2.

To verify the order of convergence given by Theorem 6.7, we would like to compute the error $\sup_{nh\leq T} \|\phi(t_n) - \phi_n\|_{L^2}$ that appears in Equation (6.20). Yet, since we do not know

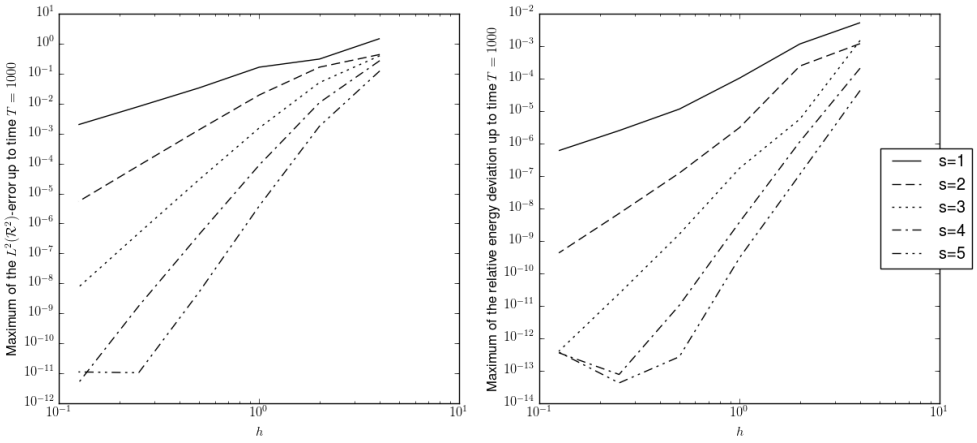

Figure 6.2 – L^2-error (left) and energy conservation (right), with respect to h, for the Lawson method for $s = 1, \dots, 5$.

the exact solution $(\phi(t))_{t\geq 0}$, we actually compute the numerical solutions $(\phi_n^{h_i})_{n\in\mathbb{N}}$ for geometric time steps h_i given by $h_i = h_0/2^i$, and we plot in the left part of Figure 6.2 the quantity $\sup_{nh\leq T} \left\|\phi_n^{h_{i+1}} - \phi_n^{h_i}\right\|_{L^2} / \left\|\phi_0^{h_i}\right\|_{L^2}$. We can observe that the slopes are close to $2s$ for all these graphs, which are the orders of convergence predicted by Theorem 6.7. We can observe for $s = 5$ and $h \lesssim 0.2$ that the error is floored around 10^{-12}. This behaviour is only due to informatic limitations arising from the way that computers encode and store real numbers. Our program is written in C, and uses double-precision floating-point format to represent real numbers. This implementation enables us to work with approximately sixteen significant digits. Since we iterate $1000/h_i \approx 10^4$ times the approximated flow over one time step h_i, we expect to keep around 10^{11} significant digits on the computation represented in this Figure, which is consistent with this floor on the error.

On the right-hand side of Figure 6.2, we plot the quantity $\sup_{nh\leq T} \left|E(\phi_n^{h_i})/E(\phi_0^{h_i}) - 1\right|$, where E is given by Equation (6.5), for the same time steps h_i as those used in the left side of the Figure 6.2. We can observe a relative deviation similar to the one for the L^2-norm. Since only the purely Hamiltonian dynamics conserves the energy, and only the purely reversible Langevin dynamics should enable to cross between the iso-energy manifolds, it seems quite important to build a numerical method that enables to reproduce this property.

Remark 6.9. *To properly take into account the projector P_K in Equation (6.4), we propose to make use of a spectral representation based on the Gauss-Laguerre modes. This spectral representation can also be used to discretize a genuine infinite dimensional Gross-Pitaevkskii equation, as was done in [14]. In this article, the authors propose a time-splitting method for the time discretisation, instead of a Lawson method. It is worth noting that, in the case of an isotropic harmonic potential, this spectral discretisation presents some advantages and drawbacks with respect to the Fourier one. The main disadvantage of our method is the fact that the computation of the nonlinearity cannot be computed using Fourier transforms,*

contrarily to the splitting method. This computation is explained in the next section. The two main arguments in favor of the Gauss-Laguerre discretisation are the following. First, Fourier transforms require to bound the support of the solution, and introduce periodic boundary conditions, that may lead to a significant error. Second, considering the potential confinement as part of the linear operator may lead to a better representation of the solution, which means that less modes may be required to describe the solution for a given error.

6.4 Metastability analysis with the AMS algorithm

As explained in the introduction, we are interested in numerically analyzing the transitions between two metastable states composed of different numbers of vortices inside the condensate. We begin by fixing the set of numerical parameters for our numerical investigation for which several metastable states exist. They are chosen to be a compromise between the typical parameters used in physical experiments and the computational cost of the numerical experiment. The dimensional reduction of the SPGPE to the two dimensional model that we use is only valid when the condensate is much more elongated in the radial dimension than in the axial one. Thus, we need to choose an harmonic trap much more confining in the axial direction. The dimensionless formulation is then given by,

$$d\phi_t = -(i + c_1 \cdot T)\nabla E^K(\phi_t)\,\mathrm{d}t + (c_1 \cdot T)c_2\, dW_t^K, \tag{6.21}$$

where T denotes the temperature and with,

$$\nabla E^K(\phi) = P_K\left[\left(\frac{1}{2}(-\Delta + |x|^2) - \mu + \Omega L_z + g\,|\phi|^2\right)\phi\right], \quad \text{and} \quad (W_t^K)_{t\geq 0} = (P_K W_t)_{t\geq 0}.$$

This formulation is called dimensionless because the coefficient $c_1 \cdot T$, the unknown and the variables (x and t) are dimensionless. The set of numerical parameters is given in Table 6.1 where the temperature T is given in nanoKelvins. We also provide, in Table 6.2

Table 6.1 – Numerical parameters of the experiment

c_1	c_2	μ	Ω	g	$\omega_{\max}$
$7.59 \cdot 10^{-6}\mathrm{nK}^{-1}$	0.302	3.5	0.77	1200	5

Table 6.2 – Physical (dimensional) parameters of the experiment

N	ω_r	ω_z	Ω	E_{cut}	μ	T	ρ
2000	$2\pi \cdot 10^4$ rad $\cdot\, \mathrm{s}^{-1}$	$5\omega_r$	$0.77\omega_r$	$5\hbar\omega_r + 0.5\hbar\omega_z$	$3.5\hbar\omega_r + 0.5\hbar\omega_z$	0.6 nK	3

the physical parameters before the adimensionalisation. We recall that N is the approximate number of atoms in the condensate. The molecular species we simulate is Rubidium 87, with mass m is given by $m = 1.44 \cdot 10^{-25}$kg and the diffusion length a is given by, $a = 5.77 \cdot 10^{-9}$m. We do not pretend that the numerical values given in Table 6.2 are phys-

ically relevant. More precisely, the trapping potential has been overestimated, whereas the cut-off level E_{cut} and the number of atoms has been underestimated. Actually, the choice of these numerical values has been made to reduce the computational cost while conserving the metastable behaviour. With this set, we count three stable vortex configurations composed by two, three or four vortices, shown in Figure 6.3. These stable configurations

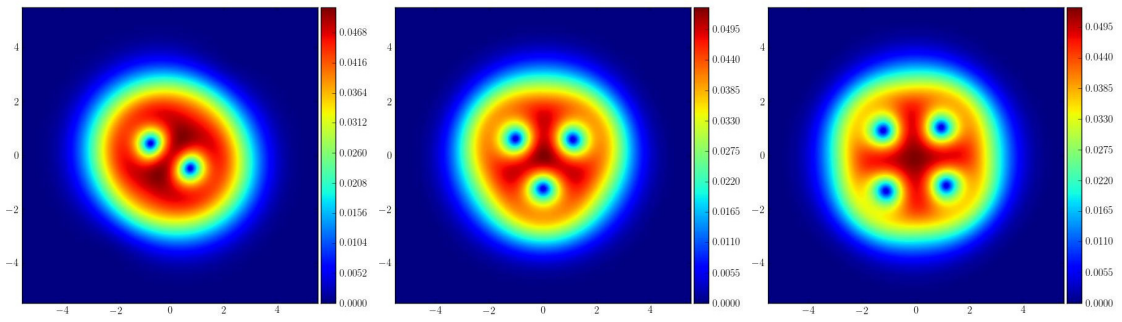

Figure 6.3 – Absolute square of the local minima of the Gross-Pitaevskii energy using the numerical parameters given in Table 6.1.

are found by using a gradient descent algorithm initialised with initial conditions sharing the same symmetries as the targeted vortex solutions. Moreover, a long time evolution of the Langevin dynamics (6.10), with a sufficiently high temperature, enables to verify that no metastable configuration has been missed. Typical evolutions of the condensate conserve the number of vortices inside the condensate over large times (for a temperature below 1nK), until one vortex exits or enters the condensate. We emphasize here that even though the numerical simulations provide (very) strong evidences that these states are stable (up to a rotation and a uniform change of phase, as explained later on), and thus separated by an energy barrier, we are not able to prove it rigorously. We refer to [156] for a theoretical study of the stability of radially symmetric minimizers of E given by (6.5).

Understanding the pathways leading to a modification of the number of vortices inside the rotating condensate is a complicated task since the observation of these events are rare in the small temperature limit. Thus, it is necessary to develop some dedicated Monte Carlo methods enabling to sample efficiently trajectories conditioned by the occurrence of these rare events. This section is devoted to the exposition of such a method based on the Adaptive Multilevel Splitting algorithm. We believe that the interest of these Monte Carlo techniques is twofold. First, it can give some qualitative insight about the rare vortex trajectories that lead to an exit of, or an entrance in the condensate. For instance, we hope to understand *how* the condensate should be prepared to favour the exit (or entrance) of a vortex. We also hope to understand to what extent the barriers between vortex configurations are of potential or entropic nature. Quantitatively, we aim at estimating the average time for a vortex to enter or exit the condensate. A lot of questions remain open. For instance, we do not know if we can write an Eyring-Kramers formula in our case. We aim at providing numerical evidences to these questions.

6.4.1 The Adaptive Multilevel Splitting Algorithm (AMS)

Let us now fix the notation before presenting the Adaptive Multilevel Splitting algorithm in an abstract setting. Let P be the transition probability kernel of a Markov chain $X = (X_n)_{n\in\mathbb{N}}$, taking values in a finite dimensional space $\mathcal{S}$. In our case, $\mathcal{S} = K$, and P is the transition probability kernel of the h-skeleton of the solution of Equation (6.21). This means that for all $x \in \mathcal{S}$, $P(x,\cdot)$ is the density with respect to the Lebesgue measure, of the law of ϕ_h, where $(\phi_t)_{t\geq 0}$ is the solution of Equation (6.21) such that $\phi_0 = x$. We could generalize the setting by assuming that $\mathcal{S}$ is only a Polish space. We also suppose that the transition kernel P is Lebesgue-irreducible, and recurrent, so that there exists a unique invariant probability measure for P. We assume in addition that the chain is aperiodic so that P is ergodic. As recalled in the introduction of the thesis, these hypotheses are satisfied in our case. We then denote by $\mathcal{P}$ the state space of the path space, that is to say,

$$\mathcal{P} = \mathcal{S}^{\mathbb{N}}$$

We consider the trajectories to be random variables taking values in $\mathcal{P}$, endowed with the product topology, generated by the distance $d_{\mathcal{P}}$, defined for all $X, Y \in \mathcal{P}$ by,

$$d_{\mathcal{P}}(X,Y) = \sum_{n\in\mathbb{N}} \frac{1}{2^n}(1 \wedge \sup_{m\leq n} d_{\mathcal{S}}(X_m, Y_m)).$$

We denote by ρ_{X_0} the law of the $\mathcal{P}$-valued random variable X.

The algorithm produces an estimator of the probability that a Markov chain (X_n), taking values in the state space $\mathcal{S}$, starting at the position $X_0 = x$, reaches a set B before reaching another disjoint set A, denoted by $\mathbb{P}_x(\tau_B < \tau_A)$, where τ_A and τ_B define respectively the hitting times of the sets A and B. The idea is to define A and B as two distinct metastable regions, with x chosen so that $x \notin A$, but sufficiently close to A so that the event $\{\tau_B < \tau_A\}$ is indeed rare (typically of probability of order 10^{-4} to 10^{-12}). Later, the starting point x will be a random variable with a law supported into the strict exterior of A (and mainly close to the boundary of A). The algorithm is based on an interacting system of replicas evolving through a mutation-selection procedure. The general idea to estimate rare events, is to define a decreasing (for the inclusion) family of events $(C_j)_{0\leq j\leq m}$ such that $C_0 = \Omega$ and $C_m = \{\tau_B < \tau_A\}$. Writing then,

$$\mathbb{P}\left(\tau_B < \tau_A\right) = \prod_{j=1}^{m} \mathbb{P}\left(C_j | C_{j-1}\right),$$

the idea is to estimate $\mathbb{P}\left(\tau_B < \tau_A\right)$ as a product of estimators of the quantities $\mathbb{P}\left(C_j|C_{j-1}\right)$. This approach is reasonable if we can estimate the quantities $\mathbb{P}\left(C_j|C_{j-1}\right)$ efficiently. This is generally true if

— $\mathbb{P}\left(C_j|C_{j-1}\right)$ is not too small, which would make classical Monte Carlo methods

inefficient,
— we know how to sample conditionally to C_{j-1}.

The Adaptive Multilevel Splitting algorithm (AMS) provides a way to construct such a family of events. This approach of stratification based on nested sampling has been used for instance in the context of nested sampling [158], RESTART [164], POP and IPS algorithms [87].

To define such a family $(C_j)_{0\leq j\leq m}$, the algorithm crucially relies on a prescribed function ξ, defined on the state space $\mathcal{S}$ and taking values in $\mathbb{R}\cup\{+\infty\}$, called *reaction coordinate*. If AMS is classified as a selection-mutation algorithm, it is because it essentially selects, among some replicas, the trajectories that maximize the reaction coordinate along them, and *mutate* them. Thus, the function ξ should be chosen such that the algorithm selects the replicas that are more likely to *mutates* into *reactive trajectories* (which means that they reach B before A in this context). We give more formal indications about the choice of the reaction coordinate in Remark 6.14. In our practical case, we only require that the reaction coordinate and the set B are compatible in the sense that we impose for all $x\in B$, $\xi(x)=+\infty$.

