Communications coopératives pour les réseaux ad hoc sans fil

Nadia Fawaz

TELECOM ParisTech (ENST)
EURECOM

THESIS

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from TELECOM ParisTech

Specialization: Communications and Electronics

Nadia FAWAZ

Cooperative Communications for Wireless Ad Hoc Networks

President	J.C. Belfiore, TELECOM ParisTech (Paris, France)
Reviewers	P. Loubaton, Université de Marne la Vallée (Marne la Vallée, France) R.R. Müller, NTNU (Trondheim, Norway)
Examiners	P. Duhamel, CNRS/LSS-SUPELEC (Gif-sur-Yvette, France) M. Médard, Massachusetts Institute of Technology (Cambridge, USA)
Guest	J. Turbert, Délégation Générale pour l'Armement (Rennes, France)
Thesis supervisors	M. Debbah, SUPELEC (Gif-sur-Yvette, France) D. Gesbert, EURECOM (Sophia-Antipolis, France)

Thesis defended on December, 16$^{\text{th}}$ 2008, in Sophia-Antipolis, France.

TELECOM ParisTech (ENST)
EURECOM

THESE

Présentée pour obtenir le Grade de Docteur
de TELECOM ParisTech

Spécialité: Communications et Electronique

Nadia FAWAZ

Communications Coopératives pour les Réseaux Ad Hoc Sans Fil

Président	J.C. Belfiore, TELECOM ParisTech (Paris, France)
Rapporteurs	P. Loubaton, Université de Marne la Vallée (Marne la Vallée, France) R.R. Müller, NTNU (Trondheim, Norvège)
Examinateurs	P. Duhamel, CNRS/LSS-SUPELEC (Gif-sur-Yvette, France) M. Médard, Massachusetts Institute of Technology (Cambridge, Etats-Unis)
Invité	J. Turbert, Délégation Générale pour l'Armement (Rennes, France)
Directeurs de Thèse	M. Debbah, SUPELEC (Gif-sur-Yvette, France) D. Gesbert, EURECOM (Sophia-Antipolis, France)

Thèse soutenue le 16 Décembre 2008, à Sophia-Antipolis, France.

A ma Famille

English Abstract

Wireless ad hoc networks have recently received significant attention because of their historical defense and security applications, as well as their more recent potential commercial applications. Most of the research on ad hoc networks has focused on higher layer (link, network, or transport) issues, especially on routing protocols designed to cope with frequent and dynamic changes in network topology. The theoretical and practical limits of the performance of dense ad hoc networks, recently unveiled in [1, 2], revealed the need for alternative techniques to improve the interference-limited performance of dense ad hoc networks.

In this thesis, we address the following issue: how can the link capacity performance in wireless ad hoc networks be improved through the use of more advanced physical layer techniques?

We first introduce the physical layer factors that impact the performance of dense ad hoc networks. In particular, we show how the throughput can be improved by using directive antennas, or by managing the number and position of relaying nodes. Then, we turn our attention to ad hoc networks in which nodes are empowered with cooperative capabilities. We first consider ad hoc networks with a small number of nodes, and we propose techniques to improve the spectral efficiency of cooperative strategies. The proposed cooperative strategies make more efficient use of the wireless resource by combining orthogonality-relaxation and dirty paper precoding. Finally, ad hoc networks with a high density of nodes are examined. We introduce cooperation in a large ad hoc network through a cooperative-clustering approach. Using tools from random matrix theory and free probability theory, we analyze the asymptotic capacity of the system when the node density increases.

This thesis shows that link capacity performance in dense wireless ad hoc networks can be improved, as long as nodes are empowered with cooperative capabilities at the physical layer, and cooperative strategies are properly

designed.

French Abstract

Les réseaux ad hoc sans fil ont reçu dernièrement une attention considérable en raison de leurs applications historiques dans le domaine de la défense et de la sécurité, et de leurs récentes applications commerciales potentielles. Jusqu'à récemment, la plupart des travaux de recherche sur les réseaux ad hoc mettaient l'accent sur des problèmes aux couches supérieures (liaison, réseau, transport), en particulier sur la conception de protocoles de routage robustes aux changements dynamiques et fréquents de topologie. Les limites théoriques et pratiques des performances des réseaux ad hoc dévoilées récemment [1, 2] ont révélé le besoin de techniques alternatives pour améliorer les performances des réseaux ad hoc denses, dont la principale limitation est due aux interférences.

Cette thèse s'attaque à la problématique suivante : comment améliorer les performances en termes de capacité de lien dans les réseaux ad hoc sans fil, en utilisant des techniques plus avancées à la couche physique ?

Dans un premier temps, les facteurs de la couche physique impactant les performances des réseaux ad hoc denses sont présentés. En particulier, on montre comment l'utilisation d'antennes directionnelles ou la gestion du nombre et de la position de nœuds-relais permettent d'améliorer le débit. Puis l'attention est portée vers les réseaux ad hoc dont les nœuds sont dotés de capacités de coopération. Les réseaux ad hoc avec un petit nombre de nœuds coopératifs sont tout d'abord considérés, et des techniques améliorant l'efficacité spectrale des stratégies de coopération sont proposées. Ces stratégies coopératives permettent une meilleure utilisation de la ressource sans fil grâce à la combinaison de la relaxation d'orthogonalité et du précodage "Dirty Paper". Puis, les réseaux ad hoc à haute densité de nœuds sont examinés. Dans ces réseaux de grande dimension, la coopération est introduite à travers une approche de clusters coopératifs. En utilisant des outils de la théorie des matrices aléatoires et des probabilités libres, la capacité asymp-

totique du réseau est analysée quand la densité de nœuds augmente.

Cette thèse montre que la capacité de lien dans les réseaux ad hoc sans fil denses peut être améliorée, lorsque les nœuds sont dotés de capacités de coopération à la couche physique, et les stratégies de coopération sont efficacement conçues.

Contents

List of Figures

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is usually indicated once, when it first occurs in the text.

AF	Amplify and Forward
AODV	Ad-hoc On-Demand Distance Vector Routing
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
BER	Bit Error Rate
BTS	Base Transceiver Station
CC	Coded Cooperation
CDMA	Code Division Multiple Access
cf.	confer
CF	Compress and Forward
CIR	Channel Impulse Response
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
DF	Decode and Forward
DFT	Discrete Fourier Transform
DMT	Diversity-Multiplexing Tradeoff
DPC	Dirty Paper Coding
DSR	Dynamic Source Protocol
EVD	EigenValue Decomposition
e.g.	exempli gratia
FIR	Finite Impulse Response
FPT	Free Probability Theory
i.e.	id est
i.i.d.	independent and identically distributed

MAC	Multiple Access Channel
MANET	Mobile Ad hoc NETworks
MCN	Multi-hop Cellular Network
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
NAF	Non-orthogonal Amplify and Forward
NC	Network Coding
OLSR	Optimized Link State Routing Protocol
PDF	Parallel Decode and Forward
RDF	Repetition Decode and Forward
RMT	Random Matrix Theory
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SOPRANO	Self-Organizing Packet Radio Networks with Overlay
STC	Space-Time Code
SVD	singular value decomposition
TR	Time Reversal
TDMA	Time Division Multiple Access
ULA	Uniform Linear Array
UMTS	Universal Mobile Telecommunications System
UWB	Ultra WideBand
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Network
ZRP	Zone Routing Protocol

Notations

Here is a list of the main notations and symbols used in this document. We have tried to keep consistent notations throughout the document, but some symbols have different definitions depending on when they occur in the text.

General Notations

$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
x	Scalar variable
$\mathbf{x}$	Vector variable
$\mathbf{X}$	Matrix variable
$\mathbf{I}_N$	identity matrix of size N
$(\cdot)^*$	Complex conjugate operator
$(\cdot)^T$	Transpose operator
$(\cdot)^H$	Hermitian transpose operator
$\mathrm{tr}(\cdot)$	Trace of a matrix
$\det(\cdot)$	Determinant of a matrix
$a_{ij}^{(k)}$	(i,j)-th entry of matrix $\mathbf{A}_k$
$\lambda_{\mathbf{A}}(1), \ldots, \lambda_{\mathbf{A}}(n)$	Eigenvalues of $n \times n$ matrix $\mathbf{A}$
$\|\mathbf{A}\|$	Operator norm of matrix $\mathbf{A}$: $\|\mathbf{A}\| \triangleq \sqrt{\max_i \lambda_{\mathbf{A}^H\mathbf{A}}(i)}$
$\|\mathbf{A}\|_F$	Fröbenius norm of matrix $\mathbf{A}$: $\|\mathbf{A}\|_F \triangleq \sqrt{\mathrm{tr}(\mathbf{A}^H\mathbf{A})} = \sum_{i,j} \|a_{ij}\|^2$
$\bigotimes_{i=1}^N \mathbf{A}_i$	Non-commutative matrix product: $\bigotimes_{i=1}^N \mathbf{A}_i \triangleq \mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_N$
$*$	Convolution operator
log	Logarithm in base 2

ln	Logarithm in base e
$u(\cdot)$	Unit-step function: $u(x) = 0$ if $x < 0$; $u(x) = 1$ if $x \geq 0$
$\mathbb{1}\{\cdot\}$	Indicator function: $\mathbb{1}\{X\}=1$ if X true; $\mathbb{1}\{X\}=0$ if X false
$K(m)$	Complete elliptic integral of the first kind: $K(m) \triangleq \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1-m\sin^2\theta}}$
o	$f(n) = o(g(n))$ means that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$
O	$f(n) = O(g(n))$ means that there exists a constant c and an integer N such that for $n > N$, $\lvert f(n)\rvert \leq c\lvert g(n)\rvert$
ω	$f(n) = \omega(g(n))$ means that $g(n) = o(f(n))$
Ω	$f(n) = \Omega(g(n))$ means that $g(n) = O(f(n))$
Θ	$f(n) = \Theta(g(n))$ means that $f(n) = O(g(n))$ and $g(n) = O(f(n))$
$\Pr\{\cdot\}$	Probability of an event
$\mathrm{E}[\cdot]$	Statistical expectation
$\mathcal{H}(\cdot)$	Entropy
$\mathcal{I}(\cdot;\cdot)$	Mutual information

Chapter 3: Non-cooperative Networks

$\bar{i}$	Complementary integer of i in set $\{1, 2\}$, e.g. if $i = 1$, $\bar{i} = 2$

Chapter 4: Small Cooperative Networks

f_i	precoding function at source S_i

Chapter 5: Large Cooperative Networks

$F_{\mathbf{A}}(\cdot)$	Empirical eigenvalue distribution of square matrix $\mathbf{A}$ with real eigenvalues
$f_{\mathbf{A}}(\cdot)$	Probability density function of the eigenvalues of square matrix $\mathbf{A}$ with real eigenvalues

Chapter 1

French Summary

Introduction et Motivations

Les réseaux ad hoc sont des réseaux sans fil fixes ou mobiles sans infrastructure, dans lesquels le transfert d'information repose sur la capacité des nœuds sans fil à relayer les données les uns pour les autres. Ces réseaux se distinguent ainsi des réseaux à infrastructure, tels que les réseaux cellulaires où une station de base centralise les communications des nœuds sans fil situés dans une cellule, ou les réseaux sans fil locaux (WLAN) où un point d'accès gère les connections entre les nœuds sans fil. Ces réseaux dynamiques ont la particularité de s'auto-organiser et de pouvoir se déployer avec une grande flexibilité et autonomie sur des terrains improvisés, les rendant appropriés pour des systèmes de communication de défense et d'urgence. Historiquement, les réseaux ad hoc comptent parmi leurs applications le déploiement sur champs de bataille, les interventions de secours, le déploiement d'urgence en cas de catastrophe naturelle ayant causé la destruction de l'infrastructure de communication préexistante. Récemment, le développement des réseaux de capteurs sans fil (WSN) et des WLAN IEEE 802.11 (WiFi) — norme possédant un mode peer-to-peer permettant à des appareils sans fil de se connecter les uns aux autres — ont encouragé l'émergence d'idées d'applications commerciales des réseaux ad hoc, tels que les réseaux locaux de jeux, les réseaux mesh communautaires, les réseaux véhiculaires sur voie, l'extension de couverture des réseaux cellulaires à travers des réseaux hybrides (cellulaire-ad hoc). De nombreuses idées sur les réseaux hybrides, mixant réseaux cellulaires et schémas multi-saut, ont d'ailleurs commence à apparaître [3, Multihop Cellular Network (MCN)], [4, iCAR], [5, Self-Organizing Packet Radio Networks with Overlay(SOPRANO)]. Réseaux ad hoc et cellulaires furent combinés dans l'espoir que la mise en commun des avantages des deux schémas permettrait d'étendre la couverture des cellules tout en supportant des topologies dynamiques [3], d'augmenter la scalabilité et la robustesse des systèmes ad hoc purs, et d'équilibrer la charge dans les réseaux cellulaires [4]. Toutefois, ces travaux traitent essentiellement de problématiques de routage.

Les changements de topologies dynamiques et imprédictibles dans les réseaux ad hoc ont déclenché des travaux de recherche extensifs sur les couches hautes, principalement sur les protocoles de routage permettant une adaptation rapide dans les réseaux ad hoc mobiles (MANETs) extrêmement dynamiques. Traditionnellement, les protocoles de routage dans les MANETs étaient focalisés sur la minimisation du nombre de sauts, approche qui ne prend pas en compte la qualité du lien et conduit à des routes de capac-

ité bien moindre que celles des chemins de haute qualité disponibles dans le réseau [6]. Dans [1], les limites théoriques sur le débit des réseaux ad hoc furent dévoilées, puis confirmées de manière pratique par des simulations et expériences dans [2]. Dans [1], considérant un réseau ad hoc de n nœuds capable de transmettre à W bits/seconde, et distribués aléatoirement sur un disque d'aire unitaire, il fut montré que le débit par nœud décroît comme $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ bits/seconde quand le nombre de nœuds n augmente. En effet, dans les réseaux ad hoc denses, la ressource sans fil doit être partagée entre les transmissions concurrentes d'un grand nombre de nœuds sans fil, et par conséquent, les performances sont limitées par l'interférence. Les transmissions sans fil de chaque nœud doivent donc être confinées au voisinage du nœud, requérant donc le multi-saut pour que l'information circule de la source à la destination. Il en résulte que la plupart des transmissions dans le réseau transportent des données relayées, ce qui conduit à une diminution importante du débit total. Cependant, cette décroissance du débit est obtenue sous des hypothèses spécifiques sur le mode de transmission à la couche physique: transmissions point à point multi-saut entre des nœuds sans fil fixes équipés d'antennes omnidirectionnelles et transmettant leurs signaux indépendamment les uns des autres sans aucune interaction coopérative entre les nœuds. Sachant que la performance à la couche physique est une borne supérieure pour les performances aux couches plus hautes, la question suivante se pose naturellement:

- Est-il possible d'améliorer les performances en termes de capacité de lien dans les réseaux ad hoc sans fil, en utilisant des techniques plus avancées à la couche physique ?

- Quels sont les facteurs à la couche physique permettant d'améliorer la capacité de lien, et quelles sont leurs limites?

Facteurs Améliorant les Performances des Reseaux Ad Hoc Networks

Il a été montré que plusieurs facteurs à la couche physique permettent d'améliorer les performances dans les réseaux ad hoc, dont la directivité des antennes, la mobilité des nœuds , le positionnement des nœuds , la coopération et les systèmes d'antennes multiples (MIMO) virtuels.

Les Antennes Directionnelles peuvent être utilisées de manière adaptative pour améliorer la fiabilité des transmissions et diminuer l'interférence [7–11]. En effet, si les nœuds ont de l'information sur leur topologie locale, la formation de faisceaux ou la sectorisation de la puissance de transmission peuvent être utilisées pour focaliser la puissance transmise dans la direction de leur destination. Ainsi, la probabilité de causer de l'interférence à un récepteur autre que la destinataire souhaité diminue. Néanmoins, la connaissance de la position du récepteur est nécessaire au transmetteur pour pouvoir focaliser le faisceau de transmission dans la bonne direction. Dans des réseaux à haute mobilité, traquer la position d'un grand nombre de nœuds requière un *feedback* non-négligeable qui augmente l'*overhead* du protocole de transmission [12–14]. Par conséquent, dans les réseaux ad hoc, on considère en général des antennes omnidirectionnelles, malgré leur impact négatif sur l'interférence. Le Chapitre 3 traite, entre autres, de la problématique suivante :

- Est-il possible de tirer bénéfice de la directivité des antennes ou de la formation de faisceau dans les réseaux ad hoc denses, tout en évitant la surcharge due au feedback?

La Mobilité des nœuds permet également d'améliorer la capacité des réseaux ad hoc sans fil [15, 16]. En permettant aux nœuds sans fil de se déplacer, un *scaling* constant du débit peut être obtenu quand le nombre de nœuds augmente. L'idée consiste à exploiter la diversité multi-utilisateurs par le relayage de paquets dans un réseau dense : une source divise son flux de paquets entre plusieurs nœuds sans fil mobiles qui vont jouer le rôle de relais et livrer le paquet qu'ils transportent dès qu'ils passent à proximité de la destination. Leur mouvement rend les nœuds relais équivalents à des nœuds dont les canaux vers la source et vers la destination varient dans le temps. Le nombre de nœuds dans le réseau étant grand, à chaque instant la probabilité qu'un nœud relai soit proche de la source et donc reçoive un nouveau paquet à transporter est grande, ainsi que la probabilité qu'un nœud soit proche de la destination et livre un paquet. Il résulte de la mobilité que l'interférence dans le réseau est diminuée, que le nombre de sauts qu'un paquet doit traverser est réduit à deux, et que le débit est considérablement augmenté. Néanmoins, ce *scaling* constant du débit dans les réseaux ad hoc mobiles fut obtenu sous l'hypothèse d'applications tolérant de larges délais. Dans [17], le compromis débit-délai est analysé dans les réseaux mobiles, et

il est montré que le délai requis pour supporter un *scaling* constant du débit dépend de la vitesse des nœuds sans fil: le délai augmente quand la vitesse diminue, ralentissement qui se produit quand le réseau devient de plus en plus bondé en raison d'une densité croissante. Cette observation soulève la question de savoir s'il est possible, par d'autres facteurs, d'améliorer le débit dans les réseaux ad hoc sans un délai de plus en plus élevé.