Remark 6.10. *In the framework of [33], the set B and the reaction coordinate are related. The authors require that there exists a constant $z_{max}\in\mathbb{R}$ such that,*

$$B\subset\xi^{-1}(]z_{\max},\infty]). \tag{6.22}$$

Our condition $\xi(B)\subset\{+\infty\}$ implies the condition (6.22). This trick gives us more freedom for the choice of B, since the condition (6.22) is always satisfied. Thus, instead of choosing A and B with respect to some level sets of the reaction coordinate, we can choose for instance A and B as level sets of the energy, which might be more natural in some applications.

We now present the algorithm using the notation of [33] in which a general formulation can be found. We only present here the simplified formulation that has been implemented. For a given reaction coordinate, and given sets A and B, the algorithm depends on four parameters:
— the number n_{rep} of replicas,
— the (minimum) number k of replicas killed at each iteration,
— the initial condition x,
— the minimum number n_{success} of replicas that reach B before A required to stop the algorithm.

We define the set $I=\{1,\dots,n_{\text{rep}}\}$. For a given reaction coordinate ξ, we introduce the function Ξ, taking values in $\mathbb{R}\cup\{+\infty\}$, defined for all discrete-time processes $X=(X_k)_{k\in\mathbb{N}}$ that take values in $\mathcal{S}$ by,

$$\Xi(X)=\sup_{k\in\mathbb{N}}\xi(X_k). \tag{6.23}$$

We also define the partial resampling method, that defines the mutation step. We introduce

the family of transition probability kernels $(\pi_z(X, dY))_{z\in\mathbb{R}}$ on $\mathcal{P}$, as the law of the $\mathcal{P}$-valued process $Y = (Y_n)_{n\in\mathbb{N}}$ defined by,

$$\begin{cases} Y_n = X_n, & \text{if } n \leq T_z(X) = \inf\{n \leq \tau_A(X), \xi(X_n) > z\}, \\ \text{Law}\,(Y_n | Y_m, 0 \leq m \leq n-1) = P(X_{n-1}, \cdot), & \text{if } n > T_z(X), \end{cases}$$

and stopped when Y reaches A or B. This kernel correspond to copying the process X up to the first time when the reaction coordinate becomes greater than the level z. After this time, the Markov chain Y is iterated with the probability transition kernel P, which is the transition kernel of the unbiased chain.

Then, the AMS algorithm is given by,

1. initialisation step
   (a) Let $(X^{n,0})_{n\in I}$ be i.i.d realizations of the stopped dynamics $(X_{n\wedge\tau_A\wedge\tau_B})$ starting from x.
   (b) Compute a permutation $\Sigma^{(0)}$ of I, such that,
   $$\Xi(X^{(\Sigma^{(0)}(1),0)}) \leq \ldots \leq \Xi(X^{(\Sigma^{(0)}(n_{\text{rep}}),0)}).$$
   (c) Set the initial selection level $Z^{(0)} = \Xi(X^{(\Sigma^{(0)}(k),0)})$.
2. Iterations. Iterate on $q \geq 0$, while
   $$Z^{(q)} \neq \Xi(X^{(\Sigma^{(q)}(n_{\text{rep}}),q)}) \text{ and } \text{Card}\left\{i; \tau_B(X^{(i,q)}) < \tau_A(X^{(i,q)})\right\} < n_{\text{success}}.$$
   (a) Selection step. We select now the (at least k) least performer replicas, which are the ones that minimizes Ξ. We define a partition $\{I_1(q), I_2(q)\}$ of I such that,
   $$I_1(q) = \{n \in I; \Xi(X^{(\Sigma^{(q)}(n),q)}) \leq Z^{(q)}\}, \quad I_2(q) = I \setminus I_1(q),$$
   and $I_1(q)$ contains the *worse* replicas. We denote by $K(q) = \text{Card}(I_1(q))$ their number.
   (b) Mutation step. This step consists in mutating only the replicas of indices in $I_1(q)$. First, we leave unchanged the replicas of indices in $I_2(q)$ by setting for all $i \in I_2(q)$, $X^{(\Sigma^{(q)}(i),q+1)} = X^{(\Sigma^{(q)}(i),q)}$. To mutate the replicas of indices in $I_1(q)$, we sample independently for all $i \in I_1(q)$ uniform random variables $u(i)$ in $I_2(q)$. Then we sample independently, for all $i \in I_1(q)$, $\mathcal{S}$-valued random variables $X^{(\Sigma^{(q)}(i),q+1)}$ according to the transition density, $\pi_{Z^{(q)}}(X^{(\Sigma^{(q)}(u(i)),q)}, \cdot)$.
   (c) We define the selection level $Z^{(q+1)}$ by computing a permutation $\Sigma^{(q+1)}$ of I such that,
   $$\Xi(X^{(\Sigma^{(q+1)}(1),q+1)}) \leq \ldots \leq \Xi(X^{(\Sigma^{(q+1)}(n_{\text{rep}}),q+1)}),$$

and set $Z^{(q+1)} = \Xi(X^{(\Sigma^{(q+1)}(k),q+1)})$.

(d) Set $q \leftarrow q+1$.

3. End of the algorithm. When the stopping condition has been reached, we define the stopping time Q_{iter} with the current value of q (without the last increment), $Q_{\text{iter}} = q$. We define $\hat{p}$, the estimator of $\mathbb{P}_x(\tau_B < \tau_A)$ as,

$$\hat{p} = \frac{1}{n_{\text{rep}}} \text{Card}\left\{i; \tau_B(X^{(i,Q_{\text{iter}})}) < \tau_A(X^{(i,Q_{\text{iter}})})\right\} \prod_{j=0}^{Q_{\text{iter}}} \left(1 - \frac{K(j)}{n_{\text{rep}}}\right). \tag{6.24}$$

Remark 6.11. *This algorithm is a simplified version of the more general one given in [33], that only enables to compute the expectation of the event $\mathbb{1}_{\{\tau_B<\tau_A\}}$. Yet, the general formulation enables to compute expectations, with respect to ρ_{X_0}, of more general statistics.*

For the algorithm to be well-posed, we need Q_{iter} to be almost surely finite. To ensure this property, we can stop the algorithm after a deterministic number of iterations. Moreover, it has been proved that it defines an unbiased estimator of $\mathbb{P}_x(\tau_B < \tau_A)$ (see Corollary 4.8 [33]), as recalled in the following proposition.

Proposition 6.12. *For any choice of n_{rep}, k, x, $n_{success}$ and reaction coordinate ξ such that $\xi(B) = +\infty$, the estimator $\hat{p}$ defined by (6.24) is an unbiased estimator of the probability $p = \mathbb{P}_x(\tau_B < \tau_A)$,*

$$\mathbb{E}\left[\hat{p}\right] = p.$$

The proof can be found in [33]. It is carried out in the general case. The strategy consists in viewing the probability p as an expectation, *i.e.* $p = \mathbb{E}_x(\mathbb{1}_{\{\tau_B<\tau_A\}})$. Then a conditional expectation on the *intermediary* estimators $\hat{p}_q$ defined for all $q \geq 0$ by

$$\hat{p}_q = \frac{1}{n_{\text{rep}}} \text{Card}\left\{i; \tau_B(X^{(i,q-1)}) < \tau_A(X^{(i,q-1)})\right\} \prod_{j=0}^{q-1} \left(1 - \frac{K(j)}{n_{\text{rep}}}\right),$$

is proved to be a martingale. Eventually the unbiasedness result can be seen as the conservation of the expectation of this martingale at the stopping time defined in the algorithm.

It is worth noting that because of the discrete time setting, the number of killed replicas at each iteration can be equal to the total number of replicas n_{rep}. This scenario occurs when $Z^{(q)} = \Xi(X^{(\Sigma^{(q)}(n),q)})$. It is referred to as *extinction* and it stops the algorithm. It is explained in detail in Remark 2.4 of [32]. This behaviour is more likely to occur for larger time steps h of the skeleton of the continuous process, and smaller temperatures. This is also true if one uses a discrete-valued reaction coordinate.

Remark 6.13. *A discrete-valued reaction coordinate might not enable to distinguish all the replicas and may lead to extinction. The typical scenario for this phenomenon corresponds to the situation where none of the the mutated trajectories get sufficiently close to B after*

the branching point to increase their reaction coordinate. In this scenario all replicas have the same maximal reaction coordinate equal to the one at the branching point, which leads to the killing of all the replicas, and the vanishing of the estimator. Large time steps also lead to a similar phenomenon. Indeed, it might happen that even though some replicas come closer to B, they end up going quickly to A. Thus, if the time step is not small enough, we may miss this behaviour of the continuous dynamics, and only observe the replicas after they got back to A.

Remark 6.14. *The question of choosing the reaction coordinate is a difficult one. Even though Equation (6.22) is the only assumption we require on ξ, its choice can greatly impact the variance of the estimator. It has been shown (in [40]), for a time continuous version of the AMS algorithm, that the best choice of reaction coordinates, in terms of minimization of the variance of the AMS estimator, is given by the committor function q defined by,*

$$q(x) = \mathbb{P}_x(\tau_B < \tau_A). \tag{6.25}$$

Because of our definition of the stopping criterion, the analogous of the reaction coordinate would formally be

$$\xi(x) = \mathbb{P}_x(\tau_B < \tau_A) + (+\infty)\mathbb{1}_{\{x \in B\}}.$$

Nevertheless, this function is not known a priori, and the AMS algorithm precisely aims at estimating such quantities. Thus, in practice, one has to construct a reaction coordinate based on some intuition. Heuristically, it is natural to look for a reaction coordinate that would be increasing along the most probable reactive trajectory. Yet, we do not always know how to do so.

6.4.2 Numerical computations of reactive trajectories

We focus in our numerical experimentations on the transition from the 3-vortex configuration, to the 2-vortex configuration. The key point of the numerical study relies on the choice of the reaction coordinate. Since we are interested in the change of the number of vortices inside the condensate, we propose to define the reaction coordinate as the the distance between the centre of the harmonic trap and the third closest vortex from the centre. We recall that, in the two dimensional setting, a vortex is defined as a point of phase singularity of the wave function. Suppose that the wave function has n_{vortex} vortices. Let $(r_i)_{1 \leq i \leq n_{\text{vortex}}}$ be the distance between these points and the origin (which is the centre of the trap), and σ be a permutation of $\{1, \dots, n_{\text{vortex}}\}$ such that,

$$r_{\sigma(1)} \leq \dots \leq r_{\sigma(n_{\text{vortex}})}.$$

We define the reaction coordinate ξ by $\xi = r_{\sigma(3)}$. Since at the end of the reactive trajectory the third vortex will draw away from the centre of the condensate, then this reaction

coordinate will globally increase along this trajectory, but certainly not monotonically. Obviously, we do not know *a priori* if it is increasing along the most favorable reactive path as discussed in Remark 6.14. Besides we can point out that this reaction coordinate shares the same symmetries as the Gross-Pitaevskii energy. More precisely, it is invariant by a uniform change of phase, and by any rotation centred at the origin. This is a good point because the reaction coordinate should encode the progress of the reactive trajectory, which is independent of these symmetries, and thus it should be independent from them as well. We refer to Appendix 6.6.3 for a practical method to locate the vortices.

We define the sets A and B respectively as the connected component of some level sets of the energy containing respectively the local minima corresponding respectively to the 3 and 2-vortex configurations. We suppose in addition that $A \cap B = \emptyset$. The sets A and B should be chosen to be included inside the basin of attraction of these two local minima. This is possible since we suppose that they are separated by an energy barrier. These sets can be chosen arbitrarily among the sets that satisfy the above conditions. We refer to Remark 6.15 for precisions about the practical choice of the sets A and B. If we are interested by the connected component of some level sets of the energy and not by those of the reaction coordinate (as it is usually done), it is because we aim at analysing the effect of the intensity of the dissipation, given by c_1 in Equation (6.21), on the average exit time of the 3-vortex configuration. In this case it is the more natural choice since it forbids the dynamics to hit the sets A and B in the limit where the temperature vanishes (where the dynamics becomes purely Hamiltonian), which may not be the case otherwise.

Remark 6.15. *We recall that the AMS estimator is unbiased whatever the choice of A and B. Thus, they can be chosen arbitrarily. We define them in such a way that entering or exiting from them is not a rare event, supposing that these sets do not contain metastable subsets. We achieve this by defining A (resp. B) such that a process initialized in the basin of attraction of the 3-vortex (resp. 2-vortex) configuration spends around half of the time inside this set and half of the time outside, as long as it does not change of basin of attraction. This choice enables to make entering events as likely as exiting events. In practice, we sample two Markov chains $(X_n^A)_{n\in\mathbb{N}}$ and $(X_n^B)_{n\in\mathbb{N}}$ initialized at the 3-vortex and 2-vortex minima. We estimate E_A and E_B such that for a given N, chosen large enough,*

$$\frac{1}{N}\sum_{n=0}^{N} \mathbb{1}_{\{E(X_n^A)\leq E_A\}} \approx 0.5, \quad \textit{and} \quad \frac{1}{N}\sum_{n=0}^{N} \mathbb{1}_{\{E(X_n^B)\leq E_B\}} \approx 0.5.$$

We also need to verify that these two Markov chains do not exit the basin of attraction of their initial starting point. To do so, we verify that,

$$\sup_{n\leq N} \xi(X_n^A) \leq \inf_{n\in\mathbb{N}} \xi(X_n^B).$$

One mathematical way to formulate this problem is to introduce the quasi-stationary

distributions (QSD) for the Markov process $(X_n)_{n\in\mathbb{N}}$ conditioned not to exit the basin of attraction of each minima. For instance, let S be a subset of the $\mathcal{S}$. The QSD for the Markov process $(X_n)_{n\in\mathbb{N}}$ conditioned not to leave the set S is a probability measure ν_S satisfying,

$$\nu_S(E) = \frac{\int_S \mathbb{P}_x(X_n \in E, n < \tau_S)\nu_S(dx)}{\int_S \mathbb{P}_x(n < \tau_S)\nu_S(dx)}, \quad \forall E \subset S, \quad \forall t > 0,$$

where τ_S denotes the first exit time of S,

$$\tau_S = \inf\{n \in \mathbb{N}; X_n \notin S\}.$$

We put aside the question of the well-posedness of these measures for the two basins of attraction containing A and B, and, supposing they are well-posed, we denote them by ρ_A and ρ_B. Then, the targeted values E_A and E_B are given by,

$$\int_K \mathbb{1}_{\{E(x)\leq E_A\}} d\rho_A(x) = 0.5, \quad \int_K \mathbb{1}_{\{E(x)\leq E_B\}} d\rho_B(x) = 0.5.$$

6.4.3 Computation of the transition times

The goal of this section is to present a numerical method based on [40], to compute the average transition time $T_{A\to B}$ to go from the set A to the set B.

The algorithm

First, let us define precisely this transition time. In all the following, we denote by $X = (X_n)_{n\in\mathbb{N}}$ a Markov chain with transition kernel P. To make the starting point clear, we denote by $X^x = (X_n^x)_{n\in\mathbb{N}}$ such a Markov chain starting from x. We also denote by $X^x_{\tau_C}$ the Markov chain X^x stopped at the first hitting time of the set $C \subset \mathcal{S}$. That is to say, $X^x_{\tau_C} = (X^x_{n\wedge\tau_C})_{n\in\mathbb{N}}$ with $\tau_C = \inf\{n \in \mathbb{N}; X_n^x \in C\}$. Let X be such a Markov chain, initialized under the invariant measure of P. Then we introduce the families of stopping times $(\tau_n^A)_{n\in\mathbb{N}^*}$ and $(\tau_n^B)_{n\in\mathbb{N}}$ defined for all $k \in \mathbb{N}$ by,

$$\begin{aligned}
\tau_0^B &= \inf\{n \geq 0; X_n \in B\},\\
\tau_{k+1}^A &= \inf\{n \geq \tau_k^B; X_n \in A\},\\
\tau_{k+1}^B &= \inf\{n \geq \tau_{k+1}^A; X_n \in B\}.
\end{aligned}$$

If the sequence of random times $(\tau_k^B - \tau_k^A)_{k\in\mathbb{N}^*}$ reaches stationarity, then the natural way to define $T_{A\to B}$ is given by,

$$T_{A\to B} = \lim_{k\to+\infty} \mathbb{E}\left[\tau_k^B - \tau_k^A\right]. \tag{6.26}$$

One way to reformulate this definition is to make use of the extracted Markov chain $(X_{\tau_{k+1}^A})_{k\in\mathbb{N}}$, if it reaches stationary.

Assumption 6.16. *The Markov chain* $(X_{\tau^A_{k+1}})_{k\in\mathbb{N}}$ *has a unique invariant measure, denoted by* ν.

Definition 6.17. *Under Assumption 6.16, we define the average transition time* $T_{A\to B}$ *by,*

$$T_{A\to B} = \mathbb{E}_\nu(\tau_0^B), \tag{6.27}$$

which is the expectation of the first hitting time of the set B*, for a Markov chain with initial condition distributed according to the measure* ν*, and transition probability kernel* P.

The definition (6.27) is somehow more practical than the expression (6.26) because if we are able to sample a random starting point according to ν, then we only have to sample estimators of one transition path from A to B without sampling transition paths from B to A. Nevertheless, it is not possible in practice to sample according to the measure ν without sampling these transition paths from B to A. We put aside this technical problem for the moment, and present a decomposition of a transition path that enables us to estimate the average transition time (supposing we are able to sample according to ν). We introduce an arbitrary closed set A' of $\mathcal{S}$ such that $A \subset A'$. The idea consists in breaking the transition path (that goes from A to B) of a Markov chain X initialized with the measure ν, into (a large number of) loops between the interior of A and the exterior of A', and the last part of the trajectory that reaches B without reaching A again. The transition time is the sum of all the times spent in the loops between A and $\mathcal{S}\setminus A'$, plus the time spent to reach B without going back to A. We denote by $(\alpha_k)_{k\in\mathbb{N}}$ the sequence of random variables that correspond to the times spent by the Markov chain X between the k-th hitting time of A and the k-th hitting time of $\mathcal{S}\setminus A'$. We also denote by $(\beta_k)_{k\in\mathbb{N}}$ the sequence of random variables that corresponds to the times spent by the Markov chain X between the k-th hitting time of $\mathcal{S}\setminus A'$ and the $(k+1)$-th hitting time of $A\cup B$. To define precisely these two sequences, we introduce the following successive hitting times defined for any Markov chain X and for all $k\in\mathbb{N}$ by,

$$\begin{aligned}
\tau^0_{A\cup B}(X) &= \inf\{n\geq 0; \quad X_n \in A\cup B\},\\
\tau^{k+1}_{A\cup B}(X) &= \inf\{n > \tau^k_{\mathcal{S}\setminus A'}; \quad X_n \in A\cup B\},\\
\tau^k_{\mathcal{S}\setminus A'}(X) &= \inf\{n > \tau^k_{A\cup B}; \quad X_n \in \mathcal{S}\setminus A'\}.
\end{aligned}$$

We also define the index τ_B of the last loop that ends up in B instead of A.

$$\tau_B(X) = \inf\{k\geq 0; \quad X_{\tau^{k+1}_{A\cup B}} \in B\}. \tag{6.28}$$

Thus, for all $k\in\mathbb{N}$, and for all Markov chain X, we define the times α_k and β_k by

$$\begin{aligned}
&\alpha_0 = \tau^0_{A\cup B}(X), \quad \alpha_k = \tau^k_{\mathcal{S}\setminus A'}(X) - \tau^k_{A\cup B}(X),\\
&\beta_k = \tau^{k+1}_{A\cup B}(X) - \tau^k_{\mathcal{S}\setminus A'}(X).
\end{aligned} \tag{6.29}$$

An example of this decomposition is depicted in Figure 6.4 in the case $\tau_B = 2$. The dotted

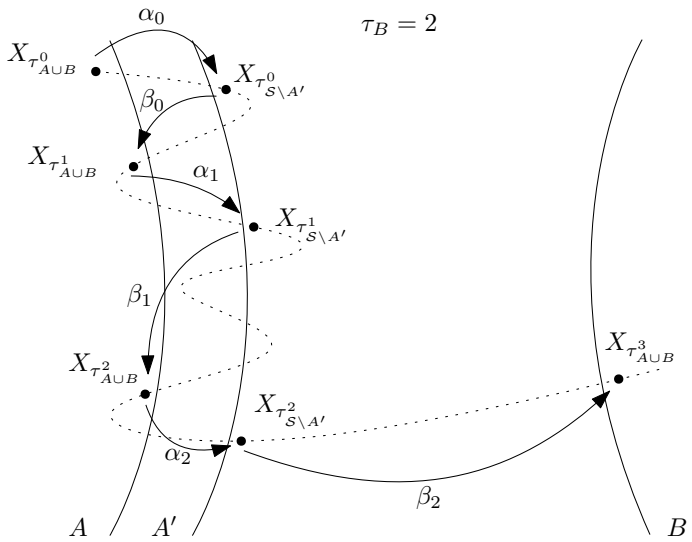

Figure 6.4 – Decomposition of a reactive trajectory between loops between A and $\mathcal{S} \setminus A'$ and A, and the last part of the trajectory.

line represents the trajectory of the underlying continuous Markov chain X, and the arrow represent the times α_k and β_k.

Under this framework, the first hitting time $T(X)$ of B for the Markov chain X is given by,

$$T(X) = \sum_{l \leq \tau_B(X)} \alpha_l(X) + \beta_l(X) = \sum_{k \in \mathbb{N}} (\alpha_k(X) + \beta_k(X)) \mathbb{1}_{\{\tau_B(X) \geq k\}}. \tag{6.30}$$

The following lemma summarises an expression of the expectation of $T(X)$, that uses conditional expectations with respect to the events $\{\tau_B \geq k\}$.

Lemma 6.18. *For any Markov chain X such that $T(X)$ is integrable, we have,*

$$\mathbb{E}(T(X)) = \sum_{k \in \mathbb{N}} u_k w_k, \tag{6.31}$$

with,

$$u_k = \mathbb{E}\left[\alpha_k(X) + \beta_k(X) | \tau_B(X) \geq k\right], \quad \textit{and} \quad w_k = \mathbb{P}\left(\tau_B(X) \geq k\right).$$

Moreover, if Assumption 6.16 holds, then,

$$T_{A \to B} = \mathbb{E}\left[T(X^\nu)\right].$$

The decomposition given by (6.30) does not provide yet an efficient method to estimate the expectation $T_{A \to B}$. Nevertheless a related quantity can be computed efficiently under

a stationarity assumption. We now define a new Markov chain whose invariant measure will be used to initialise the previous one. Consider for any $x \in A$, the two Markov chains $\tilde{U} = (\tilde{U}_n)_{n\in\mathbb{N}}$ and $\tilde{V} = (\tilde{V}_n)_{n\in\mathbb{N}}$ defined iteratively for all $k \in \mathbb{N}$ by,

$$\begin{aligned} \tilde{U}_0 &= x, \\ \mathcal{L}\left(\tilde{V}_k \middle| \tilde{U}_k\right) &= \mathcal{L}\left(X^{\tilde{U}_k}_{\tau_{A'}}\right), \quad \text{with} \quad \tau_{A'} = \inf\{n \geq 0; \quad X^{\tilde{U}_k}_n \in \mathcal{S} \setminus A'\}, \\ \mathcal{L}\left(\tilde{U}_{k+1} \middle| \tilde{V}_k\right) &= \mathcal{L}\left(X^{\tilde{V}_k}_{\tau_A} \middle| \tau_A \leq \tau_B\right), \text{with } \tau_E = \inf\{n \geq 0;\, X^{\tilde{V}_k}_n \in E\}, \end{aligned} \tag{6.32}$$

and $E = A$ or B. With such a construction, $\tilde{U}$ is supported in A and $\tilde{V}$ in $\mathcal{S} \setminus A'$. We will suppose the following stationary assumption.

Assumption 6.19. *The Markov chains $\tilde{U}$ has a unique ergodic invariant measure, denoted by ν_A.*

Suppose now that Assumption 6.19 holds. Then, the quantity $\mathbb{E}\left[T(X^{\nu_A})\right]$ can be nicely computed, using Lemma 6.18. Indeed, by construction, for all $k \in \mathbb{N}$,

$$\begin{aligned} &\mathcal{L}\left(\alpha_k(X^{\nu_A}) | \tau_B(X^{\nu_A}) \geq k\right) = \mathcal{L}\left(\alpha_0\right), \\ &\mathcal{L}\left(\beta_k(X^{\nu_A}) | \tau_B(X^{\nu_A}) \geq k\right) = \mathcal{L}\left(\beta_0 | \tau_B(X^{\nu_A}) \geq 0\right), \\ &\mathbb{P}\left(\tau_B(X^{\nu_A}) \geq k\right) = \mathbb{P}\left(\tau_B(X^{\nu_A}) \geq 1\right)^k = \left(1 - \mathbb{P}\left(\tau_B(X^{\nu_A}) = 0\right)\right)^k. \end{aligned}$$

The first equality in the last line comes from the fact that,

$$\mathbb{P}\left(\tau_B(X^{\nu_A}) \geq k+1 | \tau_B(X^{\nu_A}) \geq k\right) = \mathbb{P}\left(\tau_B(X^{\nu_A}) \geq 1\right).$$

Under Assumption 6.19, the expectation of $T(X^{\nu_A})$, defined by (6.30), is highly simplified since for all $k \in \mathbb{N}$, $u_k = u_0 = \mathbb{E}\left[\alpha_0(X^{\nu_A}) + \beta_0(X^{\nu_A})\right]$ and $w_k = \left(1 - \mathbb{P}\left(\tau_B = 0\right)\right)^k$. We obtain then,

$$\mathbb{E}(T(X^{\nu_A})) = \frac{\mathbb{E}\left[\alpha_0(X^{\nu_A}) + \beta_0(X^{\nu_A})\right]}{\mathbb{P}\left(\tau_B(X^{\nu_A}) = 0\right)}, \tag{6.33}$$

which can be reformulated in the following way,

$$\begin{aligned} \mathbb{E}(T(X^{\nu_A})) = &\frac{\mathbb{E}\left[\alpha_0(X^{\nu_A})\right] + \mathbb{E}\left[\beta_0(X^{\nu_A}) | \tau_B(X^{\nu_A}) > 0\right] \mathbb{P}\left(\tau_B(X^{\nu_A}) > 0\right)}{\mathbb{P}\left(\tau_B(X^{\nu_A}) = 0\right)} \\ &+ \mathbb{E}\left[\beta_0(X^{\nu_A}) | \tau_B(X^{\nu_A}) = 0\right]. \end{aligned} \tag{6.34}$$

In practice, the terms $\mathbb{E}\left[\alpha_0(X^{\nu_A})\right]$ and $\mathbb{E}\left[\beta_0(X^{\nu_A}) | \tau_B(X^{\nu_A}) > 0\right]$ can be efficiently sampled with a classical Monte Carlo method. Under Assumption 6.19, the initial measure ν_A can be sampled by simulating (after a burn-in period) the Markov process $\tilde{U}$ defined by (6.32), which is actually obtained by simulating the underlying Markov process X, with a rejection step if it reaches B. The terms $\mathbb{E}\left[\beta_0(X^{\nu_A}) | \tau_B(X^{\nu_A}) = 0\right]$ and $\mathbb{P}\left(\tau_B(X^{\nu_A}) = 0\right)$

can be estimated using the AMS algorithm.

This discussion can be summed up in the following algorithm. We set $N_1 \in \mathbb{N}$, the number of burn-in steps, that should be taken large enough. We set $N_2 \in \mathbb{N}^*$ the number of Monte Carlo samples for the estimation of $T_{A\to B}$. The algorithm is as follows, and it iteratively defines the Markov chains $(U_k)_{k\geq 0}$ and $(V_k)_{k\geq 0}$, and the sequences of random variables $(\alpha_k)_{k\geq 0}$, $(\beta_k)_{k\geq 0}$, $(\gamma_k)_{k\geq 0}$ and $(\delta_k)_{k\geq 0}$.