Le positionnement et le nombre de nœuds relais supportant la communication d'une paire source-destination ont un impact sur la capacité d'un réseau de relais [18–22]. Lorsqu'une seule paire source-destination dans le réseau est active, et que tous les autres nœuds relaient les données de la source, la contribution des relais à la capacité et le scaling de la capacité quand le nombre de nœuds augmente peuvent être analysés. Dans le Chapitre 3, les questions suivantes sont examinées :

- Comment la capacité de lien croit-elle avec le nombre de relais aidant une paire source-destination?
- Quel est l'impact de la topologie du réseau, en particulier de la position des nœuds relais, sur la capacité de lien?
- Quel est l'impact de l'environnement physique, à travers les réflexions, les atténuations et délais de propagation... sur la capacité de lien?

La coopération et le MIMO virtuel peuvent également contribuer à l'amélioration des performances dans les réseaux ad hoc sans fil. Dans les modèles mentionnés précédemment, on considérait que les nœuds agissaient indépendamment et la coopération était limitée à une réexpédition passive : les nœuds n'interagissaient pas, mais pouvaient simplement transférer leur signal reçu. Si des techniques de traitement plus avancées sont envisagées à la couche physique et les nœuds peuvent coopérer, on peut prévoir que les performances des réseaux ad hoc seront améliorées. En effet, dans les systèmes de communication point à point, les techniques MIMO permettent d'améliorer la fiabilité des communications grâce aux gains de diversité spatiale, et d'augmenter l'efficacité spectrale grâce aux gains de multiplexage. Ainsi, lorsqu'on considère un réseau de nœuds sans fil où les nœuds peuvent coopérer pour transmettre conjointement de l'information, un système MIMO virtuel peut être construit de manière distribuée, et la coopération

peut permettre l'exploitation des gains MIMO. Dans les Chapitres 4 et 5, les problématiques suivantes sont traitées :

- Comment faut-il concevoir les stratégies de coopération pour améliorer le débit des réseaux ad hoc, tout en utilisant la ressource sans fil de manière efficace?
- Que peut-on dire de la capacité d'un grand réseau coopératif multi-sauts, où les nœuds sans fil ont des aptitudes au MIMO virtuel?

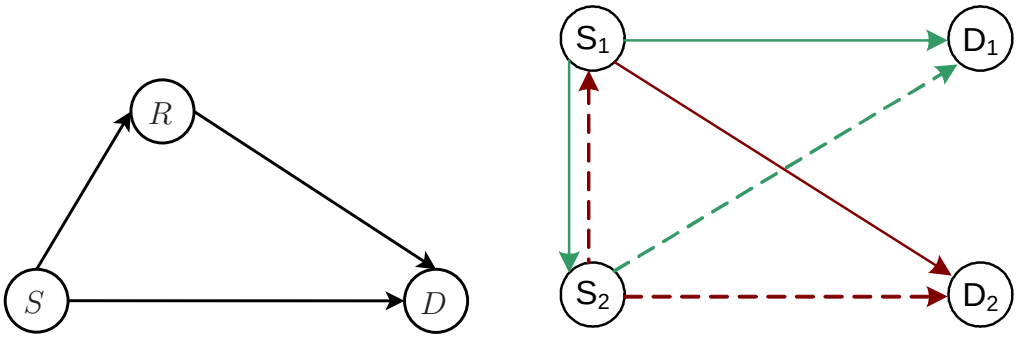
Réseaux Coopératifs de Petites Dimensions

Un système de communications coopératives est constitué lorsque des nœuds sans fil distribués interagissent pour transmettre conjointement de l'information. Plusieurs terminaux radio relayant les signaux les uns des autres forment un réseau d'antennes virtuel, et leur coopération permet d'exploiter la diversité spatiale des canaux à évanouissement, qui prend alors le nom de diversité coopérative. Les stratégies de coopération furent dans un premier temps conçues pour des réseaux de petites dimensions, représentant les briques élémentaires pour la construction de réseaux ad hoc de plus grandes dimensions. Les réseaux coopératifs de petites dimensions les plus basiques sont :

- le canal à relai (*relay channel*), cf. Fig. 1.1(a) : une paire source-destination aidée par un relai ;
- le canal à interférence coopératif (*cooperative interference channel*), cf. Fig. 1.1(b) : deux paires source-destination coopérant à la transmission et/ou à la réception.

Une pléthore de stratégies de coopération ont été proposées pour le canal à relai ou le canal à interférence coopératif [23, 24], les plus célèbres étant *Amplify and Forward* (AF), *Decode and Forward* (DF) et *Compress and Forward* (CF) [25]. La différence entre ces stratégies réside dans le traitement réalisé au nœud relai avant qu'il ne retransmette (cf. Chapitre 4).

La plupart des stratégies de coopération ont été conçues de telle sorte que la contrainte pratique de semi-duplex soit respectée: un terminal radio ne peut pas transmettre et recevoir simultanément dans la même bande de fréquence, car la puissance du signal reçu est très basse par rapport à la puissance du signal transmis. Un nœud relai utilise donc des canaux orthogonaux pour recevoir un signal de la source, et pour transmettre son signal



(a) Canal à relai: une source, un relai, une destination

(b) canal à interférence coopératif: deux sources coopérant et deux destinations

Figure 1.1: Réseaux coopératifs de petites dimensions

relayé. Par exemple, dans le canal à relai en Fig. 4.1(a), les transmissions sont réparties en deux blocs : dans le premier bloc, la source transmet et le relai et la destination reçoivent ; dans le second block, le relai transmet son signal relayé et la destination reçoit. Dans le cas du canal à interférence coopératif en Fig. 4.1(b), en général le schéma de transmission à deux blocs du canal à relai est simplement étendu à un schéma de transmission à quatre blocs [25] en répétant le schéma à deux blocs deux fois : un premier schéma à deux blocs pour la transmission de la source S_1 relayée par S_2, suivi d'un second schéma à deux blocs pour la transmission de la source S_2 relayée par S_1. Non seulement le schéma de transmission coopératif résultant respecte la contrainte de semi-duplex, mais il permet également d'éviter toute interférence. Cependant, bien que l'utilisation de canaux orthogonaux sans interférence pour les transmissions respectives de la source et du relai simplifie les algorithmes de réception, elle résulte également en une utilisation inefficace de la bande spectrale. En effet, seule la moitié des degrés de liberté est utilisée pour la transmission vers chaque destination. Une idée naturelle pour utiliser les degrés de liberté plus efficacement consisterait donc à relaxer la contrainte d'orthogonalité, mais un obstacle apparaît alors : la relaxation de la contrainte d'orthogonalité conduirait à l'introduction d'interférences dans le système.

Compte tenu de ces observations, nous examinons dans le Chapitre 4 les questions suivantes :

- Peut-on améliorer l'efficacité spectrale des stratégies de coopération en relaxant la contrainte d'orthogonalité, tout en continuant à respecter la contrainte de semi-duplex?

- Comment peut-on mitiger l'interférence due à la relaxation de la contrainte d'orthogonalité?

Réseaux Coopératifs de Grandes Dimensions

Les lois d'échelle — peu encourageantes — du débit dans les réseaux ad hoc denses dérivées dans [1] soulèvent la question de savoir si les réseaux ad hoc sont appropriés uniquement pour un petit nombre de nœuds ou pour un déploiement dans une zone limitée, ou si les communications coopératives à la couche physique peuvent permettre de déployer les réseaux ad hoc de manière viable pour un grand nombre de nœuds . [1] a ainsi ouvert la voie à des travaux de recherche récents sur les lois d'échelle dans les réseaux ad hoc où les nœuds ont des aptitudes MIMO [26–37].

Dans un réseau ad hoc idéal — mais irréaliste, tous les nœuds pourraient parfaitement coopérer à grande échelle, et la capacité du réseau croîtrait comme dans un grand système MIMO parfait : linéairement avec le nombre d'antennes [38]. Une approche plus réaliste, et toutefois pleine de sens, consiste à grouper les nœuds du réseau en clusters coopératifs : les nœuds appartenant à un cluster coopératif coopèrent pour former un réseau d'antennes virtuel et transmettre ou recevoir conjointement l'information destinée à ou provenant d'autres clusters. Des communications multi-sauts ont alors lieu entre les clusters coopératifs, au lieu de simples nœuds comme c'était le cas dans [1]. Grâce à l'approche coopérative, certaines transmissions qui apparaissaient comme de l'interférence dans [1] sont maintenant vues comme des signaux utiles pouvant être traités conjointement par les nœuds d'un cluster.

Un réseau ad hoc dense est un système très complexe, dont les métriques de performance impliquent un grand nombre de variables et de paramètres. Cependant, les simulations extensives ne sont heureusement pas les seuls moyens de gagner une certaine intuition sur la façon dont les performances du systèmes évoluent lorsque ses dimensions augmentent. Récemment, les théories des matrices aléatoires (RMT) et des probabilités libres (FPT) sont apparues comme des théories appropriées pour l'analyse et la conception de systèmes de communication complexes, et pour révéler les paramètres pertinents impactant leurs performances [39]. En effet, le transfert d'information

dans un système de communication peut souvent être modélisé par une équation matricielle aléatoire de la forme $\mathbf{y} = \mathbf{H}\mathbf{x}+\mathbf{z}$, où $\mathbf{x}$ est le vecteur d'entrée, $\mathbf{y}$ le vecteur de sortie, $\mathbf{z}$ le vecteur de bruit, et $\mathbf{H}$ la matrice de transfert du système. Pour un tel système, il s'avère que la plupart des métriques de performance issues de la théorie de l'information dépendent uniquement des valeurs et vecteurs propres de la matrice de transfert $\mathbf{H}$. RMT et FPT fournissent des résultats utiles sur les valeurs et vecteurs propres de matrices aléatoires de grandes dimensions, qui peuvent être appliqués à l'analyse de systèmes de communication de grandes dimensions. La théorie des matrices aléatoires émergea avec les travaux de Wishart [40], Wigner [41], et Marčenko et Pastur [42], et était historiquement utilisée en physique, avant de l'être dans de nombreux autres domaines. Dans les communications sans fil, RMT fut d'abord utilisée pour analyser les performances des systèmes de communications à antennes multiples, par exemple dans [38,43–45], et des systèmes CDMA, par exemple dans [46–48].

Dans le Chapitre 5, un réseau ad hoc dense où les nœuds sont groupés en clusters coopératifs est considéré, et les questions suivantes sont traitées en utilisant des outils de RMT, FPT et de l'algèbre linéaire :

- Quelle est la capacité asymptotique du système de clusters coopératifs quand le nombre de nœuds dans tous les clusters croît?
- Quels sont les paramètres pertinents impactant la capacité du système?
- Comment les nœuds dans un cluster coopératif doivent-il traiter et transmettre coopérativement leurs signaux afin de maximiser la capacité du système?

Contributions

Chapitre 3

La contribution du Chapitre 3 est double. La première partie traite de l'impact de la directivité des antennes sur le débit et la connectivité d'un réseau dense avec un grand nombre de paires source-destination. La seconde partie analyse l'impact du positionnement et du nombre de nœuds relais passifs sur la capacité d'un système où un grand nombre de nœuds relais soutiennent la communication d'une paire source-destination.

Dans la Section 3.2, un réseau dense de paires source-destination, où les sources sont équipées d'antennes directives, est considéré. La première contribution de la Section 3.2 est la proposition d'un schéma de formation de faisceaux dynamique et aveugle, permettant de tirer profit de la directivité des antennes dans un réseaux décentralisé, tout en évitant le lourd *feedback* pour traquer la position des nœuds . Le système est dynamique et aveugle car une source pointe son antenne directive successivement dans toutes les directions pour viser sûrement mais aveuglement sa destination sans connaître sa position exacte. On montre que la directivité rotationnelle a un impact positif sur la réduction d'interférence, et donc sur la capacité : en focalisant sa puissance de transmission successivement dans toutes les directions, la probabilité qu'une source interfère avec les autres destinations, i.e. que son signal atteigne une destination non-désirée en même temps que cette dernière reçoit un signal de sa propre source, est faible en raison de la focalisation spatiale et de l'asynchronisme de toutes les transmissions. Néanmoins, la directivité rotationnelle introduit un délai : lorsqu'une source ne transmet pas dans la direction de sa destination, du temps et de la puissance sont gâchés. Ces deux effets opposés conduisent à un compromis capacité-délai lorsqu'on ajuste le nombre de rotations.

La deuxième contribution de la Section 3.2 est l'analyse du débit du réseaux avec le schéma de formation de faisceaux dynamique et aveugle, et la comparaison au cas classique où les sources sont équipées d'antennes omnidirectionnelles. On montre que lorsque la densité du réseau augmente, le schéma de formation de faisceaux dynamique et aveugle surpasse les transmissions omnidirectionnelles. En effet, lorsque la densité du réseau augmente, la principale limitation au débit est due l'interférence et des faisceaux de transmission plus étroits sont nécessaires pour réduire l'interférence. Cependant, quand les faisceaux deviennent plus étroits, un nombre plus grand de rotations de l'antenne source est nécessaire pour couvrir tout l'espace, ce qui conduit à un délai accru. Pour une densité de réseau donnée, le compromis entre réduction d'interférence et augmentation du délai résulte en une largeur de faisceau optimale et un nombre de rotation optimal maximisant le débit du réseau.

La troisième contribution de la Section 3.2 est la définition d'un nouveau critère de connectivité lié au débit (*throughput-connectivity*), critère qui permet de prendre en compte l'interférence et l'accès partagé à la ressource sans fil par les différents nœuds du réseau. Ce critère de connectivité est défini pour un taux de transmission cible R comme étant la proportion de paires

dans le réseau auxquelles un débit R peut être garanti. Pour un débit cible R, le nombre de rotations et la largeur de faisceau maximisant la connectivité dépendent de la densité du réseau.

Dans la Section 3.3, un système où une source communique avec une destination avec l'aide d'un réseau dense de relais passifs est examiné. Les relais sont modélisés comme de simples diffuseurs omnidirectionnels, i.e. des nœuds passifs sans capacités d'ingénierie qui reflètent simplement de manière diffuse l'onde électromagnétique incidente provenant de l'antenne source. La première contribution de la Section 3.3 est la formulation d'un modèle prenant en compte l'environnement physique à travers les réflexions, les atténuations, les délais et la multiplicité des voies. Il est montré que ce système de relayage asynchrone peut être modélisé comme un canal virtuel à voies multiples, où le gain de chaque voie résulte de la combinaison des contributions au signal des relais appartenant à un même cluster de relayage. Les clusters de relayage sont définis par la topologie et les propriétés de la transmission, telles que la bande.

La deuxième contribution de la Section 3.3 est la dérivation de la capacité du système et son analyse lorsque le nombre de nœuds relais augmente. Il s'avère que la capacité sature quand le nombre de nœuds augmente : au-delà d'un certain nombre de nœuds relais, l'essentiel de la puissance qui pouvait être récupérée à la destination grâce aux contributions des relais a déjà été collecté — d'où une saturation, et l'augmentation de la capacité due à des voies tardives résultant de retransmissions venant de nœuds éloignés devient négligeable. Attendre les retransmissions provenant de nœuds très éloignés de la source et de la destination n'en vaut pas la peine.

La troisième contribution de la Section 3.3 est l'évaluation de l'impact de la position des relais sur leur contribution à la capacité. Des résultats numériques montrent que quelques nœuds bien positionnés, à proximité de la source ou de la destination, conduisent à de meilleures performances qu'un plus grand nombre de nœuds répartis uniformément entre la source et la destination. Dans les contributions du Chapitre 3, aucune interaction coopérative entre les nœuds , autre qu'un relayage passif, n'était considérée. Les contributions des Chapitres 4 et 5 mettent l'accent sur les performances des réseaux ad hoc ou les nœuds ont des capacités coopératives plus avancées.

Les travaux du Chapitre 3 ont été publiés en partie dans :

- "Improving Ad Hoc Networks Capacity and Connectivity using Dynamic Blind Beamforming", N. Fawaz, Z. Beyaztas, M. Debbah, D.

Gesbert, *In Proc. of the 67th IEEE Vehicular Technology Conference, VTC Spring 2008*, May 11-14th 2008, Singapore

- "Capacity and Positioning in Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2007*, June 17-20th 2007, Helsinki, Finland

- "Capacity of Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of IRAMUS Workshop*, Jan. 25-26th 2007, Val-Thorens, FRANCE

Chapitre 4

Dans le chapitre 4, le canal à interférence coopératif est étudié dans le cas de coopération à la transmission. La première contribution du Chapitre 4 est le développement de nouvelles stratégies de coopération plus efficaces spectralement que les stratégies *Decode and Forward* classiques. Les stratégies proposées relaxent la contrainte d'orthogonalité, tout en préservant la contrainte de semi-duplex, et permettent à tous les symboles transmis de bénéficier de la diversité coopérative. Dans les stratégies proposées, la relaxation d'orthogonalité est inspirée du codage réseau et est réalisée en permettant à un nœud de combiner les messages provenant de différentes origines dans un unique signal transmis, avec une allocation de puissance optimisée. Les stratégies proposées contrastent avec d'autres stratégies, dites non-orthogonales, qui relaxent la contrainte d'orthogonalité en permettant à plusieurs nœuds de transmettre simultanément. De telles stratégies non-orthogonales ne permettent pas de garantir simultanément le respect de la contrainte de semi-duplex et le gain de diversité coopérative pour tous les symboles transmis.