1. initialisation. Let $(U_k)_{k\in\mathbb{N}}$ and $(V_k)_{k\in\mathbb{N}}$ be two Markov chains. We define U_0 as an arbitrary element of A.
2. We iterate on $n = 0, \dots, N_1 + N_2 - 1$:
   (a) We sample, independently from the past, a Markov chain $X = (X_k^{U_n})_{k\in\mathbb{N}}$ up to time $\tau_{\mathcal{S}\setminus A'}(X)$ defined as the first hitting time of $\mathcal{S}\setminus A'$ for the Markov chain X. We set $V_n = X^{U_n}_{\tau_{\mathcal{S}\setminus A'}(X)}$. If $n \geq N_1$, we define $\alpha_{n-N_1} = \tau_{\mathcal{S}\setminus A'}(X)$.
   (b) If $n \geq N_1$, we compute independently from the past, an estimator γ_{n-N_1} of $\mathbb{P}\left(\tau_B(X^{V_n}) \leq \tau_A(X^{V_n})\right)$ using the AMS algorithm, where $\tau_A(X^{V_n})$ (resp. $\tau_B(X^{V_n})$) is the first hitting time of the set A (resp. B) for the Markov chain X^{V_n}. We also compute an estimator δ_{n-N_1} of $\mathbb{E}\left[\tau_B(X^{V_n})\middle|\tau_B(X^{V_n}) \leq \tau_A(X^{V_n})\right]$.
   (c) We sample, independently from the past, a Markov chain $X = (X_k^{V_n})_{k\in\mathbb{N}}$ up to time $\tau_{A\cup B}(X)$ defined as the first hitting time of $A\cup B$.
   — If $X^{V_n}_{A\cup B}(X) \in B$, we reject this trajectory and iterate this step.
   — Otherwise, $X^{V_n}_{A\cup B}(X) \in A$, and we set $U_{n+1} = X^{V_n}_{\tau_{A\cup B}(X)}$.
   If $n \geq N_1$, we define $\beta_{n-N_1} = \tau_{A\cup B}(X)$.
3. We define $\widetilde{T_{A\to B}}$ an estimator of $T_{A\to B}$ by,

$$\widetilde{T_{A\to B}} = \frac{\sum_{n=0}^{N_2-1} \alpha_n + \beta_n}{\sum_{n=0}^{N_2-1} \gamma_n} + \frac{1}{N_2}\sum_{n=0}^{N_2-1} \delta_n. \tag{6.35}$$

The rigorous justification of this algorithm, and the analysis of the bias given by $|\mathbb{E}\left[T(X^{\nu_A})\right] - \mathbb{E}\left[T(X^{\nu})\right]|$ still need to be investigated.

Numerical results

We present here some numerical results for the computation of the average time $T_{A\to B}$, using the algorithm described above. The underlying AMS estimator has been used with $n_{\text{rep}} = 32$ and $k = 1$ for the minimum number of replicas killed at each iteration. As we explained above, the algorithm actually computes the biased quantity $\mathbb{E}(T(X^{\nu_A}))$ since its expression is simply given by equations (6.33) and (6.34).

For the numerical experiments, we consider a slightly modified version of Equation (6.21) where the intensity of the non-reversibility and the temperature are independent; whereas the intensity of the non-reversibility is proportional to the temperature in the original

Equation (6.21). The numerical experiments are build on the following equation,

$$d\phi_t = -(i + c_1')\nabla E^K(\phi_t)\,\mathrm{d}t + c_2\sqrt{c_1 \cdot T c_1'}\,dW_t^K. \tag{6.36}$$

With the substitution $t \leftarrow tc_i'$, the intensity of the dissipation becomes independent from c_i', and the non-reversibility becomes proportional to $(c_i')^{-1}$.

$$d\phi_t = -((c_1')^{-1}i + 1)\nabla E^K(\phi_t)\,\mathrm{d}t + c_2\sqrt{c_1 \cdot T}\,dW_t^K. \tag{6.37}$$

Moreover, in order to reduce the computational time, we perform this computation with a dissipation c_1' much stronger than the typical value of $c_1 \cdot T$ (which is of order 10^{-7}). We choose $c_1' \in \{0.5^i, i = 0, \ldots, 4\}$, and compute these average exit times for temperatures T in the set $\{0.9^q, q = 0, \ldots, 4\}$. The result is given in Figure 6.5. We plot

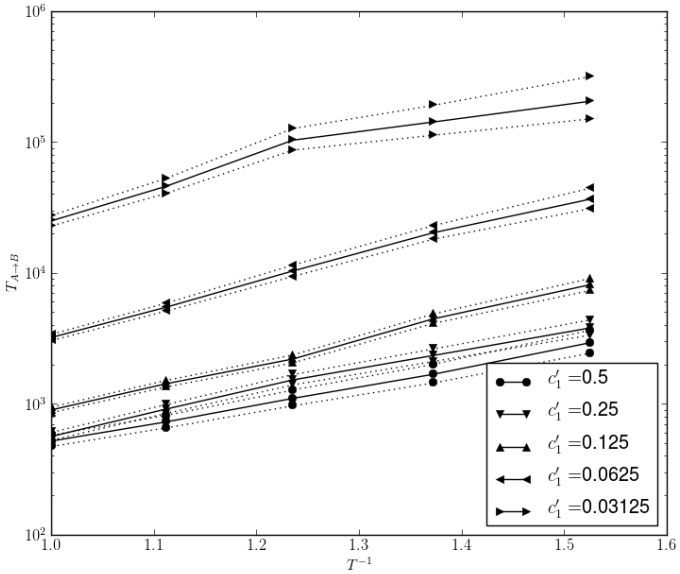

Figure 6.5 – Estimation of the average transition times $T_{A\to B}$ with respect to the inverse of the temperature for the dynamics (6.36) for various intensity c_1' of the dissipation.

the estimation of the average exit times $T_{A\to B}$ given by $\widetilde{T_{A\to B}}$ in Equation (6.35), for the dynamics (6.36), as a function of the inverse of the temperature. These averages were computed for 25500 independent realisations for each point. The straight lines correspond to the values of these estimators. They are surrounded by the dotted lines that correspond to the confidence interval at 95%. This computation was rather intensive since it ran during approximately 30 days on a cluster composed by 16 processors Intel Xeon E5-2650, which amounts to 128 threads.

First, we can observe that, as expected, the logarithm of the average exit time depends linearly on the inverse of the temperature. This kind of result is referred to as Arrhenius

law. This first point supports the fact that the method we applied here can be used (at least in the reversible case) to compute precisely the energy gaps that split up the metastable states. Let us now use a linear regression to estimate the energy gaps and the prefactors for each dissipation coefficients c_i'. For each of these coefficients, we model the average transition times $T_{A\to B}$ by

$$\log T_{A\to B} = \log \alpha_i + \beta_i T^{-1}. \tag{6.38}$$

We provide the estimators for α_i and β_i with a linear regression for $\log(T_{A\to B})$ in Table 6.3. The last column represents the coefficient of determination. We can observe a large

Table 6.3 – Mean-square estimators for $\log \alpha_i$ and β_i given by (6.38).

c_i	$\log(\alpha_i)$	$\log(c_i \cdot \alpha_i)$	β_i	r^2
0.5	2.97	2.27	3.28	0.998963
0.25	2.80	1.41	3.61	0.994054
0.125	2.56	0.58	4.24	0.997072
0.0625	3.42	0.64	4.70	0.997195
0.03125	6.30	2.84	4.02	0.946444

distribution of values for the coefficients β_i that are supposed to be equal for all coefficients c_i', in the small temperature limit (see [29]). Given the rather tight confidence interval given in Figure 6.5, we believe that these discrepancies come from the fact that the temperatures we used in the experiment are not small enough to approximate $T_{A\to B}$ by its asymptotic form. Thus it is furthermore difficult to analyse from this experiment the dependence on the prefactor α_i with respect to the non-reversibility. The most interesting case is the one given in the regime $t \leftarrow tc_i'$, where the dynamics is given by (6.37). This scaling modifies the prefactor by a multiplication by c_i. This product is given in the third column of Table 6.3. It would seem that the prefactor is not monotonic with respect to the intensity of the non-reversibility. Yet, these estimators are, in our opinion, not precise enough to validate this observation.

We discuss now the variance of the AMS estimator of $\mathbb{P}(\tau_B(X^{\nu_A}) = 0)$. We observed that a dominant part of the variance is linked to the variability of the initial conditions in the AMS estimator. This can be observed by noticing that the variance of the AMS estimator of $\mathbb{P}(\tau_B(X^{\nu_A}) = 0)$ is much larger than the variance of the AMS estimator of $\mathbb{P}(\tau_B(X^x) = 0)$ for any initial point $x \in K$. We observed that a small proportion of the initial conditions x sampled by the iterative method explained above leads to much larger probabilities $\mathbb{P}(\tau_B(X^x) = 0)$ than the other ones. In other words, the estimation of the probability of the rare event $\{\tau_B(X^{\nu_A}) = 0\}$ is partially transformed as a problem of searching for the rare initial configurations x that maximize $\mathbb{P}(\tau_B(X^x) = 0)$. We also noticed that the variance of the estimators of $\mathbb{P}(\tau_B(X^{\nu_A}) = 0)$ is reduced when we add the right amount of non-reversibility to the Langevin dynamics. We display in Figure 6.6 the relative standard deviation of these estimators. We can observe that for $c_1' \approx 0.125$,

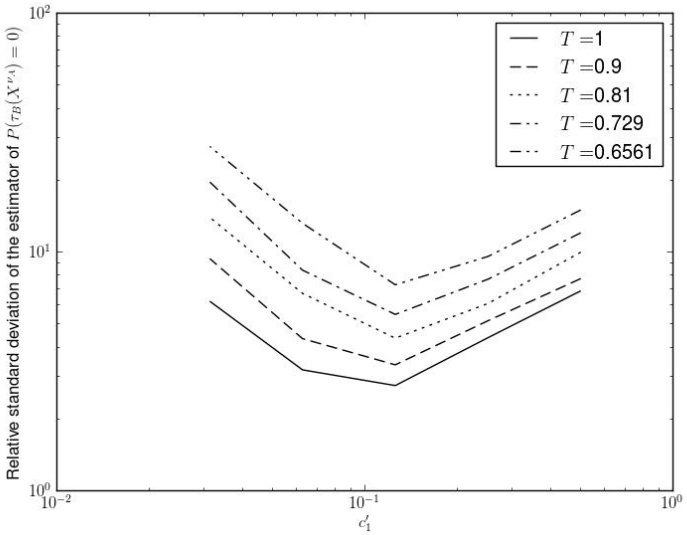

Figure 6.6 – Relative standard deviation of the estimators of $\mathbb{P}\left(\tau_B(X^{\nu_A}) = 0\right)$ with respect to the dissipation c_1' for various temperatures.

the variance is minimized for all the temperatures.

6.5 Conclusion and prospects

We proposed in this chapter a numerical scheme that enables to solve precisely the SPGPE. It would be interesting to justify rigorously the splitting method we used to decompose the two time scales linked to the purely Hamiltonian part and to the dissipative one. Since the aim of this scheme is to sample rare events (that typically occur after long integration times), it would be interesting to study its error for long times integrations.

We also proposed a (biased) method to estimate the transition times between the vortex configurations, inspired from [40]. This method remains very costly in terms of computational time. Nevertheless, it is reasonable to use it for computing the energy gaps between two metastable states. The computation of the prefactor (in the Eyring-Kramers formula) remains, in our application case, quite a numerical challenge. Moreover, we recall that we chose some numerical parameters in Table 6.1 in order to reduce the dimension of the problem, and that we used a much higher dissipation than what is predicted by the SPGPE. In addition, even if the temperatures we chose were of the order of magnitude of those used in physical experiments, they were not low enough to enable to estimate precisely the energy gap between two of the metastable states.

We also noticed that adding a non-reversibility to the Langevin dynamics enables to reduce the variance of the AMS estimators in our case. We do not know if this behaviour is generic to other settings.

As we explained previously, the model (6.4) is a nonreversible Langevin equation for which the non-reversibility depends on the temperature. It raises many questions from the theoretical point of view. For instance can we compute a generalization of the Eyring-Kramers formula that provides an asymptote of the exit time of the basin of attraction? Moreover, can we consider the possibility of simplifying the model in the small temperature limit by averaging on the Hamiltonian dynamics?

From a numerical point of view, the algorithm formulated in Section 6.4.3 requires rigorous justifications to estimate its bias and its consistence in the small temperature limit.

Eventually, to be able to compare this model with the experiments, we would require to study the time evolution of the number of vortices inside the condensate, for different temperatures. Yet, we do not know any references in this direction.