La deuxième contribution du Chapitre 4 est l'introduction du codage *Dirty Paper* (DPC) dans les stratégies coopératives proposées, afin de mitiger l'interférence résultant de la relaxation d'orthogonalité. En effet, chaque source forme son signal transmis en combinant son propre message avec le message de l'autre source, chacun des messages combinés étant destiné à un récepteur différent. Les messages combinés représentent donc de l'interférence l'un pour l'autre. Cependant, cette interférence est connue au transmetteur, et peut donc être mitigée par un codage Dirty Paper, technique spécialisée pour la mitigation d'interférence connue à la transmission.

La troisième contribution du Chapitre 4 est l'évaluation, en termes d'efficacite spectrale, des mérites relatifs des stratégies coopératives proposées par rapport à ceux des stratégies coopératives classiques*Repetetion Decode and Forward* et *Parallel Decode and Forward*. La comparaison montre que, grâce à une utilisation plus efficace de la bande, les stratégies proposées améliorent le débit du réseau par rapport aux stratégies classiques.

Comme expliqué ci-dessus, le Chapitre 4 met l'accent sur l'amélioration de l'efficacité spectrale des stratégies coopératives dans des réseaux composés d'un petit nombre de paires communicantes, tandis que le Chapitre 5 est focalisé sur les performances du réseau quand le nombre de nœuds coopérant croit.

Les travaux du Chapitre 4 ont été publiés en partie dans :

- "When Network Coding and Dirty Paper Coding Meet in a Cooperative Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *in IEEE Trans. on Wireless Communications, vol. 7, no. 5, May 2008*

- "When Network Coding and Dirty Paper Coding Cooperate", N. Fawaz, D. Gesbert, M. Debbah, *In Proc. of the 9th IEEE Winter School on Coding and Information Theory*, March 12-16th 2007, La-Colle-sur-Loup, France

Chapitre 5

Dans le Chapitre 5, un réseau ad hoc dense constitué d'un grand nombre de paires source-destination communiquant avec l'aide d'un grand nombre de relais est considéré. Les nœuds sont regroupés en trois types de clusters coopératifs : un cluster source, un cluster destination, plusieurs clusters relai. Dans chaque cluster coopératif, les nœuds coopèrent pour former un réseau d'antennes virtuel, et interagissent pour recevoir ou transmettre conjointement de l'information. L'information circule du cluster source au cluster destination en traversant une série de clusters relai intermédiaires. Ce système forme un réseau de relayage MIMO virtuel multi-saut. Le canal à chaque saut est modélisé par une matrice de canal à évanouissement par bloc corrélée, et le traitement réalisé par chaque cluster est modélisé par une matrice de précodage. Pour faciliter l'analyse, deux hypothèses simplificatrices sont faites comme première étape vers une analyse future plus complète : on suppose que les nœuds dans un cluster constituent un réseau d'antennes virtuel

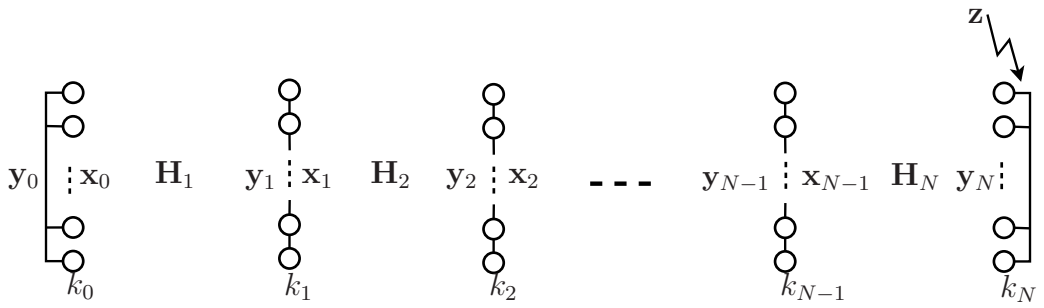


Figure 1.2: Réseau de clusters coopératifs multi-sauts

parfait, et que les communications entre clusters relai ne sont pas affectées par du bruit. Cependant, les signaux reçus par les nœuds dans le cluster destination sont altérés par du bruit.

La première contribution du Chapitre 5, résumée dans le *Théorème*[1] *1* ci-dessous, est la dérivation d'une expression explicite de l'information mutuelle bout-en-bout instantanée asymptotique entre l'entrée au cluster source et la sortie au cluster destination, quand le nombre de nœuds à tous les niveaux croît avec un taux fini. En utilisant des outils de la théorie des matrices aléatoires et des probabilités libres, on montre que l'information mutuelle instantanée par source converge vers une limite déterministe quand les dimensions du système croissent. On montre que cette expression asymptotique est indépendante de la réalisation du canal et dépend uniquement des statistiques du canal. En outre, dans le régime asymptotique, la valeur asymptotique de l'information mutuelle instantanée sert également de valeur asymptotique pour l'information mutuelle moyenne. De plus, on montre que même dans le cas d'un nombre de nœuds fini, le système se comporte comme dans le régime asymptotique. Cette observation rend la formule asymptotique un outil puissant pour gagner de l'intuition sur les performances du système, même lorsque les dimensions ne sont pas infiniment grandes.

Théorème 1. *Pour le système de clusters coopératifs décrit en Section 5.2, supposons que*

- *la destination a une connaissance parfaite du canal bout-en-bout* $\mathbf{G}_N$

[1] Le lecteur est renvoyé à la Section 5.2 pour la définition des notations impliquées dans les Théorèmes 1 et 2.

- *le nombre d'antennes à tous les niveaux* $k_0, k_1, \ldots, k_N$ *tend vers l'infini, tandis que* $\frac{k_i}{k_N} \to \rho_i, \qquad i = 0, \ldots, N$
- $\forall i \in \{0, \ldots, N\}$, *quand* k_i *tend vers l'infini,* $\mathbf{M}_i^H \mathbf{M}_i$ *a une distribution limite des valeurs propres à support compact.*

Alors l'information mutuelle par nœud source instantanée converge presque sûrement vers

$$\boldsymbol{I}_\infty = \frac{1}{\rho_0} \sum_{i=0}^{N} \rho_i \mathrm{E}\left[\log\left(1 + \eta \frac{a_{i+1}}{\rho_i} h_i^N \Lambda_i\right)\right] - N \frac{\log e}{\rho_0} \eta \prod_{i=0}^{N} h_i \tag{1.1}$$

où $a_{N+1} = 1$ *par convention,* $h_0, h_1, \ldots, h_N$ *sont les solutions du système de* $N+1$ *équations*

$$\prod_{j=0}^{N} h_j = \rho_i \mathrm{E}\left[\frac{h_i^N \Lambda_i}{\frac{\rho_i}{a_{i+1}} + \eta h_i^N \Lambda_i}\right] \qquad i = 0, \ldots, N \tag{1.2}$$

et l'espérance $\mathrm{E}[\cdot]$ *dans (1.1) et (1.2) est prise sur* Λ_i*, dont la distribution est la distribution asymptotique des valeurs propres of* $\mathbf{M}_i^H \mathbf{M}_i$*,* $F_{\mathbf{M}_i^H \mathbf{M}_i}(\lambda)$*.*

La deuxième contribution du Chapitre 5, résumée dans le *Théorème 2* ci-dessous, réside dans la dérivation de la structure des matrices de précodage optimales à chaque cluster, maximisant l'information mutuelle bout-en-bout moyenne sous l'hypothèse d'une connaissance locale statistique du canal (*local statistical CSI*) aux clusters source et relai. Notre analyse montre qu'à chaque cluster, les vecteurs singuliers à droite de la matrice de précodage optimale sont alignés avec les vecteurs propres de la matrice de corrélation de réception du canal au saut précédent, tandis que les vecteurs singuliers à gauche de la matrice de précodage optimale sont alignés avec les vecteurs propres de la matrice de corrélation de transmission du canal au saut suivant.

Théorème 2. *Considérons le système de clusters coopératifs décrit en Section 5.2. Pour* $i \in \{1, \ldots, N\}$*, soient* $\mathbf{C}_{t,i} = \mathbf{U}_{t,i} \mathbf{\Lambda}_{t,i} \mathbf{U}_{t,i}^H$ *et* $\mathbf{C}_{r,i} = \mathbf{U}_{r,i} \mathbf{\Lambda}_{r,i} \mathbf{U}_{r,i}^H$ *les décompositions en valeurs propres des matrices de corrélation de canal* $\mathbf{C}_{t,i}$ *et* $\mathbf{C}_{r,i}$*, où* $\mathbf{U}_{t,i}$ *et* $\mathbf{U}_{r,i}$ *sont unitaires et* $\mathbf{\Lambda}_{t,i}$ *et* $\mathbf{\Lambda}_{r,i}$ *sont diagonales, avec leurs valeurs propres respectives classées par ordre décroissant. Alors, sous les hypothèses de connaissance de canal* $\mathbf{A_s}$*,* $\mathbf{A_r}$ *et* $\mathbf{A_d}$*, les matrices de précodage linéaire optimales, maximisant l'information mutuelle moyenne sous*

les contraintes de puissance (5.6) sont données par :

$$\begin{aligned} \mathbf{P}_0 &= \mathbf{U}_{t,1}\mathbf{\Lambda}_{P_0} \\ \mathbf{P}_i &= \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{P_i}\mathbf{U}_{r,i}^H \text{ , for } i \in \{1,\ldots,N-1\} \end{aligned} \tag{1.3}$$

où $\mathbf{\Lambda}_{P_i}$ sont des matrices diagonales dont les elements diagonaux sont réels et positifs. En d'autres termes, les vecteurs singuliers des matrices de précodage optimales sont alignés sur les vecteurs propres des matrices de corrélation du canal. De plus, les vecteurs singuliers des matrices de précodage (1.3) sont également ceux qui maximisent l'information mutuelle moyenne asymptotique. Comme l'information mutuelle moyenne asymptotique a la même valeur que l'information mutuelle instantanée asymptotique, les vecteurs singuliers des matrices de précodage (1.3) sont finalement aussi optimaux pour l'information mutuelle instantanée asymptotique.

La troisième contribution du Chapitre 5 est l'analyse de la capacité du réseau dans plusieurs scenarios de communication, dans les cas de communication à saut unique ou multi-sauts, et de canaux décorrélés ou corrélés. Il est montré que les résultats sur la capacité des systèmes MIMO dérivés précédemment par d'autres auteurs [39, Section 3.3.2], [39, Theorem 3.7] peuvent être retrouvés en appliquant nos résultats sur l'information mutuelle asymptotique et la structure des précodeurs optimaux au cas de communications à saut unique. Nous fournissons également l'expression de la capacité asymptotique dans le cas de communications multi-sauts avec des canaux décorrélés ou à corrélation exponentielle, et nous montrons que le relayage coopératif améliore la capacité par nœud source même lorsque la taille du réseau augmente.

Les travaux du Chapitre 5 ont été publiés en partie dans :

- "Asymptotic Capacity and Optimal Precoding Strategy of Multi-level Precode & Forward in Correlated Channels", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *In Proc. of the IEEE Information Theory Workshop, ITW 2008*, May 5-9th 2008, Porto, Portugal

et ont été soumis en tant que :

- "Asymptotic Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *submitted to IEEE Trans. on Information Theory, Dec. 2008.*

Conclusion et Perspectives

Dans cette thèse, nous avons montré que la capacité de lien dans les réseaux ad hoc sans fil denses peut être améliorée, lorsque les nœuds sont dotés de capacités de coopération à la couche physique, et les stratégies de coopération sont efficacement conçues.

Les travaux de recherche de cette thèse peuvent être poursuivis dans plusieurs directions. Tout d'abord, les facteurs à la couche physique présentés dans le Chapitre 3 ne sont probablement pas les seuls facteurs permettant d'améliorer la capacité de lien dans les réseaux ad hoc sans fil dense, en réduisant l'interférence. En effet, le retournement temporel a longtemps été étudié comme méthode de focalisation d'une onde ultrasonique à la fois dans le temps et l'espace, et de récents travaux ont commencé à envisager l'application du retournement temporel aux systèmes de communication sans fil, en particulier aux systèmes à ultra large bande. Dans les réseaux ad hoc denses, la capacité de focalisation du retournement temporel pourrait permettre de mitiger l'interférence et donc améliorer la capacité de lien.

D'autre part, la plupart des protocoles pour réseaux coopératifs mentionnés dans les Chapitres 4 et 5 ont été conçus et analysés dans le régime haut SNR. Très peu de travaux ont ciblé le régime bas SNR des réseaux de relayage. Dans le régime large bande, alternativement dénommé régime bas SNR étant donné que la puissance est répartie sur un grand nombre de degrés de liberté, les performances ne sont pas limitées par l'interférence, mais par l'énergie. En tirant profit de la combinaison physique de signaux dans un lien sans fil, le codage réseau analogique apparaît comme une approche pertinente dans le régime bas SNR. Dans ce régime, le codage réseau pourrait surpasser d'autres approches cherchant essentiellement à éviter l'interférence, dont par exemple la radio cognitive.

Enfin, l'analyse de la capacité de lien dans un réseau ad hoc coopératif dense présentée dans le Chapitre 5 repose sur l'hypothèse simplificatrice d'une coopération parfaite à l'intérieur de chaque cluster. Les travaux de recherche doivent être poursuivis pour lever cette hypothèse simplificatrice, et prendre en compte le coût de la coopération intra-cluster sur les performances du réseau, en particulier en termes de connaissance du canal et feedback. De plus, une topologie particulière a été considérée dans le Chapitre 5 : les sources et les destinations pouvaient être regroupées en un unique cluster source et un unique cluster destination respectivement. Une analyse plus complète devrait fournir la capacité dans le cas de topologies plus générales.

Liste des Publications

Les travaux de recherche réalisés durant cette thèse ont conduit aux publications listées ci-dessous.

- Journaux
  - "Clip & Forward: Reaching the Min-Cut in Non-Coherent Wideband Multipath Fading Relay Channels", N. Fawaz, M. Médard, *in preparation for submission to IEEE Trans. on Information Theory*
  - "On the Asymptotic Capacity of Opportunistic Interference Alignment MIMO Networks", S. Medina Perlaza, N. Fawaz, S. Lasaulce, M. Debbah, *in preparation*
  - "Asymptotic Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *submitted to IEEE Trans. on Information Theory, 2008*
  - "When Network Coding and Dirty Paper Coding Meet in a Cooperative Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *in IEEE Trans. on Wireless Communications, vol. 7, no. 5, May 2008*

  Remarque : Deux papiers journaux n'ont pas encore été soumis, donc les travaux de recherche correspondants, bien que proches de l'achèvement, n'ont pas été inclus dans ce manuscrit.

- Conférences
  - "Improving Ad Hoc Networks Capacity and Connectivity using Dynamic Blind Beamforming", N. Fawaz, Z. Beyaztas, M. Debbah, D. Gesbert, *In Proc. of the 67th IEEE Vehicular Technology Conference, VTC Spring 2008*, May 11-14th 2008, Singapore
  - "Asymptotic Capacity and Optimal Precoding Strategy of Multi-level Precode & Forward in Correlated Channels", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *In Proc. of the IEEE Information Theory Workshop, ITW 2008*, May 5-9th 2008, Porto, Portugal

- "Large system design and analysis of protocols for decode-forward relay networks", L. Cottatellucci, T. Chan, N. Fawaz, *In Proc. of ICST 1st Workshop on Physics-Inspired Paradigms in Wireless Communications and Networks, WiOpt/PHYSCOMNET 2008*, Mar. 31st -Apr. 4th 2008, Berlin, Germany
- "Capacity and Positioning in Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2007*, June 17-20th 2007, Helsinki, Finland
- "When Network Coding and Dirty Paper Coding Cooperate", N. Fawaz, D. Gesbert, M. Debbah, *In Proc. of the 9th IEEE Winter School on Coding and Information Theory*, March 12-16th 2007, La-Colle-sur-Loup, France
- "Capacity of Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of IRAMUS Workshop*, Jan. 25-26th 2007, Val-Thorens, FRANCE

• Rapports Techniques

- "Large System Analysis of Relay Networks", L. Cottatellucci, N. Fawaz, *Chapter 5.3.2 in Deliverable DR6.1 of NEWCOM++ Work Package WPR6*, Apr. 2008
- "Network Design and Optimization: Ad-Hoc Network Infrastructure", N. Fawaz, *Chapter 4 in Deliverable D1.3.2 of BIONETS Work Package WP 1.3*, Aug. 2007
- "When Network Coding and Dirty Paper Coding Meet in a Cooperative Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *EURECOM Research Report, RR-07-199*, Sept. 2006

Chapter 2

Introduction

2.1 Overview and Motivations

Wireless ad hoc networks are flexible decentralized wireless networks where no fixed infrastructure is present. In contrast to infrastructure-based wireless networks, such as cellular networks where a base station centralizes communication of wireless nodes in a cell, or wireless local area networks (WLAN) where an access point manages the connection between wireless nodes, information transfer in ad hoc networks relies on the ability of wireless nodes to relay data for one another. Among the advantages of such networks are their flexibility, dynamism, autonomy, and self-organization capabilities that make them suitable for defense and emergency communication systems. Indeed ad hoc networks can be deployed dynamically on improvised terrains, and therefore historical applications of ad hoc networks count deployment in battle fields, rescue interventions, emergency deployment when natural disasters destroyed the pre-existing communication infrastructure. Recently, the expansion of Wireless Sensor Networks (WSN) and IEEE 802.11 (WiFi) WLAN which possesses a peer-to-peer mode enabling wireless devices to connect with each other, encouraged the emergence of commercial application ideas of ad hoc networks, e.g. gaming local area networks, community mesh networks, on-road vehicular networks, and coverage-extension of cellular networks through hybrid (cellular-ad hoc) networks. Indeed many ideas on hybrid networks, mixing cellular and multi-hop schemes, started emerging such as [3, Multi-hop Cellular Network (MCN)], [4, iCAR], [5, Self-Organizing Packet Radio Networks with Overlay(SOPRANO)]. Ad hoc and cellular networks were combined, with the prospects that bringing the advantages of both schemes together would allow to extend cell coverage while supporting dynamic topologies [3], to increase the scalability and reliability of pure ad hoc system, and to balance the load in cellular networks [4]. However those works focused mainly on routing issues.

The unpredictable and dynamic changes of topology in ad hoc networks triggered extensive research on high layers, mainly on routing protocols to provide fast adaptation in highly dynamic mobile ad hoc networks (MANETs). Traditionally, routing protocols in MANETS were focused on minimizing the hop-count, approach that did not take into account the link quality and led to routes with significantly less capacity than high-quality paths available in the network [6]. In [1] theoretical limits on the throughput of ad hoc networks were discovered, and were then confirmed practically by simulations and experiments in [2]. In [1], considering an ad hoc network of n nodes capable of

transmitting at W bits/second, randomly located on a unit-area disk, it was shown that the throughput per node decreased as $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ bits/second when the number of nodes n increased. Indeed in dense ad hoc networks, the wireless resource needs to be shared between concurrent transmissions of a large number of wireless nodes, and consequently the performance is limited by interference. The wireless transmissions of nodes need to be confined to their neighboring area, leading to multi-hopping for information to flow from a source to a destination. As a result, most transmissions in the network carry relayed data, which leads to a dramatic decrease of the total throughput. However, this scaling of the throughput was obtained under specific assumptions on the transmission mode at the physical layer: point-to-point multi-hop transmissions between fixed wireless nodes equipped with omni-directional antennas and transmitting their signals independently without any cooperative interaction with other nodes. Knowing that the performance at the physical layer upper-bounds the performance at high layers, the following questions naturally arise:

- Is it possible to improve the link capacity performance in dense ad hoc networks through the use of more advanced techniques at the physical layer?
- What are the physical layer factors that can improve the link capacity performance and what are their limits?

2.1.1 Factors Enhancing Ad Hoc Networks Performance

Several physical layer factors were proven to improve the performance of ad hoc networks, including antenna directivity, node mobility, node positioning, cooperation and virtual MIMO (Multiple-Input Multiple-Output).

Directional antennas can be used adaptively to enhance reliability and decrease interference [7–11]. Indeed, if nodes have information on their local topology, beamforming or transmission power sectorization can be used to focus the transmitted power in the direction of their receiver. Therefore the probability of interfering at non-intended receivers is decreased. Nevertheless knowledge of the position of the receiver is necessary at the transmitter in order to focus the transmission beam in the right direction. In high-mobility networks, tracking the location of a large number of nodes requires

a non-negligible feedback which increases the transmission protocol overhead [12–14]. Thus omni-directional antennas are mainly considered in ad hoc networks in spite of their negative impact on interference. One of the issues we address in Chapter 3 is the following:

- Is it possible to benefit from antenna directivity or beamforming while avoiding the feedback load in dense ad hoc networks?

Node mobility was also shown to improve the capacity of wireless ad hoc networks [15, 16]. By allowing wireless nodes to move, a constant scaling of the throughput can be obtained when the number of nodes increases. The idea consists in exploiting multi-user diversity through packet relaying in a dense network: a source splits its packet stream between several mobile wireless nodes that will act as relays and will deliver the packet they carry whenever they come close to the intended destination. Through their motion, mobile relays appear to have time-varying channels to the source and to the destination. Since the number of nodes in the network is large, at any time the probability for a relaying node to be close to the source and thus to obtain a new packet to carry is high, as well as the probability for a relaying node to be close to the destination and deliver a packet. As a result of motion, interference in the network is decreased, the number of hops a packet needs to travel through is reduced to two and the throughput is considerably increased. However, the constant scaling of the throughput in mobile ad hoc networks was obtained under the assumption of applications tolerant to large delays. In [17], the optimal throughput-delay trade-off was analyzed in mobile networks, and the delay required to sustain a constant scaling of the throughput was shown to depend on the velocity of wireless nodes: the delay increases when the velocity decreases, which happens when the network becomes more and more crowded because of an increasing density. This observation raises the question whether it is possible, through other factors, to improve the throughput in ad hoc networks without an increasingly high delay.

The position and number of relaying nodes supporting the communication of a source-destination pair have an impact on the capacity of a relaying network [18–22]. When only one source-destination pair in the network is active, and all other nodes help relaying data from the source, one

can analyze the contribution of relays to the capacity, and the scaling behavior of the capacity as the number of relaying nodes increases. In Chapter 3, we address the following questions:

- How does the link capacity grow with the number of relays helping a source-destination pair?
- What is the impact of the network topology, in particular the position of relaying nodes, on the link capacity performance?
- What is the impact of the physical environment, through reflections, propagation attenuation and delays... on the link capacity performance?

Cooperation and virtual MIMO can also contribute to performance improvements in wireless ad hoc networks. In the aforementioned models, nodes were considered to act independently and cooperation was reduced to passive forwarding: nodes did not interact, but could simply forward their received signal. If more advanced processing techniques are allowed at the physical layer and nodes can cooperate, one can expect the performance of ad hoc networks to be improved. Indeed, in point-to-point communications, MIMO techniques allow to improve the reliability of communications thanks to spatial diversity gains, and to increase the spectral efficiency through multiplexing gains. When considering a network of wireless nodes and allowing them to cooperate to jointly transmit information, a virtual MIMO system can be built in a distributed way, and cooperation can enable the exploitation of MIMO gains. In Chapters 4 and 5, the following issues are dealt with:

- How can we design cooperative strategies to improve the throughput in ad hoc networks while making efficient use of the wireless resource?
- What can we say on the capacity of a large multi-hop cooperative network, where wireless nodes have virtual MIMO capabilities?

2.1.2 Small Cooperative Networks

Cooperative communications occur when distributed wireless nodes interact to jointly transmit information. Several radio terminals relaying signals for each other form a virtual antenna array and their cooperation enables the

exploitation of spatial diversity in fading channels, which is then called cooperative diversity. Cooperative strategies were first designed for networks with small dimensions, which represent the building blocks for larger ad hoc networks. The most basic small cooperative networks are

- the relay channel (cf. Fig. 4.1(a)): one source-destination pair helped by a relay;
- the cooperative interference channel (cf. Fig. 4.1(b)): two source-destination pairs cooperating at transmit and/or receive side.

A plethora of cooperative strategies have been proposed for the relay channel or the cooperative interference channel [23, 24], the most famous ones being Amplify and Forward (AF), Decode and Forward (DF) and Compress and Forward (CF) [25]. The difference between those strategies lies in the processing performed by the relaying node before retransmission (cf. Chapter 4).

Most cooperative strategies have been designed to meet the practical half-duplex constraint: a radio terminal cannot transmit and receive simultaneously in the same frequency band, because the power of the received signal is very low compared to the power of the transmitted signal. Therefore, a relaying node uses orthogonal channels respectively to receive a signal from the source, and to transmit its relayed signal. For example, in the relay channel in Fig. 4.1(a), transmissions are divided into two blocks: in the first block, the source transmits and the relay and the destination receive; in the second block, the relay transmits its relayed signal and the destination receives. In the case of the cooperative interference channel in Fig. 4.1(b), usually the two-block scheme of the relay channel is simply extended to a four-block transmission scheme [25] by repeating twice the two-block transmission scheme: a first two-block scheme for the transmission of source S_1 relayed by S_2, followed by a second two-block scheme for the transmission of source S_2 relayed by S_1. Not only does the resulting cooperative transmission scheme meet the half-duplex constraint, but it is also interference-free. However, although the use of orthogonal interference-free channels for source and relay transmissions simplifies receivers algorithms, it results in an inefficient use of the bandwidth. Indeed only half of the degrees of freedom are used for transmission to each destination. A natural idea to use of the degrees of freedom more efficiently would be to relax the orthogonality constraint, but an obstacle appears straight: relaxing the orthogonality constraint would lead

to the introduction of interference in the system. Given those observations, in Chapter 4 we examine the following questions:

- Can we improve the spectral efficiency of cooperative strategies by relaxing the orthogonality constraint while still meeting the half-duplex constraint?

- How can we mitigate the interference due to the relaxation of the orthogonality constraint?

2.1.3 Large Cooperative Networks

The unfavorable scaling laws of the throughput in dense ad hoc networks in [1] raised the question whether ad hoc networks were only suited for small numbers of nodes or deployment in limited areas, or whether cooperative communications at the physical layer could allow ad hoc networks to be sustainably deployed for a large number of nodes. Consequently, [1] paved the way for recent research works on scaling laws in ad hoc networks where nodes have cooperative MIMO capabilities [26–37].

In an ideal— but unrealistic— ad hoc network, all nodes would be able to perfectly cooperate on a large scale and the network capacity would scale like in a perfect large MIMO system: linearly with the number of antennas [38]. A more realistic yet meaningful approach consists in grouping nodes in cooperative clusters: nodes belonging to a cooperative cluster cooperate to form a virtual antenna array and jointly transmit or receive information to or from other clusters. Then multi-hop communications occur between cooperative clusters, instead of single nodes as considered in [1]. With the cooperative approach, some of the transmissions that appeared as interference in [1] are now seen as useful signals that can be jointly processed by the nodes in a cluster.

A dense ad hoc network is a very complex system, whose performance metrics involve many variables and parameters. However, simulations are fortunately not the only way to gain some insight on how the performance of the system scales when the system dimensions increase. Recently, Random Matrix Theory (RMT) and Free Probability Theory (FPT) were discovered as appropriate theories to analyze/design complex communication systems and to reveal the relevant parameters impacting their performance [39]. Indeed, the transfer of information in a communication system can often be modeled

by a random matrix equation of the type $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$, where $\mathbf{x}$ is the input vector, $\mathbf{y}$ is the output vector, $\mathbf{z}$ is a noise vector, and $\mathbf{H}$ is the transfer matrix of the system. For such a system, most information-theoretic performance metrics can be shown to depend only on the eigenvalues and eigenvectors of the transfer matrix $\mathbf{H}$, and RMT and FPT provide useful results on the eigenvalues and eigenvectors of random matrices with large dimensions that can be applied to the analysis of large dimension communication systems. Random matrix theory emerged with the works of Wishart [40], Wigner [41], and Marčenko and Pastur [42], and was historically used in physics, before being applied in many other fields. In wireless communications, RMT was first used to analyze the performance of communication systems using multiple antennas, e.g. in [38, 43–45], and CDMA systems, e.g. in [46–48].

In Chapter 5, we consider a dense ad hoc networks where nodes are grouped in cooperative clusters, and using tools from RMT, FPT and linear algebra, we address the following questions:

- What is the asymptotic capacity of the cooperative-cluster system when the number of nodes in all clusters grow large?
- What are the relevant parameters impacting the system capacity?
- How should nodes in a cooperative cluster process and transmit cooperatively their wireless signals to maximize the system capacity?

2.2 Contributions

2.2.1 Chapter 3

The contribution of Chapter 3 is twofold. In the first part, we address the impact of antenna directivity on the throughput and connectivity of a dense network with a large number of source-destination pairs. In the second part, we analyze the impact of the position and number of passive relaying nodes on the capacity of a system where a large number of relaying nodes support the communication of a source-destination pair.

In Section 3.2, a dense network of source-destination pairs, where sources are equipped with directional antennas, is considered. The first contribution of Section 3.2 is the proposition of a dynamic blind beamforming scheme that allows to benefit from antenna directivity in a decentralized network,

while avoiding heavy feedback to track the position of nodes. The scheme is dynamic and blind since a source points its directive antenna successively in all directions to surely but blindly hit its destination without knowing its exact position. We show that rotational directivity has a positive impact on interference reduction, and thus on capacity: by focusing the transmitted power successively in different directions, the probability of interfering with other destinations, i.e. hitting a non-intended destination at the same time it is receiving a signal from its own source, is low because of both spatial focusing and asynchronism of all communications. However, rotational directivity introduces some delay: when a source is not beamforming in the direction of its intended destination, time and power are wasted. These two opposite effects lead to a capacity-delay trade-off when adjusting the number of rotations.

The second contribution of Section 3.2 is the analysis of the network throughput with dynamic blind beamforming, and the comparison to the classical case where sources are equipped with omni-directional antennas. We show that when the density of the network increases, our dynamic blind beamforming scheme outperforms omni-directional transmissions. Indeed, as the network density increases, the throughput is interference-limited and narrower transmission beams are necessary to decrease interference. However, when transmission beams become narrower, a larger number of rotations of the source antenna is necessary to cover the whole space, leading to an increased delay. For a given network density, the trade-off between interference-reduction and delay-increase results in an optimal beam-width and an optimal number of rotations maximizing the throughput of the network.

The third contribution of Section 3.2 is the definition of a new connectivity criterion, namely throughput-connectivity, that accounts for interference and for the shared access to the wireless resource by nodes in the network. The throughput-connectivity with a target rate R is the proportion of pairs in the network to which a throughput of R can be granted. For a target rate R, the number of rotations and the beam-width maximizing the connectivity depend on the network density.

In Section 3.3, we examine a system where a source communicates with a destination with the help of a dense network of passive relays. Relays are modeled as dumb omnidirectional scatterers, i.e. passive nodes without engineering capabilities that simply scatter the incident electromagnetic wave coming from the source antenna. The first contribution of Section 3.3 is the

formulation of a model taking into account the physical environment through reflections, pathloss, delays and multi-path. We show that the asynchronous relaying system can be modeled as a virtual multi-path channel, where each path gain results from the combination of signal contributions from relays belonging to a relaying cluster. Relaying clusters are defined by topology and transmission properties, such as bandwidth.

The second contribution of Section 3.3 is the derivation of the capacity of the system and its analysis as the number of relaying nodes increases. It turns out that capacity saturation occurs when the number of nodes increases: beyond a certain number of relaying nodes, most of the power that could be recovered at the destination thanks to relayed contributions has already been collected— thus a saturation, and the increase in capacity resulting from late paths due to retransmissions from very far nodes becomes negligible. Waiting for retransmissions from nodes located very far from source and destination is not worth.

The third contribution of Section 3.3 is the evaluation of the impact of the position of relaying nodes on their contribution to capacity. Numerical results show that a few relaying nodes well located, close to source or destination, lead to better performances than a larger number of nodes uniformly distributed between source and destination.

In the contributions of Chapter 3, no cooperative interaction between nodes, other than passive relaying, was considered. The contributions of Chapters 4 and 5 focus on the performance of wireless ad hoc networks where nodes have more advanced cooperation capabilities.

Part of the work in Chapter 3 has been published in:

- "Improving Ad Hoc Networks Capacity and Connectivity using Dynamic Blind Beamforming", N. Fawaz, Z. Beyaztas, M. Debbah, D. Gesbert, *In Proc. of the 67th IEEE Vehicular Technology Conference, VTC Spring 2008*, May 11-14th 2008, Singapore

- "Capacity and Positioning in Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2007*, June 17-20th 2007, Helsinki, Finland

- "Capacity of Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of IRAMUS Workshop*, Jan. 25-26th 2007, Val-Thorens, FRANCE

2.2.2 Chapter 4

In Chapter 4, we study the cooperative interference channel with transmit cooperation. The first contribution of Chapter 4 is the development of novel cooperative strategies that are more spectrally efficient than classical Decode and Forward strategies. The proposed strategies relax the orthogonality constraint, yet preserve the half-duplex constraint and allow all transmitted symbols to benefit from cooperative diversity. In the proposed strategies, the orthogonality-relaxation is inspired from network coding and is achieved by allowing a transmitting node to combine messages from different origins in a single transmitted signal, with smart power allocation. The proposed strategies are in contrast with other strategies, called non-orthogonal strategies, which relax the orthogonality constraint by allowing several nodes to transmit at the same time. Such non-orthogonal strategies cannot guaranty that both the half-duplex constraint will be met and all transmitted signals will benefit from cooperative diversity.

The second contribution of Chapter 4 is the introduction of Dirty Paper precoding in the proposed cooperative strategies, in order to mitigate the interference resulting from the orthogonality-relaxation. Indeed, each source forms its transmitted signal by combining its own message with a message from the other source, each message being intended to a different destination. Thus the combined messages act as interference for each other. However the interference is known by the transmitting source, which can mitigate it thanks to Dirty Paper Coding, a well-known technique to mitigate interference known at transmitter.

The third contribution of Chapter 4 is the evaluation of the relative merit of the proposed cooperative strategies with respect to classical cooperative strategies, in terms of spectral efficiency. The comparison shows that thanks to a more efficient use of the bandwidth, the proposed strategies improve the network throughput with respect to classical Repetition Decode and Forward and Parallel Decode and Forward strategies.

Chapter 4 laid the emphasis on improving the spectral efficiency of cooperative strategies in networks with a small number of communicating pairs. Chapter 5 focuses on the network performance when the number of cooperating nodes grows large.

Part of the work in Chapter 4 has been published in:

- "When Network Coding and Dirty Paper Coding Meet in a Cooperative

Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *in IEEE Trans. on Wireless Communications, vol. 7, no. 5, May 2008*

- "When Network Coding and Dirty Paper Coding Cooperate", N. Fawaz, D. Gesbert, M. Debbah, *In Proc. of the 9th IEEE Winter School on Coding and Information Theory*, March 12-16th 2007, La-Colle-sur-Loup, France

2.2.3 Chapter 5

In Chapter 5, we consider a dense ad hoc network with a large number of source-destination pairs communicating with the cooperative support of a large number of relays. Nodes are grouped in three types of cooperative clusters: source cluster, destination cluster, relaying clusters. In each cooperative cluster, nodes cooperate to form a virtual antenna array, and interact to jointly receive or transmit information. The information flows from the source cluster to the destination cluster through a series of intermediary relaying clusters. This system forms a virtual MIMO multi-hop relay network. The channel at each hop is modeled by a block-fading correlated channel matrix, and the processing performed by each cluster is model by a precoding matrix. To ease the analysis, a couple simplifying assumptions are made, as a first step towards a future more complete analysis: nodes in a cluster are assumed to form a perfect virtual antenna array, and communications between relaying clusters is assumed to be non-noisy. However, signals received by nodes in the destination cluster are assumed to be impaired by noise.