6.6 Appendix

6.6.1 Computation of the nonlinearity

This section is devoted to the explanation of the exact computation of the nonlinearity $F(\phi) = P_K |\phi|^2 \phi$ for all $\phi \in K$. The argument has been used in [30]. The problem consists in computing the quantities $\langle F(\phi), Y_{n,l} \rangle$ for all $(n,l) \in K_{\text{indices}}$ defined by,

$$\langle F(\phi), Y_{n,l} \rangle = \int_{\mathbb{R}_+} \int_{\theta=0}^{2\pi} F(\phi)(r,\theta) \overline{Y_{n,l}}(r,\theta) d\theta r dr. \tag{6.39}$$

We define for all $(n,l) \in K_{\text{ind}}$, the polynomial $Q_{n,l}$ of degree n by

$$Y_{n,l}(r,\theta) = e^{il\theta} e^{-r^2/2} r^{|l|} Q_{n,l}(r^2),$$

where $Y_{n,l}$ is given by (6.7), and define the function $Z_{n,l}$ for all $r \geq 0$ and $\theta \in \mathbb{R}$ by,

$$Y_{n,l}(r,\theta) = e^{il\theta} Z_{n,l}(r).$$

Then, we define the constants l_-, l_+ and $n(l)$ for all $l \in [\![-l_-, l_+]\!]$ by

$$l_\pm = \left\lfloor \frac{\omega_{\max}}{1 \mp \Omega} \right\rfloor, \quad n(l) = \left\lfloor \frac{\omega_{\max} - |l| + \Omega l}{2} \right\rfloor.$$

With this notation the set K_{ind} can be written as

$$K_{\text{ind}} = \{(n,l); -l_- \leq l \leq l_+, 0 \leq n \leq n(l)\}.$$

From now on, we fix ϕ. Then, we define for all $l \in [\![-l_-, l_+]\!]$, the polynomial $Q_l(r)$ of degree at most $n(l)$ by

$$Q_l(r) = \sum_{n=0}^{n(l)} \langle \phi, Y_{n,l} \rangle Q_{n,l}(r).$$

Using this notation we can now write for all $r \geq 0$ and $\theta \in \mathbb{R}$,

$$\phi(r,\theta) = \sum_{l=-l_-}^{l_+} e^{il\theta} r^{|l|} Q_l(r^2) e^{-r^2/2},$$

and it follows that,

$$\begin{aligned} &(|\phi|^2 \phi \overline{Y_{n,l}})(r,\theta) \\ &\quad = \sum_{l_1,l_2,l_3=-l_-}^{l_+} e^{i(l_1-l_2+l_3-l)\theta} r^{|l_1|+|l_2|+|l_3|+|l|} Q_{l_1}(r^2) \overline{Q_{l_2}(r^2)} Q_{l_3}(r^2) Q_{n,l}(r^2) e^{-2r^2}. \end{aligned} \tag{6.40}$$

Integrating this equation with respect to θ cancels all the terms such that $l_1 - l_2 + l_3 - l \neq 0$. Thus we obtain,

$$\begin{aligned} &\int_0^{2\pi} (|\phi|^2 \phi \overline{Y_{n,l}})(r,\theta) d\theta \\ &\quad = 2\pi \sum_{(l_1,l_2,l_3) \in J} r^{|l_1|+|l_2|+|l_3|+|l|} Q_{l_1}(r^2) \overline{Q_{l_2}(r^2)} Q_{l_3}(r^2) Q_{n,l}(r^2) e^{-2r^2}, \end{aligned} \tag{6.41}$$

with $J = \big\{(l_1,l_2,l_3) \in [\![-l_-, l_+]\!]^3; l_1 - l_2 + l_3 - l = 0\big\}$. The key point here is to notice that, for any fixed radius r, the quantity given by Equation (6.41) can be computed explicitely by means of a Fourier transform. To make this point clear, we can notice that the expression given by Equation (6.40) is a power expansion in $e^{i\theta}$, with powers bounded from below by $-2(l_- + l_+)$ and bounded from above by $2(l_- + l_+)$. Thus, the integration with respect to θ in (6.39) can be exactly computed with a Fourier transform of degree $2(l_- + l_+) + 1$. Thus, we set $(\theta_m)_{0 \leq m \leq 2(l_- + l_+)}$ the uniform subdivision of $[0, 2\pi]$ given for $0 \leq m \leq 2(l_- + l_+)$ by $\theta_m = 2\pi m / (2(l_- + l_+) + 1)$. Doing so, we obtain for all $r \geq 0$,

$$\int_0^{2\pi} (|\phi|^2 \phi \overline{Y_{n,l}})(r,\theta) d\theta = \frac{2\pi Z_{n,l}(r)}{2(l_- + l_+) + 1} \sum_{m=0}^{2(l_- + l_+)} e^{-il\theta_m} (|\phi|^2 \phi)(r, \theta_m).$$

Then, noticing that the nullity of $l_1 - l_2 + l_3 - l$ implies that $|l_1| + |l_2| + |l_3| + |l|$ is even, we deduce that the above expression can be written as a polynomial of degree at most $(|l_1| + |l_2| + |l_3| + |l|)/2 + n(l_1) + n(l_2) + n(l_3) + n(l)$ (which is upper bounded by $2l_+$) in r^2. Thus, the integral in r that appears in Equation (6.39) can also be computed

exactly. To do so, we perform the following change of variable,

$$\langle F(\phi), Y_{n,l}\rangle = \frac{1}{2}\int_{\mathbb{R}_+}\int_{\theta=0}^{2\pi} F(\phi)(\sqrt{r},\theta)\overline{Y_{n,l}}(\sqrt{r},\theta)d\theta dr.$$

According to Equation (6.41), the integrand (for the integral with respect to r) is a polynomial of degree at most $2l_+$ with the measure $e^{-2r}dr$, and thus can be exactly computed using a Gauss-Laguerre quadrature of order $l_+ + 1$. More precisely, by setting ω_p and r_p respectively the weights and the roots of a Gauss-Laguerre quadrature of order $l_+ + 1$ that we recall therafter, we obtain,

$$\langle F(\phi), Y_{n,l}\rangle = \sum_{p=0}^{l_+} \frac{\omega_p e^{r_p}\pi}{2}\frac{Z_{n,l}(r_p)}{2(l_- + l_+)+1}\sum_{m=0}^{2(l_-+l_+)} e^{-il\theta_m}(|\phi|^2\phi)(\sqrt{r_p}/2,\theta_m).$$

The nodes and the weights of this quadrature are chosen such that for all polynomial P of degree at most $2l_+$, $\int_{\mathbb{R}^+} P(r)e^{-r}dr = \sum_{p=0}^{l_+}\omega_p P(r_p)$, and the r_p are the roots of the Laguerre polynomial L_{l_+}, and the weights ω_p are given by,

$$\omega_p = \frac{r_p}{(p+1)^2[L_{l_++1}(r_p)]^2}.$$

6.6.2 Approximation of the phase in the neighborhood of a vortex

The objective of this section is to compute an approximation of the phase of the solution of Equation (6.4) around its vortices. It will be used to design a practical algorithm to precisely locate these vortices in section 6.6.3. The result of this section is given in Proposition 6.20. It is only a formal result because we do not make precise in what sense the approximation holds and because the proof involves formal arguments. This result is only given for critical points on the energy E defined over K by Equation (6.5). Nevertheless, this result seems to be accurate enough in practice, even for the solution of Equation (6.4), to provide satisfactory results for the vortex localisation algorithm. In this section only, we denote the vectors of $\mathbb{R}^2$ in bold font. We denote by $\mathbf{x}$ the cartesian space variable. We denote by $\cdot$ the canonical scalar product in $\mathbb{R}^2$, and define the cross product $\times$ for all $(\mathbf{a},\mathbf{b}) \in (\mathbb{R}^2)^2$ by $\mathbf{a}\times\mathbf{b} = a_x b_y - a_y b_x$, where $\mathbf{a} = (a_x, a_y)$ and $\mathbf{b} = (b_x, b_y)$.

Let u be a critical point of the energy E defined on K that exhibits a vortex centred at $\mathbf{a}$ of degree n. We recall that, in the two dimensional case, vortices are singularities of the argument, but not of the complex field (the wave function). These singularities are characterized by their degrees, which are formally the number of turns of the phase around the singularity. We are looking for an approximation of u on a small neighborhood of $\mathbf{a}$. It is then natural to introduce the change of variable

$$\mathbf{z} = \mathbf{x} - \mathbf{a}, \quad v(\mathbf{z}) = u(\mathbf{x}). \tag{6.42}$$

We set $\mathbf{z} = (z_x, z_y)$, and introduce the polar coordinates (r,θ) for the variable $\mathbf{z}$, so that

$(z_x, z_y) = (r\cos(\theta), r\sin(\theta))$.

Proposition 6.20. *If u is a critical point of the energy E defined on K that exhibits a vortex centred at $\mathbf{a}$ of degree n, then an approximation of u in the neighbourhood of $\mathbf{a}$, or equivalently an approximation of v in the neighbourhood of 0, is given by,*

$$v(r,\theta) \approx r^{|n|} \exp\left(i\left(\theta + r\Omega\,|\mathbf{a}|\sin(\theta - \arg(\mathbf{a}))\right)\right). \tag{6.43}$$

This approximation is obtained using an asymptotic expansion of v in x.

Formal proof of Proposition 6.20. We begin by defining an asymptotic expansion of v by setting $v = R(r,\theta)e^{i\eta(r,\theta)}$, and by looking for an approximation of v (in the limit where r vanishes) in the form

$$\begin{aligned} R(r,\theta) &= r^{|n|}(1 + r^2 h_2(\theta) + r^4 h_4(\theta) + \ldots), \\ \eta(r,\theta) &= n\theta + r g_1(\theta) + r^2 g_2(\theta) + \ldots. \end{aligned} \tag{6.44}$$

The term of order one follows from the fact that we suppose the vortex in $\mathbf{a}$ to be of degree n. The dominant part in the expansion of R is taken this way to be consistent with the degree of the vortex. It enables to keep a bounded energy gradient in a neighborhood of the vortex. We recover this fact in the following. The idea of the proof is now to plug this expression in the gradient of the energy E (defined by Equation (6.5)), and to express this quantity as an asymptotic expansion in r. We will see this way that a good choice of g_1 enables to cancel the first order of this expansion.

We begin by performing the change of variable given by Equation (6.42) on the gradient of the energy E given by Equation (6.5). We define $L_{\mathbf{a}} = z_x\partial_y - z_y\partial_x$, and obtain,

$$\nabla E(v) = \left(\frac{1}{2}(-\Delta + |\mathbf{z}|^2) - \mu + i\Omega L_{\mathbf{a}} + g\,|v|^2\right) v + \frac{1}{2}(2\mathbf{z}\cdot\mathbf{a} + |\mathbf{a}|^2)v + i\Omega(\mathbf{a}\times\nabla)v. \tag{6.45}$$

We write now Equation (6.45) in polar coordinates. To do so, we recall the following identities,

$$\begin{aligned} \Delta &= \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}, \\ L_a &= \frac{\partial}{\partial\theta}, \\ \mathbf{a}\times\nabla &= \frac{\mathbf{a}\times\mathbf{z}}{r}\frac{\partial}{\partial r} + \frac{\mathbf{a}\cdot\mathbf{z}}{r^2}\frac{\partial}{\partial\theta}. \end{aligned}$$

Thus,

$$\begin{aligned}\nabla E(v) = &\left(-\frac{1}{2}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)+\frac{1}{2}r^2-\mu+i\Omega\frac{\partial}{\partial\theta}+g\left|v^2\right|\right)v\\ &+\frac{1}{2}(2r\left|\mathbf{a}\right|\cos(\theta-\arg(\mathbf{a}))-\left|\mathbf{a}\right|^2)v\\ &+i\Omega\left(+\left|\mathbf{a}\right|\sin(\theta-\arg(\mathbf{a}))\frac{\partial}{\partial r}+\frac{1}{r}\left|\mathbf{a}\right|\cos(\theta-\arg(\mathbf{a}))\frac{\partial}{\partial\theta}\right)v.\end{aligned} \tag{6.46}$$

We now use the asymptotic expansion of v proposed in Equation (6.44). The computation of the derivatives of v give,

$$\begin{aligned}\frac{\partial v}{\partial r} &= \left(\left|n\right|r^{|n|-1}+ig_1r^{|n|}+((|n|+2)h_2+2g_2)r^{|n|+1}+\dots\right)e^{i\eta(r,\theta)},\\ \frac{\partial^2 v}{\partial r^2} &= \left(\left|n\right|(|n|-1)r^{|n|-2}+2i\left|n\right|g_1r^{|n|-1}+\dots\right)e^{i\eta(r,\theta)},\\ \frac{\partial v}{\partial\theta} &= \left(inr^{|n|}+i\dot{g}_1r^{|n|+1}+(\dot{h}_1+inh_2+i\ddot{g}_2)r^{|n|+2}+\dots\right)e^{i\eta(r,\theta)},\\ \frac{\partial^2 v}{\partial\theta^2} &= \left(-\left|n\right|^2r^{|n|}+(i\ddot{g}_1-2n\dot{g}_1)r^{|n|+1}+\dots\right)e^{i\eta(r,\theta)},\end{aligned}$$

By plugging these expressions into Equation (6.46) we obtain,

$$\begin{aligned}e^{-i\eta(r,\theta)}\nabla E(v) = &\, r^{|n|-2}\frac{1}{2}\left(\left|n\right|(|n|-1)+|n|-n^2\right)\\ &+r^{|n|-1}\left(n\dot{g}_1-\Omega n\left|\mathbf{a}\right|\cos(\theta-\arg(\mathbf{a}))\right)\\ &+ir^{|n|-1}\left(\frac{1}{2}[-(2\left|n\right|+1)g_1-\ddot{g}_1]+\Omega\left|n\right|\left|\mathbf{a}\right|\sin(\theta-\arg(\mathbf{a}))\right)\\ &+\dots\end{aligned} \tag{6.47}$$

We can notice that the term of order $r^{|n|-2}$ vanishes, and that the term of order $r^{|n|-1}$ vanishes if and only if,

$$\begin{cases}\dot{g}_1=\dfrac{\Omega\left|\mathbf{a}\right|}{2}\cos(\theta-\arg(\mathbf{a})),\\ \ddot{g}_1+(2\left|n\right|+1)g_1=2\Omega\left|n\right|\left|\mathbf{a}\right|\sin(\theta-\arg(\mathbf{a})).\end{cases} \tag{6.48}$$

Equation (6.48) has a unique solution given by,

$$g_1(\theta)=\Omega\left|\mathbf{a}\right|\sin(\theta-\arg(\mathbf{a})). \tag{6.49}$$

The approximation (6.43) is the truncation of the expansion (6.44) of v at first order in amplitude and phase obtained by plugging Equation (6.49) in Equation (6.44). □

6.6.3 Practical vortex localisation in rotating BEC

The objective of this part is to locate a vortex of degree $n = \pm 1$ in a phase field ϕ. We recall that the phase of the wave function is defined as the imaginary part of the complex logarithm of the complex field. Yet, we point out that because the complex exponential is 2π periodic, one has to make precise the meaning of the logarithm. In practice, the phase ϕ is defined as the imaginary part of the principal value of the logarithm, which amounts to supposing that ϕ takes values in $(-\pi, \pi]$. Then, this function is not continuous outside all neighbourhoods of the singularity. More precisely the phase along all closed paths that strictly contain the singularity might *jump* between $-\pi$ and π. One way to define a continuous phase along all these closed paths, is classically to define the logarithm on a Riemann surface which is constructed by gluing all the branches of the complex logarithm. This way, it is also possible to define a continuous gradient of the phase, denoted in the following by $\nabla\phi$, defined outside the singularity. Therefore, one way to locate the phase singularities is to perform an integration of the gradient on an arbitrary closed contour whose interior contains the vortex. The result gives the number of vortices inside the contour, counted with their degree n, up to a factor 2π,

$$\oint_{\partial P} \nabla\phi(\mathbf{a}) \cdot \mathbf{da} = 2\pi n. \tag{6.50}$$

This gives a way to detect the existence of a vortex inside the condensate.