The first contribution of Chapter 5 is the derivation of a closed-form expression of the asymptotic instantaneous end-to-end mutual information between the input of the source cluster and the output of the destination cluster, as the number of nodes at all levels grow large with a finite rate. Using tools from random matrix theory and free probability theory, the instantaneous mutual information per source is shown to converge to a deterministic value as the system dimensions grow large. This asymptotic expression is shown to be independent from the channel realizations and to only depend on the channel statistics. Besides, in the asymptotic regime, the asymptotic value of the instantaneous mutual information is shown to also serve as the asymptotic value of the average mutual information. Furthermore, we show that with a finite number of nodes, the system behaves closely to the asymptotic regime, making the asymptotic formula a powerful tool to gain insight

on the system performance even when the dimensions are not infinitely large.

The second contribution of Chapter 5 consists in providing the structure of the optimal precoding matrices, at each cluster, maximizing the end-to-end average mutual information under the assumption of local statistical CSI at source and relaying clusters. Our analysis shows that at each cluster, the right singular vectors of the optimal precoding matrices are aligned to the eigenvectors of the receive correlation matrix of the backward channel, while the left singular vectors of the optimal precoding matrices are aligned to the eigenvectors of the transmit correlation matrix of the forward channel.

The third contribution of Chapter 5 is the analysis of the network capacity in several communication scenarios in the case of single-hop or multi-hop communications, and uncorrelated or correlated channels. We show that results on the capacity of MIMO systems formerly derived by other authors [39, Section 3.3.2], [39, Theorem 3.7] can be recovered by applying the aforementioned results on the asymptotic mutual information and the optimal precoder structure to the case of single-hop communications. We also provide the expression of the asymptotic capacity in the case of multi-hop communications with uncorrelated or exponentially correlated channels, and show that cooperative relaying improves the capacity per source node even when the network size increases.

The work in Chapter 5 has been published in part in:

- "Asymptotic Capacity and Optimal Precoding Strategy of Multi-level Precode & Forward in Correlated Channels", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *In Proc. of the IEEE Information Theory Workshop, ITW 2008*, May 5-9th 2008, Porto, Portugal

and has been submitted as:

- "Asymptotic Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *submitted to IEEE Trans. on Information Theory, Dec. 2008.*

As a conclusion, the link capacity performance in dense wireless ad hoc networks can be improved, as long as nodes are empowered with cooperative capabilities at the physical layer, and cooperative strategies are properly designed.

2.3 Publications

The research work performed during this thesis led to the publications listed below.

- Journals
  - "Clip & Forward: Reaching the Min-Cut in Non-Coherent Wideband Multipath Fading Relay Channels", N. Fawaz, M. Médard, *in preparation for submission to IEEE Trans. on Information Theory*
  - "On the Asymptotic Capacity of Opportunistic Interference Alignment MIMO Networks", S. Medina Perlaza, N. Fawaz, S. Lasaulce, M. Debbah, *in preparation*
  - "Asymptotic Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *submitted to IEEE Trans. on Information Theory, Dec. 2008*
  - "When Network Coding and Dirty Paper Coding Meet in a Cooperative Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *in IEEE Trans. on Wireless Communications, vol. 7, no. 5, May 2008*

  Note that two journal papers have not yet been submitted, thus the corresponding research work, though close to completion, could not be included in this thesis.

- Conferences
  - "Improving Ad Hoc Networks Capacity and Connectivity using Dynamic Blind Beamforming", N. Fawaz, Z. Beyaztas, M. Debbah, D. Gesbert, *In Proc. of the 67th IEEE Vehicular Technology Conference, VTC Spring 2008*, May 11-14th 2008, Singapore
  - "Asymptotic Capacity and Optimal Precoding Strategy of Multi-level Precode & Forward in Correlated Channels", N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, *In Proc. of the IEEE Information Theory Workshop, ITW 2008*, May 5-9th 2008, Porto, Portugal

- "Large system design and analysis of protocols for decode-forward relay networks", L. Cottatellucci, T. Chan, N. Fawaz, *In Proc. of ICST 1st Workshop on Physics-Inspired Paradigms in Wireless Communications and Networks, WiOpt/PHYSCOMNET 2008*, Mar. 31st -Apr. 4th 2008, Berlin, Germany
- "Capacity and Positioning in Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2007*, June 17-20th 2007, Helsinki, Finland
- "When Network Coding and Dirty Paper Coding Cooperate", N. Fawaz, D. Gesbert, M. Debbah, *In Proc. of the 9th IEEE Winter School on Coding and Information Theory*, March 12-16th 2007, La-Colle-sur-Loup, France
- "Capacity of Dense Scattering Environments", N. Fawaz, M. Debbah, D. Gesbert, *In Proc. of IRAMUS Workshop*, Jan. 25-26th 2007, Val-Thorens, FRANCE

• Technical Reports

- "Large System Analysis of Relay Networks", L. Cottatellucci, N. Fawaz, *Chapter 5.3.2 in Deliverable DR6.1 of NEWCOM++ Work Package WPR6*, Apr. 2008
- "Network Design and Optimization: Ad-Hoc Network Infrastructure", N. Fawaz, *Chapter 4 in Deliverable D1.3.2 of BIONETS Work Package WP 1.3*, Aug. 2007
- "When Network Coding and Dirty Paper Coding Meet in a Cooperative Ad Hoc Network", N. Fawaz, D. Gesbert, M. Debbah, *EURECOM Research Report, RR-07-199*, Sept. 2006

Chapter 3

Factors Improving Ad Hoc Networks Performance

3.1 Introduction

Wireless ad hoc networks are decentralized communication networks without a fixed infrastructure or centralized administration, as opposed to cellular networks for example. In ad hoc networks, wireless nodes are not only in charge of transmitting their own information, but also responsible for relaying packets from other nodes, to the benefit of the network. These networks have the particularity to be self-organized, autonomous and flexible systems. Thus ad hoc networks can be deployed with a great flexibility on improvised terrains, and are naturally relevant for defense and security applications: dynamic deployment in battle fields, security and rescue interventions, emergency deployment for example in the case of a natural catastrophe which destroyed the pre-existing communication infrastructure... Besides historical military applications of ad hoc networks, ideas of commercial applications started emerging recently, such as on-road vehicular networks, community mesh networks, gaming local area networks, hybrid networks where ad hoc networks would help extending the coverage of cellular networks... The perspective of adopting ad hoc networks in commercial applications has been encouraged by the rapid expansion of IEEE 802.11 (WiFi) networks and Wireless Sensor Networks (WSN).

On the other hand, the topology of wireless ad hoc networks can evolve in an unpredictable way and difficulties occur in those wireless networks because of constraints in energy autonomy, delay, received signal quality in presence of interference... For long, research on ad hoc networks focused on high layer— link, network, and transport layers— issues in an extensive way. In particular, at the network layer, distributed algorithms are necessary to cope with frequent wireless topology changes, that generate frequent updates of routing tables. As a result, several adaptive routing protocols for Mobile Ad Hoc NETworks (MANETS) were developed, such as DSR (Dynamic Source Protocol) [49], AODV (Ad-hoc On-Demand Distance Vector Routing) [50], OLSR (Optimized Link State Routing Protocol) [51], ZRP (Zone Routing Protocol) [52]... Nevertheless, in these works focused on routing, the physical layer remains unchanged: transmissions at the physical layer are point-to-point though possibly multi-hop, wireless nodes are treated as single processing units, and cooperation between nodes is strictly limited to storing and forward packets.

Recently, theoretical [1] and practical [2] limits to the performance of ad hoc networks were revealed, and they gave the impression that the per-

formance of ad hoc networks was doomed to a dramatic decrease when the density of the network increased. In [1], dense ad hoc networks with n randomly placed nodes, capable of transmitting at a rate W bits/second and operating in the multi-hop store-and-forward mode, were considered. It was shown that the throughput per node $T(n)$ decreased as $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ when the node density increased. This dramatic throughput scaling in dense networks is due to the need for nodes to concurrently share the wireless channel with neighboring nodes, thus the performance of such dense ad hoc networks is interference limited. However, these results were obtained under the assumptions that transmissions at the physical layer were point-to-point multi-hop transmissions between fixed wireless nodes, equipped with omni-directional antennas, and transmitting their signals independently without any cooperative interaction. Given that observation and given that the performance at high layers is upper-bounded by the performance at the physical layer, the following questions naturally arise:

- Is it possible to improve the link capacity performance in dense ad hoc networks through the use of more advanced techniques at the physical layer?
- What are the physical layer factors that can improve the link capacity performance and what are their limits?

Recent works revealed some of the factors improving the performance of ad hoc networks, including antenna directivity, node mobility, node positioning, cooperation. We review the aforementioned factors hereunder, and open the way to our contributions on the impact of antenna directivity and node positioning in the following sections.

3.1.1 Antenna Directivity and Beamforming

In ad hoc networks, in particular dense thus interference-limited networks, directional antennas can improve performance when the destination position is known [7–11]. Indeed, directive antennas have the capability to focus the transmitted power in a given direction, which has a double positive effect. First, focusing the power allows to increase the transmission range of a node or the received power at the destination node, thus to combat pathloss effects. This is particularly helpful in low-density widely-spread ad hoc networks, whose performance is known to be power/coverage limited (cf. Section 5.1.1).

Second, by focusing the transmitted power in a narrow beam, the probability of a non-intended destination to be in the transmission beam is smaller than with omni-directional antennas. Thus, the probability to interfere with a non-intended destination is lower with directional antennas. This is all the more relevant in dense ad hoc networks, whose performance is limited by interference (cf. Section 5.1.1).

Nevertheless the use of directive antennas requires the transmitter to have knowledge of the position of the receiver, in order to focus the transmission beam in the right direction. In a high-mobility context, the feedback required for tracking the location of a large number of nodes increases the transmission protocol overhead [12–14], and thus reduces the useful rate and consequently leads to considering mainly omnidirectional antennas in ad hoc networks. The question we address in Section 3.2 is how to benefit from directional antennas or beamforming while avoiding the feedback load in dense ad hoc networks.

3.1.2 Node Mobility

Mobility can increase the capacity of wireless ad hoc networks [15,16]. In [15], a network of n mobile nodes moving in a disk of unit area is considered. Assuming that node motion processes are mutually independent and have a stationary ergodic uniform distribution, it is shown that the average long-term throughput per source-destination pair can have a constant scaling when the number of nodes increases. The constant scaling is achieved by having a source split its data stream between many mobile nodes. Each mobile node carries a different packet from the source, and delivers it to the destination whenever its trajectory leads it near the destination. Therefore, mobile nodes appear as nodes with a time-varying channel to the destination, and they deliver only when their channel quality is good, keeping the number of hops per packet to two and interference low. Because of the large number of nodes in the network, the probability that at least one mobile node is near the destination is high. Thus, mobility allows to exploit a form of multi-user diversity through packet relaying. In [16] those scaling results were shown to hold even when the nodes have a more restrictive one-dimensional mobility pattern. The constant scaling of the throughput obtained thanks to mobility is far more encouraging than the throughput scaling in fixed networks in [1].

However, the previous throughput-scaling results in [15,16] were obtained under the loose assumption of delay-tolerant applications. The scaling of the

delay in fixed and mobile ad hoc networks was analyzed in [17], and the optimal throughput-delay tradeoff were provided.

- For the fixed network described in [1], the throughput-delay tradeoff is $D(n) = \Theta(nT(n))$, where $D(n)$ and $T(n)$ denote the average delay and throughput respectively. For an optimal throughput $T(n) = \Theta\left(\frac{1}{\sqrt{n \log n}}\right)$, the delay increases as $D(n) = \Theta\left(\sqrt{\frac{n}{\log n}}\right)$. Fixed dense networks are interference-limited, therefore transmissions are confined to neighboring nodes to limit interference, resulting in multi-hopping and high delays.

- For the mobile network in [15], with a throughput scaling as $T(n) = \Theta(1)$, the delay scales as $D(n)O(\frac{\sqrt{}(n)}{v(n)})$ where $v(n)$ is the velocity of mobile nodes. By moving, nodes can come close to each other, which allows to decrease the number of hops per packet to two, i.e. the use of a single relay per packet and lead to a constant scaling of the throughput. Nevertheless the speed at which the node moves becomes the main factor impacting the delay. Node velocity scales down as the density increases— intuitively, when the density increases, the area is more crowded and nodes will move slower— thus the delay will increase as the network becomes denser.

Mobility allows to improve ad hoc networks throughput, but at the cost of an increasing delay. The question arises to know whether other factors or techniques could further improve the performance of ad hoc networks, without requiring an increasingly large delay.

3.1.3 Node Positioning

Considering a source-destination pair in a fixed network, the number and position of nodes relaying information for the pair has an impact on the performance of the resulting system [18–22]. The questions we pose in Section 3.3 are:

- How does the link capacity grow with the number of relays helping a source-destination pair?

- What is the impact of the network topology, in particular the position of relaying nodes, on the link capacity performance?

- What is the impact of the physical environment, through reflections, propagation attenuation and delays... on the link capacity performance?

It turns out that when the number of non-noisy relaying nodes increases, capacity saturation occurs, more or less faster depending on the position of relaying nodes. Thus having a very large number of relays serving a communicating pair is not necessary, only a few well-located relays per pair may be sufficient. However, those results rely on the assumption that nodes do not interact, but simply forward their received signal as single isolated units. In a network with several communicating pairs, helped by several relays, allowing cooperation between nodes and enabling more advanced transmission techniques may be required for a sustainable scaling of the throughput with the network density.

3.1.4 Cooperation and Virtual MIMO

In previous models, nodes were considered to act independently and were not allowed to cooperate actively to transmit information. Cooperative communications occur when distributed wireless terminals interact to jointly transmit information. When several radio terminals relay signals for each other, they form a virtual antenna system and their cooperation enables the exploitation of gains, that are usually the prerogative of MIMO systems, such as spatial diversity in fading channels and multiplexing. Cooperation leads to improved performance in wireless networks, and the reader is referred to Chapters 4 and 5 for a more detailed discussion on cooperative networks and virtual MIMO.

Our contributions on the impact of antenna directivity and node positioning on the performance of ad hoc networks are presented in Sections 3.2 and 3.3 respectively, while our contributions on cooperation and virtual MIMO span Chapters 4 and 5.

3.2 Antenna Directivity impact

3.2.1 Introduction

In this section, we analyze the impact of antenna directivity on the throughput and connectivity of a dense network.

Contribution

We propose a dynamic blind beamforming scheme which allows to benefit from antenna directivity in large wireless ad hoc networks while avoiding heavy feedback to track mobile nodes localization, usually due to mobility and density. The scheme is dynamic and blind since a source uses a rotating antenna successively in all directions to surely but blindly hit its destination without knowing its exact position. If position is known with a certain accuracy, for example thanks to limited feedback, a source can semi-blindly form beams in a subset of directions. Rotational directivity has a major impact on interference and thus on capacity: by focusing the transmitted power successively in different directions, the probability of interfering with other destinations, i.e. hitting a non-intended destination at the same time it is receiving a signal from its own source, is low because of both spatial focusing and asynchronism of all communications. Nevertheless when rotating the antenna, some time and power is wasted when the source is not beamforming in the direction of its intended destination. These two opposite effects lead to a capacity-delay trade-off when tuning the number of rotations. We analyze performance in terms of total network throughput and connectivity and we show that our scheme can outperform omnidirectional transmissions in interference-limited dense ad hoc networks both in terms of capacity and connectivity. The optimal number of rotations, maximizing the network performance, depends on the density of the network and results from a trade-off between delay and improvements in terms of interference

Related Works

Recently Sharif and Hassibi [53] proposed a random beamforming scheme for the Multi-user MIMO Broadcast channel in which the transmitter constructs random beams and transmits to the users with the highest SINRs, fedback to the transmitter. When the number of users increases, the capacity was shown to scale as with perfect CSI at the transmitter. Nevertheless this random beamforming model relies on feedbacks from mobile units to a base station in a cellular system and the served-destinations are chosen according to the quality of their link to the base station for a given set of random beams.

On the other hand Bettstetter et al. [54] showed that random beamforming in ad hoc networks can improve received-power-connectivity: sources send

random beams in a random direction they chose once and for all at the beginning and any node whose received power from a source is above a threshold is considered connected to the source. No source and destination are associated in a communicating pair idea, i.e. a source ignores not only the position but also the identity of the destinations. Any nodes whose signal reaches a receiver with a high enough power is considered connected to the receiver, and that model also ignores interference that occurs when two sources hit a destination at the same time. The received-power-connectivity criterion does not take into account interference as an SINR (Signal-to-Interference-plus-Noise Ratio) criterion would do, nor reliable decoding issues that are usually illustrated by capacity or BER (Bit Error Rate).

.

3.2.2 System Model

Consider the $2D$-network of M communicating pairs $\{S_iD_i\}_{i\in\{1,\ldots,M\}}$ uniformly distributed over a square of area a^2 square meters, illustrated in Fig. 3.1. $d_{ji} = S_iD_j$ denotes the distance between source S_i and destination D_j. All nodes are equipped with a single antenna, directional at sources and omnidirectional at destinations.

At time $t = kT$, the signal transmitted by source S_i is denoted $x_i(k)$ whereas $y_j(k)$ represents the signal received by destination D_j. The channel between transmitter S_i and receiver D_j is represented by h_{ji} which includes the effects of shadowing and slow flat fading. These channel coefficients are modeled by independent circularly-symmetric complex gaussian random variables with zero mean and variance σ_{ji}^2, i.e. Rayleigh fading. z_j denotes the i.i.d circularly-symmetric complex gaussian noise at destination D_j , with zero mean and variance σ^2.

Each source S_i generates a sequence of packets, and each packet consists of N_s symbols $s_i(m)$, $m \in \{0, \ldots, N_s - 1\}$. These symbols are modeled by independent identically distributed (i.i.d.) circularly-symmetric complex gaussian random variables with zero mean.

Sources have the ability to rotate their directional antenna, selecting a different transmission direction at each time-slot. The duration of a time-slot is N_sT and corresponds to the transmission of the N_s symbols contained in a packet. N denotes the number of times a source rotates its directional antenna to transmit the same packet of N_s symbols repetitively in N time-slots, pointing at a different direction during each time slot with a beamwidth

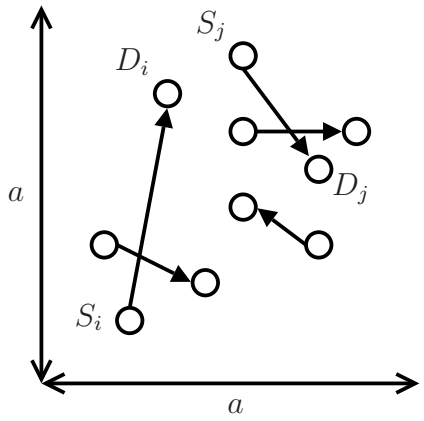


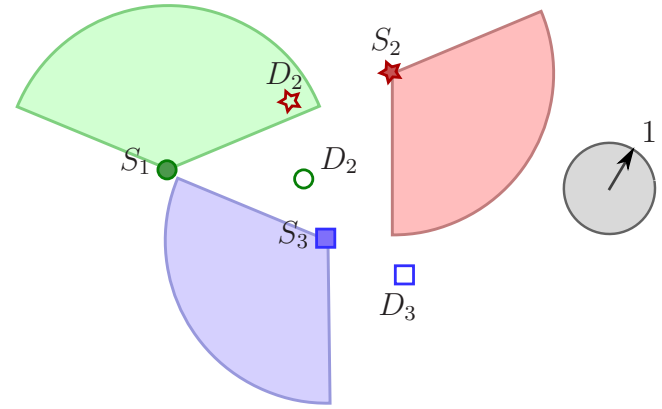
Figure 3.1: $2D$-Network

$\alpha = \frac{2\pi}{N}$. $N = 1$ corresponds to the case where source antennas are omnidirectional and transmit a packet of N_s symbols only once. After N time-slots, the source has sent the same packet N times in N successive directions, covering the whole 2π-space, like a lighthouse operating in a discrete fashion. Fig. 3.2.2 illustrate the blind beamforming scheme in a network with $M = 3$ source-destination pairs and $N = 3$ rotations. Note that when rotating the antenna, one cannot switch at the symbol level, but only at the packet level. Indeed, switching direction at symbol level would create a Doppler shift larger than the bandwidth of the signal.

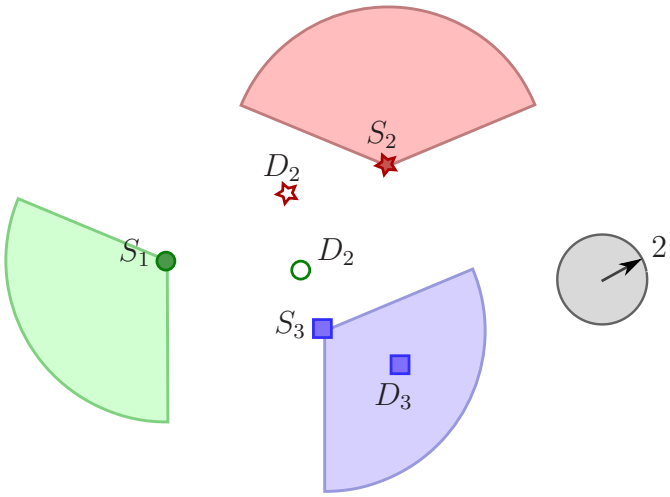
Interferers Groups

Consider the communicating pair S_iD_i and an arbitrary packet. For each symbol $s_i(m)$ in the packet that S_i transmits N times, D_i receives the symbol only once, during the time-slot when D_i is located in the rotating beam of S_i. Any other source whose beam would cover D_i during the time-slot when D_i receives a signal from S_i belongs to the group $\mathcal{I}_i$ of interferers of D_i. A source whose beam would cover D_i in a slot where D_i is not receiving any signal from S_i is not an interferer.

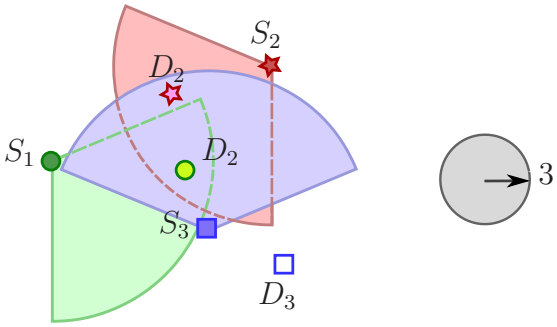
In the omnidirectional case $N = 1$, all other sources are interferers, therefore $\forall i \in \{1, \ldots, M\}$, $\mathcal{I}_i = \{S_j / j \neq i\}$ contains $M - 1$ interferers. When $N > 1$, rotating the antennas clearly allows to decrease the number of interferers per destination with respect to the omnidirectional case and the group of interferers of a destination depends on the network topology and the initial



(a) First Rotation



(b) Second Rotation



(c) Third Rotation

Figure 3.2: Dynamic blind beamforming with $N = 3$

transmission direction chosen by each source.

When D_i receives a signal that does not contain any component coming from S_i, recognized for example thanks to an embedded signature identifying S_i, D_i simply discards the received signal. One could argue that in a static network, when D_i recognizes a signal component from S_i, it could send a feedback to S_i which would then identify the direction in which to beamform. But in a high mobility context, tracking D_i moving position would lead to heavy overhead, which our blind dynamic beamforming strategy intends to avoid.

When D_i is in the transmission beam of S_i at time $t = kT$, the received signal $y_i(k)$ at D_i is the sum of the signals transmitted by S_i and all sources in $\mathcal{I}_i$ filtered by their respective channels, and noise $z_i(k)$.

Transmitted Power and Energy

We use the simple ideally-sectorized directional antenna model to describe the gain pattern, as proposed in [54] Equation (4). As illustrated in Fig. 3.3, it is assumed that at time t, the transmit antenna of S_i forms a beam of width α in the direction $\alpha_i(t)$ with a certain gain. $\alpha_i(0)$ denotes the initial direction chosen at random by S_i during the first time-slot, then every time-slot, S_i rotates its antenna anti-clockwise of an angle α to get the new direction. Thus $\alpha_i(t) = \alpha_i(0) + \lfloor \frac{t}{N_s T} \rfloor \alpha$.

Each source has a power constraint in the continuous time-channel of P_0 Joule/s. In the omnidirectional case, P_0 is transmitted over the whole 2π space with an angular density of power $P_0/2\pi$, whereas in the directional case $N > 1$, P_0 is focused in an angle $\alpha = 2\pi/N$ leading to the power angular density at time t in direction θ:

$$\frac{\partial P_i}{\partial \theta}(\theta, t) = \frac{P_0}{\alpha} \mathbb{1}_{[\alpha_i(t)-\alpha/2;\alpha_i(t)+\alpha/2[}(\theta) \tag{3.1}$$

Indeed since a source transmits only in $(1/N)^{th}$ of the space, it can increase its transmit power in its transmission beam to $P_0/\alpha = NP_0/2\pi$ and remain within its average power constraint for the whole space. The power angular density depends on time since the transmit antenna rotates. On the contrary, the total power transmitted by source S_i does not depend on time, nor on the number of rotations N and respects the power constraint by definition:

$$P_i = \int_{\theta=0}^{2\pi} \frac{\partial P_i}{\partial \theta}(\theta, t) \partial\theta = \int_{\theta=\alpha_i(t)-\alpha/2}^{\alpha_i(t)+\alpha/2} \frac{P_0}{\alpha} \partial\theta = P_0 \tag{3.2}$$

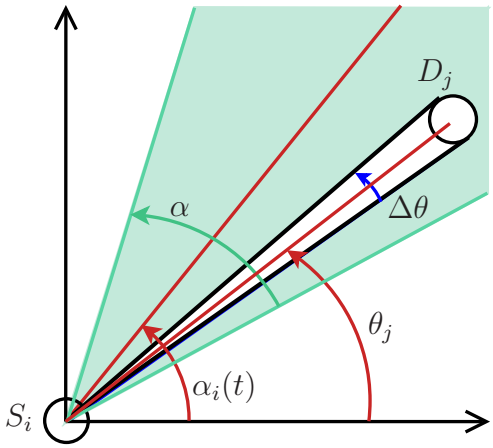


Figure 3.3: Transmission between S_i and D_j

The transmitted power is by definition $P_i = \frac{\partial \varepsilon_i}{\partial t}$, where ε_i is the energy transmitted by S_i. Since source S_i transmits the same symbol $s_i(m)$ in N time intervals $[(nN_s + m)T, (nN_s + m + 1)T]$, $n \in \{0, \ldots, N-1\}$, the total transmitted energy for symbol $s_i(m)$ is:

$$\varepsilon_i = \sum_{n=0}^{N-1} \int_{t=(nN_s+m)T}^{(nN_s+m+1)T} P_i \partial t = NTP_0 \tag{3.3}$$

Thus in the rotational directional case $N > 1$, the transmitted energy is N times greater than in the omnidirectional case, but only part of the transmitted energy will be collected at the destination, during the single time-slot where the destination is in the transmission beam.

Received Power and Energy

The effective aperture of the omnidirectional antenna at a destination is an area A_e and the associated angular aperture is $\Delta\theta$, whereas θ_j represents D_j angular position in polar coordinates in the plane as in figure 3.3. We assume that the effective aperture A_e is small with respect to distances between nodes, so that the variations of the angular aperture with the position of the node can be neglected.

The received power $P_{ji}(t)$ at D_j coming from S_i depends on time, since the destination needs to be in the rotating beam to receive power from S_i.

$$
\begin{aligned}
P_{ji}(t) &= \frac{|h_{ji}|^2}{d_{ji}^2} \int_{\theta=\theta_j-\frac{\Delta\theta}{2}}^{\theta_j+\frac{\Delta\theta}{2}} \frac{\partial P_i}{\partial\theta}(\theta,t)\partial\theta \\
&= \frac{|h_{ji}|^2 P_0}{d_{ji}^2\alpha} \int_{\theta=\theta_j-\frac{\Delta\theta}{2}}^{\theta_j+\frac{\Delta\theta}{2}} \mathbb{1}_{[\alpha_i(t)-\alpha/2;\alpha_i(t)+\alpha/2[}(\theta)\partial\theta \\
&= \begin{cases} \frac{|h_{ji}|^2}{d_{ji}^2}\frac{P_0\Delta\theta}{\alpha} & \text{if } \theta_j \text{ is in } S_i \text{ beam at time } t \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{3.4}
$$

where $\frac{\Delta\theta}{\alpha}$ represents the fraction of power that the destination receives from the beam of width α, due to the finite size of the receive antenna.

S_i transmits a total energy ε_i for symbol $s_i(m)$ during the N time intervals $[(nN_s+m)T,(nN_s+m+1)T]$, $n \in \{0,\ldots,N-1\}$, but D_j receives energy ε_{ji} for symbol $s_i(m)$ only during the time interval of duration T when S_i beamforms in the direction of D_j, leading to the expression:

$$
\varepsilon_{ji} = \sum_{n=0}^{N-1} \int_{t=(nN_s+m)T}^{(nN_s+m+1)T} P_{ji}(t)\partial t = \frac{P_0|h_{ji}|^2\Delta\theta}{d_{ji}^2\alpha}T = N\varepsilon_{ji}^{omni} \tag{3.5}
$$

where $\varepsilon_{ji}^{omni} = \frac{P_0|h_{ji}|^2\Delta\theta}{d_{ji}^2 * 2\pi}T$ is the energy received by D_j for a symbol $s_i(m)$ transmitted during only once in the omnidirectional case. When $N > 1$ the received energy ε_{ji} for symbol $s_i(m)$ at D_j is N times greater than in the omnidirectional case because of the spatial focusing effect of the directional transmit antenna. Without a loss of generality, we will consider that $T = 1$ and simplify expressions in the sequel.

3.2.3 Performance Analysis

In this section, we derive the performance criteria to compare the dynamic blind beamforming strategy to the omnidirectional transmission in terms of total network throughput and throughput-based connectivity.

Total Network Throughput

For a communicating pair S_iD_i, the mutual information [55] between input s_i and output y_i at D_i, according to (3.5), is given by:

$$\begin{aligned} I(s_i;y_i) &= \log\left(1+\frac{\varepsilon_{ii}}{\sigma^2+\sum_{j\in\mathcal{I}_i}\varepsilon_{ij}}\right) \\ &= \log\left(1+\frac{\rho N\frac{|h_{ii}|^2\Delta\theta}{2\pi d_{ii}^2}}{1+\rho\sum_{j\in\mathcal{I}_i}N\frac{|h_{ij}|^2\Delta\theta}{2\pi d_{ij}^2}}\right) \end{aligned} \tag{3.6}$$

where the input SNR is $\rho = P_0/\sigma^2$. Since the source hits its intended destination only once in N successive rotational trials, the throughput of user i is given by

$$C_i = \frac{1}{N}I(s_i;y_i) \tag{3.7}$$

where the factor $1/N$ in front of the log accounts for the waste of time in the transmission of a symbol.

The total network throughput is given by:

$$C = \frac{1}{N}\sum_{i=1}^{M}\log\left(1+\frac{\rho N\frac{|h_{ii}|^2\Delta\theta}{2\pi d_{ii}^2}}{1+\rho N\sum_{j\in\mathcal{I}_i}\frac{|h_{ij}|^2\Delta\theta}{2\pi d_{ij}^2}}\right) \tag{3.8}$$

As previously mentioned, the use of rotating directional antennas allows to decrease the number of interferers in a group $\mathcal{I}_i$ and to focus the power in a direction, increasing the received power at the destination. But the spatial focusing also makes an interferer hit a non-intended destination stronger than in the omnidirectional case. The greater N, the narrower the beam thus the smaller the number of interferers and the higher the useful received power, but also the stronger the power of interference and the greater waste of time, suggesting a trade-off. The positive impact of the dynamic blind beamforming on the network throughput might not look obvious a priori, but it is shown in section 5.6.

Throughput-based Connectivity

Several definitions of connectivity exist, they have in common that two nodes are said to be connected if some criterion is above a threshold. In [54], connectivity is defined with respect to the level of received power, but this definition

does not take into account interference. To take into account interference, an SINR-based definition of connectivity can be considered. Nevertheless in the case of the dynamic blind beamforming technic we propose, defining the connectivity in terms of SINR above a threshold would lead to ignore the waste of time represented by the factor $1/N$ in the throughput formula. Indeed, it would be as if a pair was said to be always connected with a certain SINR, when the pair is actually discontinuously connected, only once every N time-slots.

The Information-theoretic point of view of connectivity, considering rate as the criterion to define connectivity, appears to be a more relevant and appropriate definition of connectivity. In particular, the notion of rate threshold makes sense in a quality-of-service approach, where users have target rates that need to be satisfied whatever happens. Inspired by [56] and taking into account the factor $1/N$, we define connectivity with target throughput R as follows: "A pair is connected if the source can communicate with its intended destination with a throughput at least R".

The network throughput-based connectivity κ is defined as the number of connected pairs divided by the number of pairs in the network, i.e. the proportion of pairs to which a throughput R can be guaranteed:

$$\kappa = \frac{|\{i/I(s_i; y_i) \geq R\}|}{M} \tag{3.9}$$

3.2.4 Numerical Results

In this section, numerical results are presented to compare the different transmission strategies. Monte-Carlo Simulations of $10,000$ different topologies were performed for different values of input SNR ρ, load M/N, number of pairs of nodes M in the network i.e. density. The edge of the area was $a = 100$ meters and the case of symmetric networks, i.e. in which the fading variances are identical $\sigma_{ji}^2 = 1$, was considered.

Total Network Throughput

We first analyze how the rotational directivity impacts the network throughput. Figures 3.4(b), 3.4(a) and (3.5) illustrate the total network throughput obtained by averaging the throughput over all generated topologies.

Figure 3.4(a) shows the evolution of the network throughput when the load increases, for different values of N. The network throughput reaches a

maximum corresponding to an optimal load, then saturates when the load of the network increases. Nevertheless a high number of rotations N allows to support a higher network throughput for a given load and to reach the saturation level later when the load increases.

Figure 3.4(b) plots the throughput versus the density of the network, for different values of N at high SNR. A similar behavior as in 3.4(a) - maximum then saturation - is observed, but the successive intersections of the curves show that the number of rotations maximizing the network throughput increases progressively when the density increases. Indeed at low densities, interference occurring in the network is low and the impact of dynamic blind beamforming on interference is not high enough to compensate the $1/N$ factor in front of the log in expression (3.8). On the contrary, at higher densities, the network becomes interference limited, omnidirectional transmission is not optimal anymore and the improvements in SINR via reduction of the interference thanks to dynamic blind beamforming are important enough to mitigate the $1/N$ loss. We would like to point out that although we present graphs at high SNR only, for the sake of conciseness, the same behavior is observed in the case of lower input SNR, except that intersections occur at higher densities.

The gains in total network throughput thanks to dynamic blind beamforming for increasing densities are clearly illustrated in figure 3.5, plotting the network throughput versus the number of rotations N, each curve representing a density δ. The curves at the bottom represent low densities, and the curves move toward the top of the graph when density increases. Clearly there exists an optimal N which maximizes the network throughput for each density, illustrating the trade-off between interference reduction and delay. Using the optimal N allows to dramatically improve the network sum-rate, from 30% at $M = 60$ pairs, up to 70% at very high densities ($M = 350$) with respect to omnidirectional transmissions. As the density of the network increases, the optimal N increases indicating that beams need to get narrower, but not too quickly so that improvements in terms of interference are not done at the expense of an infinite delay.