In practice, we can only compute discrete versions of these integrals. More precisely, let $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}$ be the n vertices of a polygon P. We denote by ∂P the closed contour formed by its boundary. Let $\phi(\mathbf{a_i})$ be the principal value of the phase at vertex $\mathbf{a_i}$. Then, we approximate the integral of the (continuous) gradient of the phase along the boundary of the polygon in the following way,

$$\oint_{\partial P} \nabla\phi(\mathbf{a}) \cdot \mathbf{da} \approx \sum_{i=1}^{n} \mathrm{pv}(\phi(\mathbf{a_{i+1}}) - \phi(\mathbf{a_i})), \tag{6.51}$$

where $\mathbf{a_{n+1}} = \mathbf{a_n}$, and for all $x \in \mathbb{R}$, $\mathrm{pv}(x)$ denotes the only real number belonging to $]-\pi, \pi]$ and such that x and $\mathrm{pv}(x)$ are congruent modulo 2π. This limitation implies that we may falsely detect some vortices, or on the contrary miss some of them.

In practice, as a by-product of our numerical scheme, we can get at each time step the value of the phase function on each nodes of the Fourier and Laguerre quadratures used in the numerical scheme. These nodes define a polar grid of the disk centred at $\mathbf{0}$, whose radius is the highest node of the Laguerre quadrature, and whose cells take the form of ring sectors.A first approximation can be obtained by computing the approximate circulation of the gradient of the phase on the boundaries of each of the cells, given by the right-hand side of equation (6.51) where the $(\mathbf{a_i})_{1\leq n}$ are the (four) vertices that define the cells. The computational cost of this operation is linear in the number of cells, and thus

is negligible with respect to the cost associated with the computation of the nonlinearity that occurs at each time step.

Yet, this method is not sufficient for two reasons. First we need to avoid missing (or misdetecting) a vortex. This may impact greatly the AMS algorithm with our choice of reaction coordinate. For instance, suppose that we are interested in the exit of the nth vortex from the inside of the condensate. Then, we choose as a reaction coordinate the distance (from the centre of the trap) of the nth most distant vortex. In this case, missing a vortex may lead to an overestimated measure of this distance, and thus the trajectory may be considered falsely as a reactive trajectory. Moreover, we explained in Remark 6.13 that a discrete-valued reaction coordinate may lead to a big variance of the AMS estimator. Thus, we require a more precise computation of the positions of the vortices. In practice we propose to refine the computation of the wave function on a thinner mesh of the cells for which the circulation does not vanish. Actually, we may also include in this mesh a neighbouring cell if we have a doubt about the precise location of the vortex (as it will be explained in the following). Moreover, this method enables to detect if a cell does not actually contain a vortex.

If we denote by $\mathbf{z}$ and (r, θ) the affix and the polar coordinates of a same point, then the wave function ψ around a vortex centred at $\mathbf{a}$ of index m is approximated (up to a uniform change of phase) at first order by,

$$\psi(\mathbf{z}) \approx \tilde{\psi}(\mathbf{z} - \mathbf{a}), \quad \text{where} \quad \tilde{\psi}(r, \theta) = re^{im\theta}.$$

We also denote by $\phi(\mathbf{z})$ the principal value of the phase of $\psi(\mathbf{z})$. Suppose that the wave function ψ is exactly given by the above approximation, and that its index is equal to one. Then, the principal value $\mathrm{pv}(\phi(\mathbf{a_{i+1}}) - \phi(\mathbf{a_i}))$ that appears in (6.51) is exactly the angle $\widehat{\mathbf{a_{i+1}aa_i}}$, and the approximation (6.51) is actually exact. Typical errors in the approximation of the circulation of the gradient of the phase occur when the vortex is close to the boundary, as shown in Figure 6.7. In this figure, two cells are represented in continuous lines, and the vortex is supposed to be at position $\mathbf{a}$. The dash dotted lines in the lower cells represent the iso-values of the phase that intersect the phase of ψ at the vertices $\mathbf{a_0}$, $\mathbf{a_1}$, $\mathbf{a_4}$ and $\mathbf{a_5}$ of the lower cell. They are not represented as lines since the wave function is supposed to be slightly perturbed from its first order approximation. It is clear in this example that the phase difference between $\phi(\mathbf{a_4}) - \phi(\mathbf{a_1})$, which is equal to the angle α is greater than π, and its principal value is negative. Thus, the circulation around the lower cell will be found to be equal to zero in this case. Moreover, it is likely that the circulation around the upper cell will be found to be equal to 2π since the principal value $\mathrm{pv}(\phi(\mathbf{a_1}) - \phi(\mathbf{a_4}))$ is actually equal to the phase difference $\phi(\mathbf{a_1}) - \phi(\mathbf{a_4})$. Another way of seeing this is to say that all the points such that the angle α is equal to $\widehat{\mathbf{a_4aa_1}}$ belong to the upper cell. For instance, the point $\mathbf{a'}$ is a possible position of such a point. Morally, the more "curved" the isovalues of the phase are, and the more likely the localisation error is.

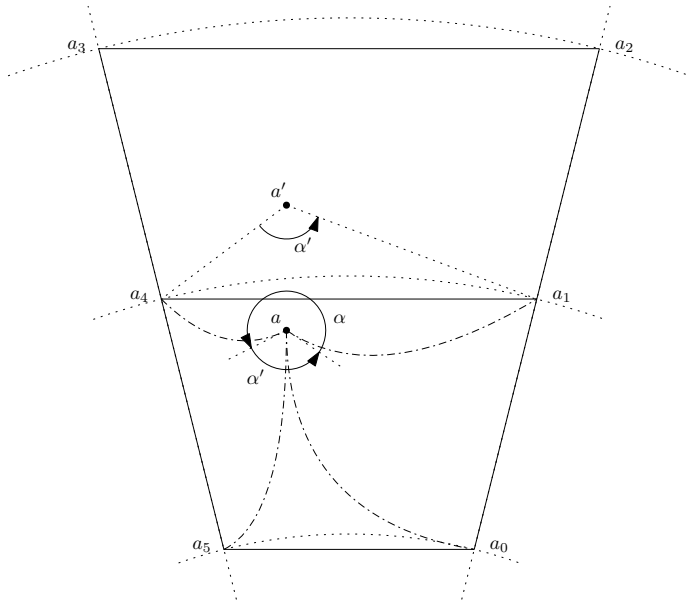

Figure 6.7 – A typical problem of vortex misdetection when the vortex is close to the boundary of a cell.

In the light of the previous observations, we propose two improvements for the algorithm. First, we propose to include the neighbouring cell in the refinement process if a principal value that appears in the right-hand side of (6.51) is close to $-\pi$ or π. Then, we propose to make use of an approximation of the phase of the wave function in the neighbourhood of the vortices given by Equation (6.43), to improve the approximation (6.51). This approximation states that an approximation of the angle $\alpha = \widehat{\mathbf{a_{i+1}aa_i}}$ is given by

$$\alpha = \phi(\mathbf{a_{i+1}}) - \phi(\mathbf{a_i}) - \Omega\mathbf{a} \times (\mathbf{a_{i+1}} - \mathbf{a_i}).$$

Thus, we replace the approximation given by Equation (6.51) by the following approximation,

$$\oint_{\partial P} \nabla\phi(\mathbf{a}) \cdot \mathbf{da} \approx \sum_{i=1}^{n} \mathrm{pv}(\phi(\mathbf{a_{i+1}}) - \phi(\mathbf{a_i}) - \Omega\mathbf{a}' \times (\mathbf{a_{i+1}} - \mathbf{a_i})), \tag{6.52}$$

where $\mathbf{a}'$ denotes the centre of the cell, or any approximation of the position of the vortex. This last approximation is motivated by the fact that the typical length of the cell $\|\mathbf{a_{i+1}} - \mathbf{a_i}\|$ is supposed to be small. Thus, if the error of the approximation $\mathbf{a}'$ of the position $\mathbf{a}$ is of order of the size of the cell, then the error is of order $\|(\mathbf{a_{i+1}} - \mathbf{a_i})\|^2$, which should be at least the same as the error associated with the approximation given by Equation (6.43).

Bibliography

[1] F. Kh. Abdullaev, B. B. Baizakov, and V. V. Konotop, *Dynamics of a Bose-Einstein Condensate in Optical Trap*, Nonlinearity and Disorder: Theory and Applications (Fatkhulla Abdullaev, Ole Bang, and Mads Peter Sørensen, eds.), NATO Science Series, vol. 45, Springer Netherlands, 2001, pp. 69–78.

[2] F. Kh. Abdullaev, J. C. Bronski, and G. Papanicolaou, *Soliton perturbations and the random Kepler problem*, Phys. D **135** (2000), no. 3-4, 369–386.

[3] J. R. Abo-Shaeer, C. Raman, J. M. Vogels, and W. Ketterle, *Observation of Vortex Lattices in Bose-Einstein condensates*, Science **292** (2001), no. 5516, 476–479.

[4] A. Agarwal, S. De Marco, E. Gobet, and G. Liu, *Rare event simulation related to financial risks: efficient estimation and sensitivity analysis*, preprint, October 2015.

[5] J. O. Andersen, *Theory of the weakly interacting Bose gas*, Rev. Mod. Phys. **76** (2004), 599–639.

[6] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor*, Science **269** (1995), no. 5221, 198–201.

[7] X. Antoine, W. Bao, and C. Besse, *Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations*, Computer Physics Communications **184** (2013), no. 12, 2621 – 2633.

[8] X. Antoine and R. Duboscq, *Gpelab, a Matlab toolbox to solve Gross–Pitaevskii equations II: Dynamics and stochastic simulations*, Computer Physics Communications **193** (2015), 95 – 117.

[9] M. Banterle, C. Grazian, A. Lee, and C. P. Robert, *Accelerating Metropolis-Hastings algorithms by delayed acceptance*, arXiv preprint arXiv:1503.00996 (2015).

[10] W. Bao and Y. Cai, *Uniform error estimates of finite difference methods for the nonlinear Schrödinger equation with wave operator*, SIAM Journal on Numerical Analysis **50** (2012), no. 2, 492–521.

[11] ———, *Optimal error estimates of finite difference methods for the Gross-Pitaevskii equation with angular momentum rotation*, Mathematics of Computation **82** (2013), no. 281, 99–128.

[12] W. Bao and Q. Du, *Computing the ground state solution of Bose-Einstein condensates by a normalized gradient flow*, SIAM J. Sci. Comput. **25** (2004), no. 5, 1674–1697.

[13] W. Bao, D. Jaksch, and P. A. Markowich, *Numerical solution of the Gross–Pitaevskii equation for Bose–Einstein condensation*, Journal of Computational Physics **187** (2003), no. 1, 318 – 342.

[14] W. Bao, H Li, and J. Shen, *A Generalized-Laguerre–Fourier–Hermite Pseudospectral Method for Computing the Dynamics of Rotating Bose–Einstein Condensates*, SIAM Journal on Scientific Computing **31** (2009), no. 5, 3685–3711.

[15] W. Bao and J. Shen, *A fourth-order time-splitting Laguerre–Hermite pseudospectral method for Bose–Einstein condensates*, SIAM Journal on Scientific Computing **26** (2005), no. 6, 2010–2028.

[16] W. Bao and Y. Zhang, *Dynamics of the group state and central vortex states in Bose-Einstein condensation*, Math. Models Methods Appl. Sci. **15** (2005), no. 12, 1863–1896.

[17] C. Bardos, F. Golse, and N. J. Mauser, *Weak coupling limit of the N-particle Schrödinger equation*, Methods Appl. Anal. **7** (2000), no. 2, 275–293, Cathleen Morawetz: a great mathematician.

[18] M. D. Barrett, J. A. Sauer, and M. S. Chapman, *All-Optical Formation of an Atomic Bose-Einstein Condensate*, Phys. Rev. Lett. **87** (2001), 010404.

[19] J.-L. Basdevant and J. Dalibard, *Mécanique quantique*, Cours de l'Ecole Polytechnique, 2002.

[20] R. Belaouar, A. de Bouard, and A. Debussche, *Numerical analysis of the nonlinear Schrödinger equation with white noise dispersion*, Stochastic Partial Differential Equations: Analysis and Computations **3** (2015), no. 1, 103–132.

[21] F. A. Berezin and M. A. Shubin, *The Schrödinger Equation*, Mathematics and its Applications (Soviet Series), vol. 66, Kluwer Academic Publishers Group, Dordrecht, 1991.

[22] C. Besse, G. Dujardin, and I. Lacroix-Violet, *High order exponential integrators for nonlinear Schrödinger equations with application to rotating Bose-Einstein condensates*, SIAM Journal on Numerical Analysis **55** (2017), no. 3, 1387–1411.

[23] J. Bierkens, *Non-reversible Metropolis-Hastings*, Statistics and Computing (2015), 1–16.

[24] J. Bierkens, P. Fearnhead, and G. Roberts, *The Zig-Zag process and super-efficient sampling for Bayesian analysis of Big Data*, arXiv preprint arXiv:1607.03188 (2016).

[25] P. B. Blakie, A. S. Bradley, M. J. Davis, R. J. Ballagh, and C. W. Gardiner, *Dynamics and statistical mechanics of ultra-cold Bose gases using c-field techniques*, Advances in Physics **57** (2008), no. 5, 363–455.

[26] S. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Zeitschrift für Physik **26** (1924), no. 1, 178–181.

[27] N. Bou-Rabee and E. Vanden-Eijnden, *Pathwise accuracy and ergodicity of metropolized integrators for SDEs*, Comm. Pure Appl. Math. **63** (2010), no. 5, 655–696.

[28] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet, *The bouncy particle sampler: A non-reversible rejection-free Markov chain Monte Carlo method*, Journal of the American Statistical Association (2017).

[29] F. Bouchet and J. Reygner, *Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes*, Annales Henri Poincaré **17** (2016), no. 12, 3499–3532.

[30] A. S. Bradley, C. W. Gardiner, and M. J. Davis, *Bose-Einstein condensation from a rotating thermal cloud: Vortex nucleation and lattice formation*, Phys. Rev. A **77** (2008), 033616.

[31] C. C. Bradley, C. A. Sackett, J. J. Tollett, and R. G. Hulet, *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions*, Phys. Rev. Lett. **75** (1995), 1687–1690.

[32] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre, and M. Rousset, *Unbiasedness of some generalized Adaptive Multilevel Splitting algorithms*, arXiv preprint arXiv:1505.02674 (2015).

[33] ———, *Unbiasedness of some generalized adaptive multilevel splitting algorithms*, Ann. Appl. Probab. **26** (2016), no. 6, 3559–3601.

[34] V. Bretin, P. Rosenbusch, F. Chevy, G. V. Shlyapnikov, and J. Dalibard, *Quadrupole oscillation of a single-vortex Bose-Einstein condensate: Evidence for Kelvin modes*, Phys. Rev. Lett. **90** (2003), 100403.

[35] E. Cancès, C. Le Bris, and Y. Maday, *Méthodes mathématiques en chimie quantique. Une introduction*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 53, Springer, Berlin, 2006.

[36] E. Cancès, F. Legoll, and G. Stoltz, *Theoretical and numerical comparison of some sampling methods for molecular dynamics*, ESAIM: Mathematical Modelling and Numerical Analysis **41** (2007), no. 2, 351–389.

[37] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*, Springer Series in Statistics, Springer, New York, 2005.

[38] T. Cazenave, *Semilinear Schrödinger Equations*, Courant Lecture Notes in Mathematics, vol. 10, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003.

[39] F. Cérou and A. Guyader, *Adaptive Multilevel Splitting for rare event analysis*, Stochastic Analysis and Applications **25** (2007), no. 2, 417–443.

[40] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier, *A multiple replica approach to simulate reactive trajectories*, The Journal of Chemical Physics **134** (2011), no. 5, 054108.

[41] S.-M. Chang, S. Gustafson, K. Nakanishi, and T.-P. Tsai, *Spectra of linearized operators for NLS solitary waves*, SIAM J. Math. Anal. **39** (2007/08), no. 4, 1070–1111.

[42] C. Chen, J. Hong, and A. Prohl, *Convergence of a ϑ-scheme to solve the stochastic nonlinear Schrödinger equation with Stratonovich noise*, Stochastic Partial Differential Equations: Analysis and Computations (2015), 1–45.

[43] F. Chen, L. Lovász, and I. Pak, *Lifting Markov chains to speed up mixing*, Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '99, ACM, 1999, pp. 275–281.

[44] C. Cohen-Tannoudji, *Nobel lecture: Manipulating atoms with photons*, Rev. Mod. Phys. **70** (1998), 707–719.

[45] C. Cohen-Tannoudji, J. Dalibard, and F. Laloë, *La condensation de Bose-Einstein dans les gaz*, Einstein aujourd'hui, EDP Sciences et CNRS Editions, 2005, pp. 87–127.

[46] P. Collet, S. Martínez, and J. San Martín, *Quasi-Stationary Distributions*, Springer Science & Business Media, 2012.

[47] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Applied Mathematical Sciences, Springer Berlin Heidelberg, 1997.

[48] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, *Theory of Bose-Einstein condensation in trapped gases*, Rev. Mod. Phys. **71** (1999), 463–512.

[49] J. Dalibard, *Statistique de Bose–Einstein et condensation*, Cours du collège de France, 2016.

[50] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Bose-Einstein condensation in a gas of sodium atoms*, Phys. Rev. Lett. **75** (1995), 3969–3973.

[51] A. De Bouard and A. Debussche, *A semi-discrete scheme for the stochastic nonlinear Schrödinger equation*, Numerische Mathematik **96** (2004), no. 4, 733–770.

[52] A. de Bouard and A. Debussche, *Random modulation of solitons for the stochastic Korteweg-de Vries equation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **24** (2007), no. 2, 251–278.

[53] ———, *Soliton dynamics for the Korteweg-de Vries equation with multiplicative homogeneous noise*, Electron. J. Probab. **14** (2009), no. 58, 1727–1744.

[54] A. de Bouard and R. Fukuizumi, *Stochastic fluctuations in the Gross–Pitaevskii equation*, Nonlinearity **20** (2007), no. 12, 2823.

[55] ———, *Modulation analysis for a stochastic NLS equation arising in Bose-Einstein condensation*, Asymptot. Anal. **63** (2009), no. 4, 189–235.

[56] ———, *Representation formula for stochastic Schrödinger evolution equations and applications*, Nonlinearity **25** (2012), no. 11, 2993.

[57] A. de Bouard, R. Fukuizumi, and R. Poncet, *Vortex solutions in Bose-Einstein condensation under a trapping potential varying randomly in time*, Discrete and Continuous Dynamical Systems - Series B **20** (2015), no. 9, 2793–2817.

[58] A. de Bouard and E. Gautier, *Exit problems related to the persistence of solitons for the Korteweg-de Vries equation with small noise*, Discrete Contin. Dyn. Syst. **26** (2010), no. 3, 857–871.

[59] A. del Campo, A. Retzker, and M. B. Plenio, *The inhomogeneous Kibble–Zurek mechanism: vortex nucleation during Bose–Einstein condensation*, New Journal of Physics **13** (2011), no. 8, 083022.

[60] M. Delfour, M. Fortin, and G. Payr, *Finite-difference solutions of a non-linear Schrödinger equation*, Journal of Computational Physics **44** (1981), no. 2, 277 – 288.

[61] C. Dellago and P. G. Bolhuis, *Transition path sampling and other advanced simulation techniques for rare events*, pp. 167–233, Springer Berlin Heidelberg, 2009.

[62] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010.

[63] J. Denschlag, J. E. Simsarian, D. L. Feder, Charles W. Clark, L. A. Collins, J. Cubizolles, L. Deng, E. W. Hagley, K. Helmerson, W. P. Reinhardt, S. L. Rolston, B. I. Schneider, and W. D. Phillips, *Generating solitons by phase engineering of a Bose-Einstein condensate*, Science **287** (2000), no. 5450, 97–101.

[64] G. Di Gesu, T. Lelièvre, D. Le Peutrec, and B. Nectoux, *Jump Markov models and transition state theory: the quasi-stationary distribution approach*, Faraday Discuss. **195** (2016), 469–495.

[65] L. Di Menza, *Numerical computation of solitons for optical systems*, M2AN Math. Model. Numer. Anal. **43** (2009), no. 1, 173–208.

[66] P. Diaconis, S. Holmes, and R. M. Neal, *Analysis of a nonreversible Markov chain sampler*, Ann. Appl. Probab. **10** (2000), no. 3, 726–752.

[67] B. Diu, C. Guthmann, D. Lederer, and B. Roulet, *Elements de Physique Statistique*, Collection enseignements des sciences, 37, 1989, 2001.

[68] D. Down, S. P. Meyn, and R. L. Tweedie, *Exponential and uniform ergodicity of Markov processes*, Ann. Probab. **23** (1995), no. 4, 1671–1691.

[69] J. R. Driscoll, D. M. Healy, Jr., and D. N. Rockmore, *Fast Discrete Polynomial Transforms with applications to data analysis for distance transitive graphs*, SIAM J. Comput. **26** (1997), no. 4, 1066–1099.

[70] S. Duane, A. D. Kennedy, B. J. Pendleton, and R. Duncan, *Hybrid Monte Carlo*, Physics Letters B **195** (1987), no. 2, 216 – 222.

[71] R. Duboscq and R. Marty, *Analysis of a splitting scheme for a class of random nonlinear partial differential equations*, ESAIM: PS **20** (2016), 572–589.

[72] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis, *Variance reduction using nonreversible Langevin samplers*, Journal of Statistical Physics **163** (2016), no. 3, 457–491.

[73] M. Eichhorn, M. Mudrich, and M. Weidemüller, *Optical dipole trap inside a laser resonator*, Opt. Lett. **29** (2004), no. 10, 1147–1149.

[74] L. Erd˝ os, B. Schlein, and H.-T. Yau, *Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems*, Invent. Math. **167** (2007), no. 3, 515–614.

[75] M. Fathi and G. Stoltz, *Improving dynamical properties of metropolized discretizations of overdamped Langevin dynamics*, Numerische Mathematik **136** (2017), no. 2, 545–602.

[76] M. Fauquembergue, *Réalisation d'un dispositif de condensation de Bose–Einstein et de transport d'un echantillon cohérent d'atomes*, Ph.D. thesis, Universite Paris Sud - Paris XI, 2004.

[77] G. Fibich and N. Gavish, *Theory of singular vortex solutions of the nonlinear Schrödinger equation*, Phys. D **237** (2008), no. 21, 2696–2730.

[78] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer, Heidelberg, 2012.

[79] J. Fröhlich, S. Gustafson, B. L. G. Jonsson, and I. M. Sigal, *Solitary wave dynamics in an external potential*, Comm. Math. Phys. **250** (2004), no. 3, 613–642.

[80] C. W. Gardiner, J. R. Anglin, and T. I. A. Fudge, *The stochastic Gross–Pitaevskii equation*, Journal of Physics B: Atomic, Molecular and Optical Physics **35** (2002), no. 6, 1555.

[81] C. W. Gardiner and M. J. Davis, *The stochastic Gross–Pitaevskii equation: II*, Journal of Physics B: Atomic, Molecular and Optical Physics **36** (2003), no. 23, 4731.

[82] J. Garnier, F. Kh. Abdullaev, and B. B. Baizakov, *Collapse of a Bose–Einstein condensate induced by fluctuations of the laser intensity*, Phys. Rev. A **69** (2004), 053607.

[83] M. Gazeau, *Analyse de modèles mathématiques pour la propagation de la lumière dans les fibres optiques en présence de biréfringence aléatoire*, Ph.D. thesis, 2012.

[84] ———, *Probability and pathwise order of convergence of a semidiscrete scheme for the stochastic Manakov equation*, SIAM J. Numer. Anal. **52** (2014), no. 1, 533–553.

[85] M. E. Gehm, K. M. O'Hara, T. A. Savard, and J. E. Thomas, *Dynamics of noise-induced heating in atom traps*, Phys. Rev. A **58** (1998), 3914–3921.

[86] J. Ginibre and G. Velo, *On the global Cauchy problem for some nonlinear Schrödinger equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **1** (1984), no. 4, 309–323.

[87] E. Gobet and G. Liu, *Rare event simulation using reversible shaking transformations*, SIAM J. Sci. Comput. **37** (2015), no. 5, A2295–A2316.

[88] D. G. Greif, *Evaporative cooling and Bose–Einstein condensation of Rb-87 in a moving–coil TOP trap geometry*, Ph.D. thesis, 2010.

[89] E. P. Gross, *Structure of a quantized vortex in boson systems*, Il Nuovo Cimento (1955-1965) **20** (1961), no. 3, 454–477.

[90] H. Haario, E. Saksman, and J. Tamminen, *An adaptive Metropolis algorithm*, Bernoulli **7** (2001), no. 2, 223–242.

[91] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, New York, 2002.

[92] ———, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer, Dordrecht, 2006.

[93] B. Helffer and F. Nier, *Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach: the case with boundary*, (2006), 1–89.

[94] K. Hukushima and Y. Sakai, *An irreversible Markov-chain Monte Carlo method with skew detailed balance conditions*, Journal of Physics: Conference Series **473** (2013), no. 1, 012012.

[95] D. Hutchinson, E. Zaremba, and A. Griffin, *Finite temperature excitations of a trapped Bose gas*, Physical Review Letters **78** (1997), no. 10.

[96] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu, *Accelerating Gaussian diffusions*, Ann. Appl. Probab. **3** (1993), no. 3, 897–913.

[97] ———, *Accelerating diffusions*, Ann. Appl. Probab. **15** (2005), no. 2, 1433–1444.

[98] C.-R. Hwang, R. Normand, and S.-J. Wu, *Variance reduction for diffusions*, Stochastic Processes and their Applications **125** (2015), no. 9, 3522 – 3540.

[99] J. Iaia and H. Warchall, *Nonradial solutions of a semilinear elliptic equation in two dimensions*, J. Differential Equations **119** (1995), no. 2, 533–558.

[100] B. Jackson, N. P. Proukakis, C. F. Barenghi, and E. Zaremba, *Finite-temperature vortex dynamics in Bose–Einstein condensates*, Phys. Rev. A **79** (2009), 053615.

[101] B. Jackson and E. Zaremba, *Modeling Bose–Einstein condensed gases at finite temperatures with N–body simulations*, Phys. Rev. A **66** (2002), 033606.

[102] A. Jentzen and P. Kloeden, *Taylor Approximations for Stochastic Partial Differential Equations*, Society for Industrial and Applied Mathematics, 2011.

[103] D. S. Jin, M. R. Matthews, J. R. Ensher, C. E. Wieman, and E. A. Cornell, *Temperature-dependent damping and frequency shifts in collective excitations of a dilute Bose–Einstein condensate*, pp. 493–496, WORLD SCIENTIFIC, 2012.

[104] B. L. G. Jonsson, J. Fröhlich, S. Gustafson, and I. M. Sigal, *Long time motion of NLS solitary waves in a confining potential*, Ann. Henri Poincaré **7** (2006), no. 4, 621–660.

[105] A. D. Kennedy and B. Pendleton, *Cost of the generalised hybrid Monte Carlo algorithm for free field theory*, Nuclear Physics B **607** (2001), no. 3, 456 – 510.

[106] W. Ketterle, D. S. Durfee, and D. M. Stamper-kurn, *Making, probing and understanding Bose–Einstein condensates*, Proceedings of the international school of physics

"Enrico Fermi", course CXL, edited by M. Inguscio, S. Stringari and C.E. Wieman (IOS, Press, 1999, p. 67.

[107] T. W. B. Kibble, *Topology of cosmic domains and strings*, Journal of Physics A: Mathematical and General **9** (1976), no. 8, 1387.

[108] R. Kollar, *Existence and stability of vortex solutions of certain nonlinear Schrödinger equations*, Ph.D. thesis, 2004, Thesis (Ph.D.)–University of Maryland, College Park, p. 147.