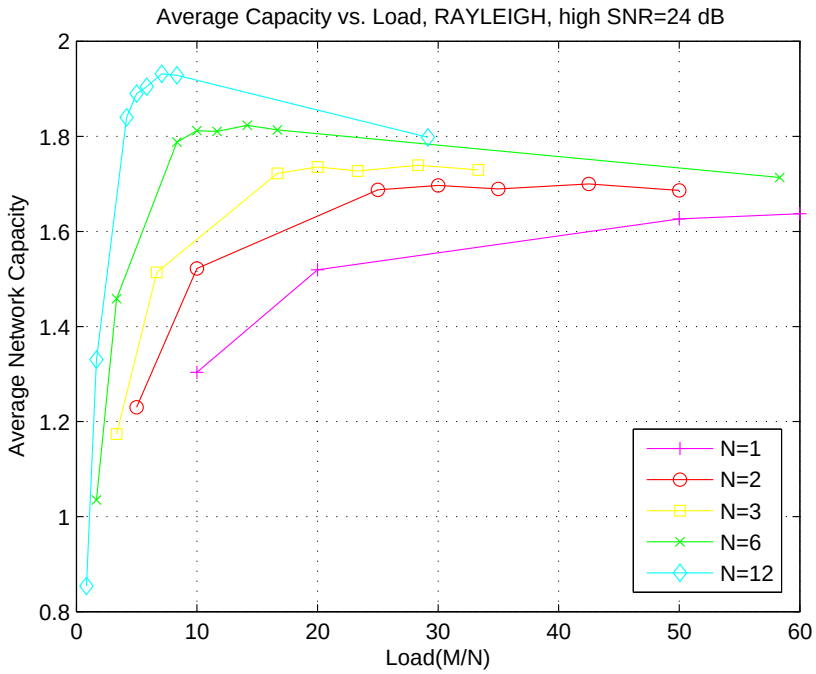
In large networks, omnidirectional transmissions are not optimal, and the use of directional antennas even blindly and dynamically allows to enhance the network performance.

Network Connectivity

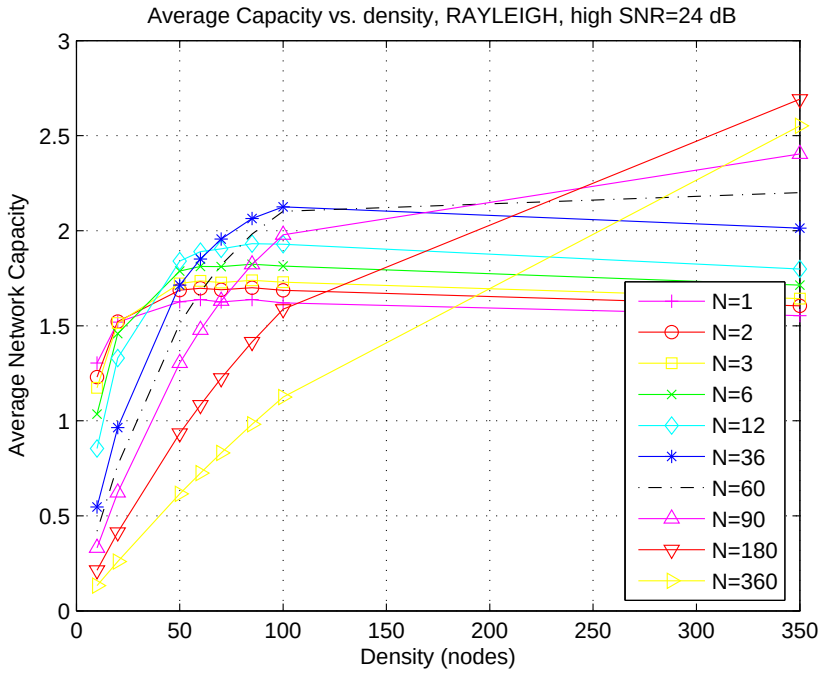
In terms of network connectivity, a trade-off is illustrated by figure 3.6, which plots the average connectivity versus the density for different N. Indeed curves can be grouped in two sets: for $2 \leq N \leq 12$ connectivity curves are above the omnidirectional case $N = 1$ and for $N > 36$ curves are below. When N increases, connectivity is increased up to a certain point, then increasing N decreases connectivity. Network connectivity can thus be improved thanks to dynamic blind beamforming.

3.2.5 Conclusion

We proposed a dynamic blind beamforming technique that allows to benefit from directional antennas while avoiding the feedback load for localization tracking in ad hoc networks with a large number of nodes. We analyzed performance in terms of total network throughput and throughput-connectivity and showed that our scheme can outperform omnidirectional transmissions in ad hoc networks. Depending on the density of the network, an optimal number of rotations allows to maximize the network performance. This optimal number of rotations results from a trade-off between introduction of delay and reduction of interference. In large ad hoc networks, which are known to be interference limited, omnidirectional transmissions are not optimal and the use of directional antennas even blindly and dynamically allows to fight against interference and to enhance the network performance. Future work may include analysis of the impact of limited feedback of the positions on the performance of dynamic blind beamforming in particular in a high mobility environment.



(a) Average Network Capacity versus Load



(b) Average Network Capacity versus Density

Figure 3.4: Comparison of Network Capacities For Different Rotational Scenarios

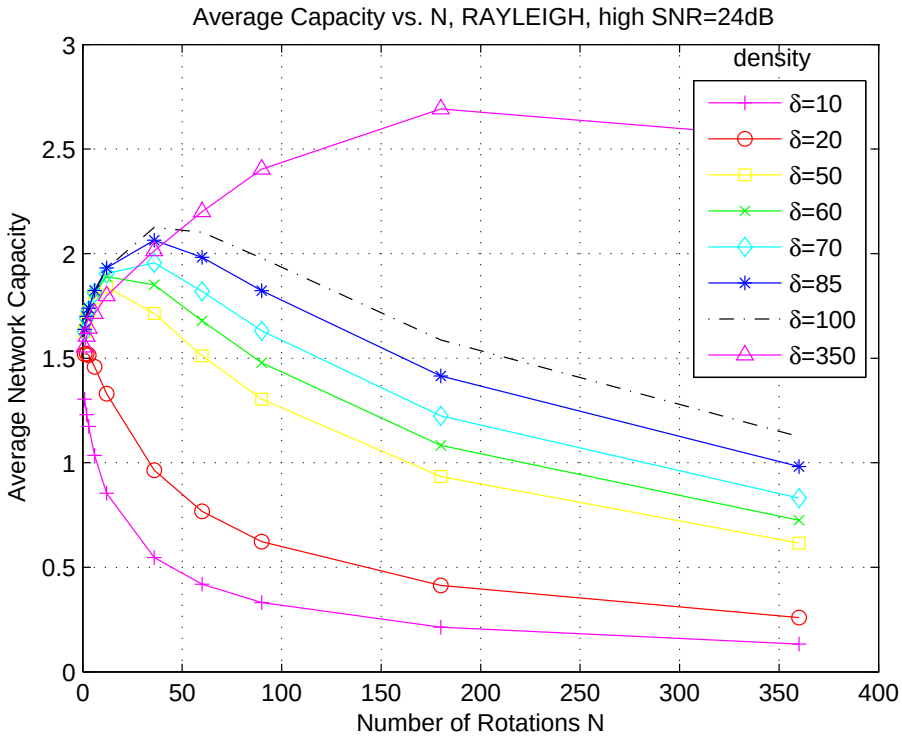


Figure 3.5: Average Network Capacity versus N

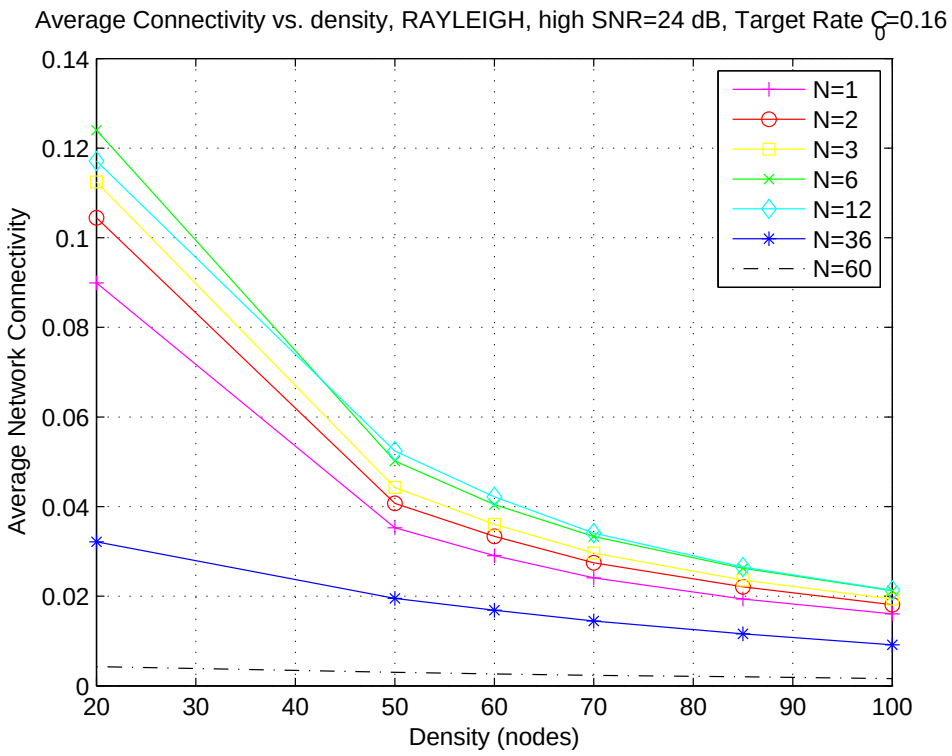


Figure 3.6: Average Network Connectivity versus Density

3.3 Node Positioning Impact

3.3.1 Introduction

In this section, we study the impact of the number and the position of nodes on the capacity of a wireless network.

Contribution

We analyze a system where a dense network of scatterers helps a source to communicate with a destination and we look at the network from a physical propagation point of view: relays are modeled as dumb omnidirectional scatterers, i.e. passive nodes without engineering capabilities that simply scatter the incident electromagnetic wave coming from the source antenna. Capacity expressions accounting for physical characteristics of the environment (topology, frequency band...) are provided and an asymptotic analysis is performed for an increasing number of scattering nodes. We study how capacity scales when the number of scattering nodes increases, as well as the point at which asymptotic regime is reached depending on the nodes positioning. The capacity is shown to reach a saturation level in the asymptotic regime. This saturation is due to the fact that signal contributions coming from peripheral nodes very far from source and destination do not lead to much increased performance. Moreover the saturation point depends on the positioning of scatterers, in particular in wideband systems where topology impacts capacity in terms of pathloss, delay and multipath. Waiting very long for retransmissions from an infinite number of scatterers is not worth and a few well located scatterers around source and destination lead to better performances than more scatterers uniformly distributed on a square area between source and destination. However saturation is obtained under the assumption that relaying nodes are passive and do not interact when transmitting their signals, and the capacity scaling may be different when considering active relays with virtual MIMO capabilities.

Related Works

In [57], a source and destination both equipped with M antennas communicate with the help of K relays performing amplify and forward. When the number of relays K grows to infinity, the capacity of the system scales as $C = (M/2)\log K + O(1)$ when CSI is available at relays, while it scales as

$C = (M/2)\log \mathrm{SNR} + O(1)$ in the non-coherent setting, i.e. when no CSI is available at relays. Interestingly, capacity appears to reach a saturation level in the non-coherent case as with the model considered in this chapter. Nevertheless, to the best of our knowledge no previous work focused on the scaling of capacity in dense wideband – thus with high resolution in time and space – networks of dumb scatterers taking into account topology not only in terms of pathloss but also of multi-path. Moreover the environment impact on the scaling laws, through reflections, diffraction effects... has never been studied in details, although this aspect is used for example in MIMO communications to create different spatial multiplexing streams, and we went deeper in the analysis in [18, 58].

3.3.2 System Model

We focus on a geometrical network model by considering a finite square grid of scatterers, equally spaced vertically and horizontally. Source S and Destination D are located on vertices of the grid separated by a distance d_0.

Scatterers Clustering

In a wideband system, resolution in time and space are high and propagation delays cannot be neglected as in narrow band systems. Consider a band-limited input signal with band W and carrier frequency f_c, and let us model it as a complex baseband process of band $W/2$. Using the sampling theorem, the input signal can be represented as a sampled complex time process, with samples at sampling rate $W = 1/T$. We transmit the input signal in a fading channel. At the output of the channel, given sampling rate W, differential propagation delays larger than $T = 1/W$ seconds can be discriminated, which corresponds to differential propagation distances larger than $\lambda_s = c/W$ meters. In a wideband system, λ_s may be small compared to the distance between communicating terminals and reflectors, therefore reflections leading to differential propagation distances larger than λ_s will appear at the destination as multiple paths arriving in different time-slots. To take into account propagation delays, scatterers are grouped in N_{clus} clusters, depending on the time-slot of reception of the scattered wave. Thus a cluster gathers scatterers whose scattered waves are received by D during the same interval T.

Considerer a square grid of scatterers spaced by $\lambda_s/2 = c/(2W)$ vertically

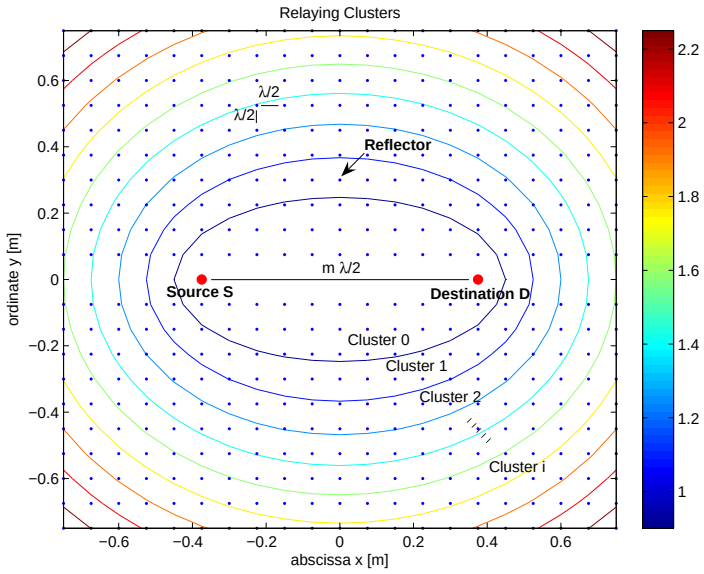


Figure 3.7: Relaying Clusters

and horizontally, where W is the sampling rate, and Source S and Destination D located on vertices of the grid separated by a distance $d_0 = m\lambda_s/2$. The shortest distance to go from source to destination is the straight path without reflection of length d_0 and propagation duration $\tau_0 = d_0/c = (m/2)T$. Cluster Δ_0 contains the straight path from source to destination, as well as the reflected paths received during the first time-slot $TS_0 = [\tau_0, \tau_0 + T[$. Those paths correspond to waves propagating on a total distance, sum of the distances from S to the scatterer and from the scatterer to D between d_0 and d_0+T: $d_r^{(s)}+d_r^{(d)} \in [d_0, d_0+\lambda_s[$. For $i \geq 1$, Cluster Δ_i contains only scattered waves, such that $(d_r^{(s)} + d_r^{(d)}) \in [d_0 + i\lambda_s, d_0 + (i+1)\lambda_s[$. Those waves are received during the i^{th} time-slot defined by $TS_i = [\tau_0 + iT, \tau_0 + (i+1)T[$. For a scatterer k in cluster i, the propagation delay of the scattered wave from S to D via this scatterer $R_{k,i}$ will be denoted $\tau_{k,i} \in TS_i$. This duration consists of three parts: $\tau_{k,i} = \tau_0 + iT + \tau'_{k,i}$ with $\tau'_{k,i} \in [0, T[$. Note that the sampling rate being $W = 1/T$, the destination cannot discriminate between the delays $\tau'_{k,i} \in [0, T[$ of scatterers in the same cluster, thus the destination sees all reflected paths from scatterers in the same cluster as a single combined path.

Fig. (3.7) illustrates the clusters in the case of a 13×13 grid and a source-destination distance $d_0 = 4\,\lambda_s/2$. A cluster is geometrically represented by a surface bounded by two ellipses whose focus are S and D, and whose equations are given by $d_r^{(s)} + d_r^{(d)} = d_0 + i\lambda_s$ and $d_r^{(s)} + d_r^{(d)} = d_0 + (i+1)\lambda_s$.

Power Attenuation

To model the attenuation of propagating waves, the far field propagation model holding for distances $d \gg \frac{\lambda_c}{2\,\pi}$ will be considered, where λ_c is the carrier wavelength. Since $f_c \geq W/2$, the inequality $\lambda_s \geq \lambda_c/2$ holds and the far field condition is fulfilled for distances $d \geq \lambda_s/2$. The power received by D depends on the path taken by the signal to reach D: one-hop direct path from S to D without reflection or two-hop-path from S to D via one scatterer. Multiple reflections before reaching D are not taken into account.

If no reflection occurs, the direct received power [59,60] is

$$P_d = \frac{\lambda_c^2}{(4\pi)^2\, d_0^2}\, P_{rad} = K_1^2 \frac{\lambda_c^2}{d_0^2}\, P_{rad} \tag{3.10}$$

with P_{rad} the power transmitted by the source antenna and the constant $K_1 = 1/(4\pi)$.

If one reflection occurs before reaching D, according to the radar equation [59,60] with omnidirectional antennas, the reflected power received by D is

$$P_r = \frac{\lambda_c^2\, s}{(4\pi)^3\, (d_r^{(s)}\; d_r^{(d)})^2}\, P_{rad} = K_2^2\, \frac{\lambda_c^2}{(d_r^{(s)}\; d_r^{(d)})^2}\, P_{rad} \tag{3.11}$$

where s is the radar cross section, and $K_2 = \sqrt{s/4\pi} K_1 = \sqrt{s/(4\pi)^3}$. Scatterers can be seen as very minimalist and dumb wireless nodes with a very restrictive power constraint. They are minimalist because they do not have any engineering capabilities, they can only scatter a wave; and their power constraint is restrictive since the only power available to them for transmission is the power they collected from the source wirelessly.