[109] R. Kollár and R. L. Pego, *Spectral stability of vortices in two–dimensional Bose–Einstein condensates via the Evans function and Krein signature*, Appl. Math. Res. Express. AMRX (2012), no. 1, 1–46.

[110] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1997.

[111] G. Lamporesi, S. Donadello, S. Serafini, F. Dalfovo, and G. Ferrari, *Spontaneous creation of Kibble–Zurek solitons in a Bose–Einstein condensate*, Nat Phys **9** (2013), no. 10, 656–660.

[112] G. Leibon, D. N. Rockmore, W. Park, R. T., and G. S. Chirikjian, *A Fast Hermite Transform*, Theor. Comput. Sci. **409** (2008), no. 2, 211–228.

[113] T. Lelièvre, F. Nier, and G. A. Pavliotis, *Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion*, Journal of Statistical Physics **152** (2013), 237–274.

[114] T. Lelièvre, M. Rousset, and G. Stoltz, *Free Energy Computations: a Mathematical Perspective*, Imperial College Press, London, Hackensack (N.J.), Singapore, 2010.

[115] T. Lelièvre and G. Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*, Acta Numerica **25** (2016), 681–880.

[116] M. Lewin, P. T. Nam, and N. Rougerie, *Derivation of Hartree's theory for generic mean-field Bose systems*, Advances in Mathematics **254** (2014), no. Supplement C, 570 – 621.

[117] E.H. Lieb, *The Mathematics of the Bose Gas and Its Condensation*, Oberwolfach Seminars, Springer Basel AG, 2005.

[118] J. Liu, *Order of convergence of splitting schemes for both deterministic and stochastic nonlinear Schrödinger equations*, SIAM J. Numer. Anal. **51** (2013), no. 4, 1911–1932.

[119] K. W. Madison, F. Chevy, W. Wohlleben, and J. Dalibard, *Vortex formation in a stirred Bose–Einstein condensate*, Phys. Rev. Lett. **84** (2000), 806–809.

[120] M. Maeda, *Symmetry breaking and stability of standing waves of nonlinear Schrödinger equations*, Master Thesis, Kyoto University, 2008.

[121] G. Mastroianni and G. Monegato, *Error estimates for Gauss-Laguerre and Gauss-Hermite quadrature formulas*, pp. 421–434, Birkhäuser Boston, Boston, M. A., 1994.

[122] M. R. Matthews, B. P. Anderson, P. C. Haljan, D. S. Hall, C. E. Wieman, and E. A. Cornell, *Vortices in a Bose-Einstein condensate*, Phys. Rev. Lett. **83** (1999), 2498–2501.

[123] J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic Processes and their Applications **101** (2002), no. 2, 185 – 232.

[124] H. J. Metcalf and P. van der Straten, *Laser Cooling and Trapping*, Graduate Texts in Contemporary Physics, Springer New York, 2001.

[125] S. P. Meyn and R. L. Tweedie, *Stability of Markovian processes II: continuous-time processes and sampled chains*, Advances in Applied Probability **25** (1993), no. 3, 487–517.

[126] T. Mizumachi, *Vortex solitons for 2D focusing nonlinear Schrödinger equation*, Differential Integral Equations **18** (2005), no. 4, 431–450.

[127] ———, *Instability of vortex solitons for 2D focusing NLS*, Adv. Differential Equations **12** (2007), no. 3, 241–264.

[128] P. Monmarché, *Hypocoercive relaxation to equilibrium for some kinetic models via a third order differential inequality*, arXiv preprint arXiv:1306.4548 (2013).

[129] C. A. Newell, *Inelastic collisions in cold dipolar gases*, 2010.

[130] Y.-G. Oh, *Cauchy problem and Ehrenfest's law of nonlinear Schrödinger equations with potentials*, J. Differential Equations **81** (1989), no. 2, 255–274.

[131] M. Ottobre, N. S. Pillai, and K. Spiliopoulos, *Optimal scaling of the MALA algorithm with irreversible proposals for Gaussian targets*, arXiv preprint arXiv:1702.01777 (2017).

[132] Cockburn S. P., *Bose gases in and out of equilibrium within the stochastic Gross-Pitaevskii Equation*, Ph.D. thesis, Newcastle University, 2010.

[133] G. A. Pavliotis, *Stochastic Processes and Applications*, Texts in Applied Mathematics, vol. 60, Springer, New York, 2014, Diffusion processes, the Fokker-Planck and Langevin equations.

[134] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.

[135] R. L. Pego and H. A. Warchall, *Spectrally stable encapsulated vortices for nonlinear Schrödinger equations*, J. Nonlinear Sci. **12** (2002), no. 4, 347–394.

[136] J. Pérez-Ríos and A. S. Sanz, *How does a magnetic trap work?*, American Journal of Physics **81** (2013), no. 11, 836–843.

[137] E. A. J. F. Peters and G. de With, *Rejection–free Monte Carlo sampling for general potentials*, Phys. Rev. E **85** (2012), 026703.

[138] C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases*, 2 ed., Cambridge University Press, 2008.

[139] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation*, International Series of Monographs on Physics, vol. 116, The Clarendon Press, Oxford University Press, Oxford, 2003.

[140] L. P. Pitaevskii, *Vortex lines in an imperfect Bose gas*, Sov. Phys.—JETP **13** (1961), 451.

[141] R. Poncet, *Generalized and hybrid Metropolis-Hastings overdamped Langevin algorithms*, working paper or preprint, January 2017.

[142] ———, *Numerical analysis of the Gross-Pitaevskii Equation with a randomly varying potential in time*, working paper or preprint, January 2017.

[143] D. Potts, G. Steidl, and M. Tasche, *Fast algorithms for Discrete Polynomial Transforms*, Math. Comput. **67** (1998), no. 224, 1577–1590.

[144] M.-Z. Qin and W.-J. Zhu, *Volume-preserving schemes and numerical experiments*, Computers & Mathematics with Applications **26** (1993), no. 4, 33–42.

[145] M. Quiroga-Teixeiro and H. Michinel, *Stable azimuthal stationary state in quintic nonlinear optical media*, J. Opt. Soc. Am. B **14** (1997), no. 8, 2004–2009.

[146] Duboscq R., *Analyse et simulation d'equations de Schrödinger déterministes et stochastiques. Applications aux condensats de Bose-Einstein en rotation*, Ph.D. thesis, Université de Lorraine, 2013.

[147] M. Reed and B. Simon, *Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975.

[148] ———, *Methods of Modern Mathematical Physics. IV. Analysis of Operators*, Academic Press, New York-London, 1978.

[149] ———, *Methods of Modern Mathematical Physics. III*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1979, Scattering theory.

[150] L. Rey-Bellet and K. Spiliopoulos, *Irreversible Langevin samplers and variance reduction: a large deviations approach*, Nonlinearity **28** (2015), no. 7, 2081.

[151] ———, *Variance reduction for irreversible Langevin samplers and diffusion on graphs*, Electron. Commun. Probab. **20** (2015), 16 pp.

[152] G. O. Roberts and R. L. Tweedie, *Exponential convergence of langevin distributions and their discrete approximations*, Bernoulli **2** (1996), no. 4, 341–363.

[153] S. J. Rooney, P. B. Blakie, and A. S. Bradley, *Numerical method for the stochastic projected Gross-Pitaevskii equation*, Phys. Rev. E **89** (2014), 013302.

[154] P. J. Rossky, J. D. Doll, and H. L. Friedman, *Brownian dynamics as smart Monte Carlo simulation*, The Journal of Chemical Physics **69** (1978), no. 10, 4628–4633.

[155] T. A. Savard, K. M. O'hara, and J. E. Thomas, *Laser-noise-induced heating in far-off resonance optical traps*, Physical Review A **56** (1997), no. 2, R1095.

[156] R. Seiringer, *Gross-Pitaevskii theory of the rotating Bose gas*, Comm. Math. Phys. **229** (2002), no. 3, 491–509.

[157] A. E. Siegman, *Lasers*, University Science Books, 1986.

[158] J. Skilling, *Nested sampling for general Bayesian computation*, Bayesian Anal. **1** (2006), no. 4, 833–859.

[159] H. T. C. Stoof and M. J. Bijlsma, *Dynamics of fluctuating Bose–Einstein condensates*, Journal of Low Temperature Physics **124** (2001), no. 3, 431–442.

[160] S. K. Suslov, *Dynamical invariants for variable quadratic hamiltonians*, Physica Scripta **81** (2010), no. 5, 055006.

[161] V Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, vol. 1054, Springer, 1984.

[162] L. Tierney and A. Mira, *Some adaptive Monte Carlo methods for Bayesian inference*, Statistics in Medicine **18** (1999), no. 17-18, 2507–2515.

[163] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second-Order Differential Equations. Part I*, Second Edition, Clarendon Press, Oxford, 1962.

[164] J. Villen-Altamirano and M. Villen-Altamirano, *Restart: a method for accelerating rare event simulations*, Queueing, Performance and Control (1991), 71–76.

[165] M. Vucelja, *Lifting–A nonreversible Markov chain Monte Carlo algorithm*, American Journal of Physics **84** (2016), no. 12, 958–968.

[166] C. N. Weiler, *Spontaneous formation of quantized vortices in Bose-Einstein condensates*, Ph.D. thesis, The University of Arizona, 2008.

[167] C. N. Weiler, T. W. Neely, D. R. Scherer, A. S. Bradley, M. J. Davis, and B. P. Anderson, *Spontaneous vortices in the formation of Bose-Einstein condensates*, Nature **455** (2008), no. 7215, 948–951.

[168] M. I. Weinstein, *Modulational stability of ground states of nonlinear Schrödinger equations*, SIAM J. Math. Anal. **16** (1985), no. 3, 472–491.

[169] G. K. Woodgate, *Elementary Atomic Structure*, Oxford science publications, Clarendon Press, 1980.

[170] S.-J. Wu, C.-R. Hwang, and M. T. Chu, *Attaining the optimal Gaussian diffusion acceleration*, Journal of Statistical Physics **155** (2014), no. 3, 571–590.

[171] X.-C. Yao, H.-Z. Chen, Y.-P. Wu, X.-P. Liu, X.-Q. Wang, X Jiang, Y. Deng, Y.-A. Chen, and J.-W. Pan, *Observation of coupled vortex lattices in a mass-imbalance Bose and Fermi superfluid mixture*, Phys. Rev. Lett. **117** (2016), 145301.

[172] W. H. Zurek, *Cosmological experiments in superfluid helium?*, Nature **317** (1985), no. 6037, 505–508.

[173] M. W. Zwierlein, J. R. Abo-Shaeer, A. Schirotzek, C. H. Schunck, and W. Ketterle, *Vortices and superfluidity in a strongly interacting Fermi gas*, Nature **435** (2005), no. 7045, 1047–1051.

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Méthodes numériques pour la simulation d'équations aux dérivées partielles stochastiques non-linéaires en condensation de Bose-Einstein

Mots clefs : analyse numérique, équations aux dérivées partielles stochastiques, condensation de Bose-Einstein, méthodes de Monte Carlo

Résumé : Cette thèse porte sur l'étude de méthodes numériques pour l'analyse de deux modèles stochastiques apparaissant dans le contexte de la condensation de Bose-Einstein. Ceux-ci constituent deux généralisations de l'équation de Gross-Pitaevskii. Cette équation aux dérivées partielles déterministe modélise la dynamique de la fonction d'onde d'un condensat de Bose-Einstein piégé par un potentiel extérieur confinant.
Le premier modèle étudié permet de modéliser les fluctuations de l'intensité du potentiel confinant et prend la forme d'une équation aux dérivées partielles stochastiques. Nous construisons, dans un premier temps, un schéma numérique de type Crank-Nicolson, basé sur une discrétisation spectrale en espace. Nous démontrons qu'il converge fortement en probabilité à l'ordre au moins 1 en temps. Le chapitre suivant est consacré à l'étude théorique et numérique de la dynamique d'une solution stationnaire (pour l'équation déterministe) de type vortex, soumise à des perturbations aléatoires du potentiel de confinement.
Le deuxième modèle permet de modéliser les effets de la température sur la dynamique d'un condensat. Lorsque celle-ci n'est pas nulle, la condensation n'est pas complète et le condensat interagit avec les particules non condensées. Nous nous sommes intéressés dans les chapitres suivants à des questions relatives à la simulation de la distribution des solutions de cette équation en temps long. D'abord, nous construisons une méthode d'échantillonnage sans biais pour des mesures connues à une constante multiplicative près. C'est une méthode de Monte Carlo par chaînes de Markov non-réversibles basée sur une équation de type Langevin sur-amortie. Ensuite, nous nous sommes consacré à l'étude numérique de dynamiques métastables liées à la nucléation de vortex dans des condensats en rotation, à l'aide de simulations d'événements rares correspondant aux changements de configurations métastables.

Title : Numerical methods for the simulation of nonlinear stochastic partial differential equations in Bose-Einstein condensation

Keywords : numerical analysis, stochastic partial differential equations, Bose-Einstein condensation, Monte Carlo methods

Abstract : This thesis is devoted to the numerical study of two stochastic models arising in Bose-Einstein condensation. They constitute two generalisations of the Gross-Pitaevskii equation. This deterministic partial differential equation models the wave function dynamics of a Bose-Einstein condensate trapped in an external confining potential. The first chapter contains a simple presentation of the Bose-Einstein condensation phenomenon and of the experimental methods used to construct such systems.
The first model considered enables to model the fluctuations of the confining potential intensity, and takes the form of a stochastic partial differential equation. We propose to build a Crank-Nicolson numerical scheme to solve this model, based on a spectral space discretisation. We show that the proposed scheme converges strongly at order at least one in probability. The next chapter is devoted to the numerical and theoretical study of the dynamics of a stationary solution (for the deterministic equation) of vortex type under random disturbances of the confining potential.
The second model can be used to model the effects of the temperature on the dynamics of a Bose-Einstein condensate. In the case of finite temperature, the Bose-Einstein condensation is not complete and the condensate interacts with the non-condensed particles. We have studied some questions linked to the long time simulation of this model. First, we construct an unbiased sampling method for measures known up to a multiplicative constant. It is a non-reversible Markov-Chain Monte Carlo algorithm based on an overdamped Langevin equation. Then, we numerically study metastable dynamics linked to the nucleation of vortices in rotating Bose-Einstein condensates.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France