3.3.3 Analysis

Signal Expressions in Time Domain

Source S produces a sequence $s(t) = \sum_{n=0}^{N-1} s_n\, \delta(t - nT)$ of N complex symbols with power $E[|s_n|^2] = P_{rad}$ at sampling rate $W = 1/T$ (thus an average

energy $\varepsilon = P_{rad}/W$ per symbol) and transmits a linearly modulated signal with complex envelope $x(t) = s(t) * g(t) = \sum_{n=0}^{N-1} s_n\, g(t - nT)$, where $g(t)$ is the pulse shaping filter, satisfying Nyquist's criterion $\int g(t)g^*(t - kT)dt = \delta_{0,k}$. We assume hereafter that $g(t)$ is the rectangle function of amplitude $1/\sqrt{T}$ over $[0, T[$. The received signal at D is the superposition of the signal coming directly from S and the signals scattered once:

$$y(t) \;=\; \sum_{i=0}^{N_{clus}-1} \sum_{k\in\Delta_i} a_{k,i}\, x(t - \tau_{k,i}) \;+\; n'(t) \tag{3.12}$$

where n'(t) is additive white Gaussian noise and coefficients $a_{k,i}$ are given by (3.10) and (3.11):

$$a_{k,i} = \begin{cases} K_1 \frac{\lambda_c}{d_0} e^{j\varphi_{0,0}} & \text{for } (k,i) = (0,0) \quad \text{direct path} \\ K_2 \frac{\lambda_c}{d_{k,i}^{(s)}\, d_{k,i}^{(d)}} e^{j\varphi_{k,i}} & \text{for } (k,i) \neq (0,0) \quad \text{reflected path} \end{cases}$$

$\varphi_{k,i} \in [0, 2\pi[$ are phase shifts due to propagation and reflections. Phases can be modeled as independent random variables provided nodes are sufficiently spaced. If the network becomes denser and denser, phases may not be independent anymore but correlated and should be expressed in function of optical path differences, which is out of the scope of this section. After matched-filtering, the received signal becomes:

$$\begin{aligned} r(t) \;&=\; y(t) * g^*(-t) \;+\; n(t) \\ &=\; s(t) * \underbrace{g(t) * g^*(-t) * \sum_{i=0}^{N_{clus}-1} \sum_{k\in\Delta_i} a_{k,i}\, \delta(t - \tau_{k,i})}_{h(t-\tau_0)} + n(t) \\ r(t) \;&=\; h(t) * s(t - \tau_0) \;+\; n(t) \end{aligned} \tag{3.13}$$

where we define the scatterers network equivalent channel by:

$$h(t) \;=\; \sum_{i=0}^{N_{clus}-1} \sum_{k\in\Delta_i} a_{k,i}\, g(t - \tau_{k,i} + \tau_0) * g^*(-t) \tag{3.14}$$

Time shift τ_0 in equation (3.13) illustrates the minimum propagation delay corresponding to direct path. D starts receiving signals only τ_0 seconds after

S started emitting. The introduction of τ_0 in the definition of h simplifies notations since all the delays $\tau_{k,i} = \tau_0 + iT + \tau'_{k,i}$ contain τ_0.

By sampling at rate $W = 1/T$, the received sequence is the convolution between the symbol sequence and the channel impulse response (CIR):

$$\begin{aligned} r_l &= r(lT) = \sum_{n=0}^{N-1} s_n h(lT - nT - \tau_0) + n(lT) && (3.15)\\ &= \sum_{n=0}^{N-1} s_n h_{l-n-m/2} + n_l && (3.16)\\ &= \sum_{n=0}^{N_{clus}} h_n \, s_{l-n-m/2} + n_l \end{aligned}$$

The sum in (3.15) is finite and contains only $N_{clus} + 1$ non-null terms. Indeed, as shown in Appendix 3.A, the CIR has finite length $N_{clus} + 1$ and its coefficients are given by:

$$\begin{aligned} h_0 &= \sum_{k \in \Delta_0} a_{k,0} \left(1 - \frac{\tau'_{k,0}}{T}\right) && (3.17)\\ h_l &= \sum_{k \in \Delta_l} a_{k,l} \left(1 - \frac{\tau'_{k,l}}{T}\right) + \sum_{k \in \Delta_{l-1}} a_{k,l-1} \left(\frac{\tau'_{k,l-1}}{T}\right) \text{ , for } l \in 1, \ldots, N_{clus} - 1\\ h_{N_{clus}} &= \sum_{k \in \Delta_{N_{clus}-1}} a_{k,N_{clus}-1} \left(\frac{\tau'_{k,N_{clus}-1}}{T}\right) \end{aligned}$$

This means that the scatterer network is equivalent to a multi-path fading channel, where each path h_l is the combination of contributions from scatterers in clusters l and $l-1$. Note that the $m/2$ first samples $r_0 = ...r_{m/2-1} = 0$ are null because of the minimum propagation delay $\tau_0 = mT/2$. Under matrix notation, the system can be reduced to the equation

$$\mathbf{R}_N = \mathbf{H}_N \, \mathbf{S}_N \; + \; \mathbf{N}_N \qquad (3.18)$$

where $\mathbf{R}_N = [r_{m/2}, ..., r_{N+m/2-1}]^T$, $\mathbf{S}_N = [s_0, ..., s_{N-1}]^T$, and $\mathbf{N}_N = [n_0, ..., n_{N-1}]^T$ are columns of size N and the channel is represented by the $N \times N$ lower

triangular banded Toeplitz matrix:

$$\mathbf{H}_N = \begin{bmatrix} h_0 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ h_{N_{clus}} & & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & h_{N_{clus}} & \dots & h_0 \end{bmatrix}_{N\times N} \tag{3.19}$$

Asymptotic analysis for Capacity expression

The capacity of the grid of scatterers is defined [55] as:

$$C = \lim_{N\to+\infty} \frac{W}{N} \log_2 \det\left(\mathbf{I}_N + \rho \mathbf{H}_N \mathbf{H}_N^H\right) \tag{3.20}$$

where $\rho = \frac{P_{rad}}{W\sigma^2} = \frac{\varepsilon}{\sigma^2}$ is the SNR. Using Toeplitz and circulant matrices properties, we show in appendix 3.B the following proposition:

Proposition 1. *Let $\{\lambda_{k,N}\}_{k\in[0,N-1]}$ denote the N-point DFT (Discrete Fourier Transform) of the first column of matrix $\mathbf{H}_N$. Then the capacity of the grid of scatterers is given by*

$$C = \lim_{N\to+\infty} \frac{W}{N} \sum_{k=0}^{N-1} \log_2(1+\rho|\lambda_{k,N}|^2) \tag{3.21}$$

Expressing the average capacity in terms of the network physical characteristics is intricate. Nevertheless Jensen's inequality allows to give an upperbound (3.22) to the average capacity $\bar{C} = \mathbb{E}_\varphi[C]$ in function of physical characteristics of the network

$$\frac{\bar{C}}{W} \leq \log\left(1+\rho\left(K_1^2\frac{\lambda_c^2}{d_0^2} + K_2^2\frac{\lambda_c^2}{\lambda_s^2}\sum_{i=0}^{N_{clus}-1}\sum_{k\in\Delta_i}\frac{((i+1)\lambda_s + d_0 - d_{k,i}^{(s)} - d_{k,i}^{(d)})^2 + (i\lambda_s + d_0 - d_{k,i}^{(s)} - d_{k,i}^{(d)})^2}{(d_{k,i}^{(s)} d_{k,i}^{(d)})^2}\right)\right) \tag{3.22}$$

Besides at low SNR $\rho = \frac{\varepsilon}{\sigma^2}$, using Taylor expansion of the log around zero at order 1 and performing some manipulations of (3.20), that we skip for sake of conciseness and readability, $\bar{C}$ can be written as

$$\frac{\bar{C}}{W} \approx \frac{\rho}{\ln 2}\left(K_1^2\frac{\lambda_c^2}{d_0^2} + K_2^2\frac{\lambda_c^2}{\lambda_s^2}\sum_{i=0}^{N_{clus}-1}\sum_{k\in\Delta_i}\frac{((i+1)\lambda_s + d_0 - d_{k,i}^{(s)} - d_{k,i}^{(d)})^2 + (i\lambda_s + d_0 - d_{k,i}^{(s)} - d_{k,i}^{(d)})^2}{(d_{k,i}^{(s)}d_{k,i}^{(d)})^2}\right) \tag{3.23}$$

Those expressions illustrate the SNR-gain due to scatterers in function of the topology and the system band. We would like to point out that those formulas are valid for any topology and are not specific to the grid model, nor to the $\lambda_s/2$-node spacing.

3.3.4 Simulations and Results

In this section, numerical results illustrate the growth of average capacity when the number of scatterers increases. We consider a wide band system with $W = 2GHz$ and a carrier $f_c = 2GHz$, at SNR $\rho = 10dB$, for different values of distance d_0. Scatterers are located on the vertices of a square grid covering an area of $21\lambda_s/2 \times 21\lambda_s/2$. Nevertheless we consider that the line (SD) does not contain any scatterer, since when three nodes are aligned, one link among the links S-R, S-D, R-D is blocked.

The number of scatterers is increased in two different ways:

- **Centered-Grid Positioning** : N_r scatterers are uniformly distributed on the vertices of a grid of size $\sqrt{N_r} \times \sqrt{N_r}$, centered on the midpoint I between S and D. Increasing N_r corresponds to increasing the edge of the square grid. The average capacity (3.21) obtained by Monte-Carlo simulations over many independent channel realizations as well as Jensen's upperbound (3.22) are plotted in Fig. (3.8) for the centered grid positioning.

- **Optimal Positioning** : Considering a grid 21×21, we select the N_r optimal vertices , i.e. the positions that give the highest capacity for a given number N_r of scatterers. Increasing N_r corresponds to adding a scatterer at the available vertex which gives the next highest increase of capacity. Jensen's upperbound (3.22) only is plotted in this case in Fig. (3.8).

Fig. (3.8) shows that the passive relaying achieved by scatterers allows to increase the capacity with respect to the case without scatterers, and that capacity saturation occurs as the number of scatterers increases. The level

of saturation corresponds to a 30%-increase in capacity at $d_0 = 3\lambda_s$ and 40%-increase at $d_0 = 5\lambda_s$. Saturation is due to the following reasons:

- scatterers act as power collectors and forwarders. Indeed, the destination has a finite size antenna and does not receive all the power that was radiated omni-directionally by the source antenna. Scatterers help the destination by collecting part of the energy transmitted by the source that would be lost otherwise, and radiate it again.

- reflection achieved by scatterers is diffuse, not specular. Consequently the power collected by a scatterer is reflected in broad range of directions, which is the reason why the power received at destination after a reflection on a scatterer R decreases as $\frac{1}{(d_r^s d_d^r)^2}$ instead of only $\frac{1}{(d_d^s)^2}$ for the direct path.

- the destination, even helped by scatterers, cannot collect more power than transmitted by the source, which represents a physical constant upper-bound on the capacity. Given the diffuse-nature of reflection at scatterers and the consequent power attenuation as $\frac{1}{(d_r^s d_d^r)^2}$, the destination will never be able to recover all the power transmitted by the source. Thus the actual saturation level of the capacity is lower than the capacity that would be obtained if the destination could collect all the source power.

- When the number of scatterers increases, the number of scattering clusters increases and consequently, the length of the channel impulse response, as well as its delay-spread, increase. Nevertheless, the paths with the largest delays are due to scatterers located at an increasing distance from source and destination, and thus with a strong pathloss attenuation. Signals coming from peripheral nodes very far from source and destination lead to small contributions. After a certain point, most of the power that could be recovered at the destination thanks to scatterers contributions has already been collected— thus a saturation, and the increase in capacity resulting from late paths due to retransmissions from very far nodes is negligible.

Note that allowing a little cooperation between nodes within a cluster l to combine coherently their scattered signals would obviously lead to a gain in capacity, since their scattered powers would add coherently in the path h_l

and thus increase the path gain. The channel between such a cooperative cluster and the destination would be similar to a MISO channel and the coherent combining would provide transmit diversity at the destination.

Fig. 3.9 shows how the contribution to SNR of each scatterer depends on its position. Scatterers positions affect the capacity not only in terms of path loss but also of delay, leading to a notable difference between performances in the centered-grid case and the optimal positioning case. A few scatterers well located, close to source or destination according to fig. (3.9), lead to better performances than a large number of scatterers uniformly distributed between source and destination.

3.3.5 Conclusion

In this contribution, the link capacity in a scattering network is analyzed from a physical point of view, taking into account characteristics such as topology and transmission band. Asymptotic analysis shows that capacity saturates when the size of the network increases and that topology affects the saturation point, in particular in wide band systems where the impact of topology on capacity is not only a matter of pathloss but also of delays that cannot be neglected. Capacity saturation suggests that it is not worth waiting for infinite retransmissions and that a few well located scatterers around source and destination lead to better performance than more scatterers uniformly distributed on a square area centered between source and destination.

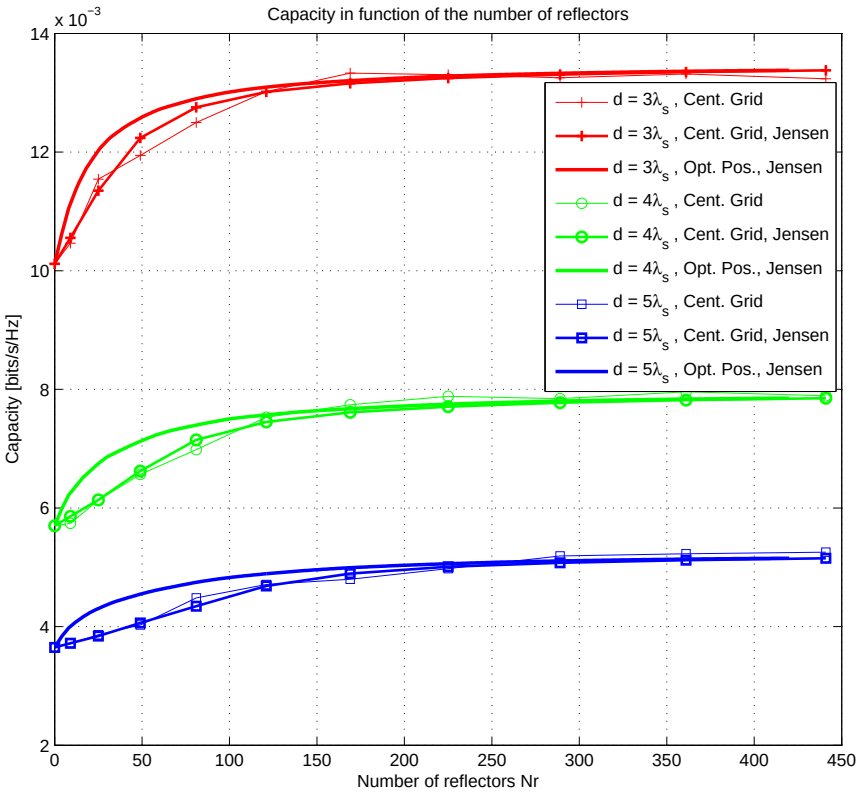


Figure 3.8: Capacity in function of the number of scatterers

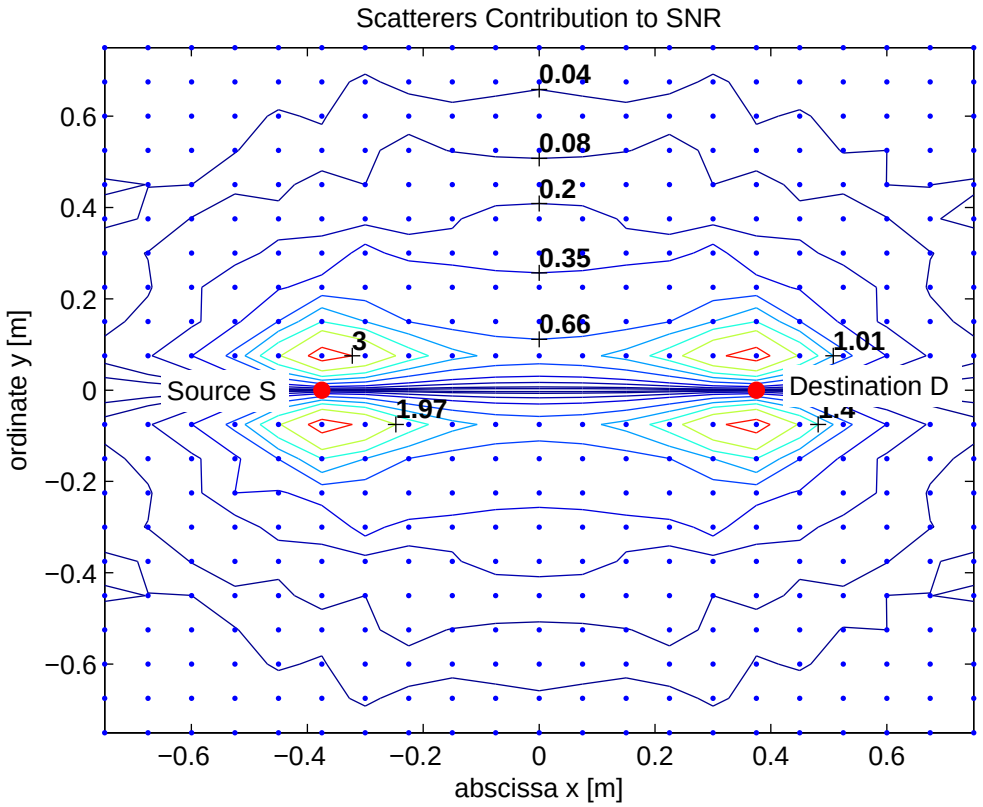


Figure 3.9: Contribution to SNR of scatterers vs. their position

3.4 Conclusion

In this Chapter, we presented physical layer factors which can improve the link capacity performance in wireless ad hoc networks. We showed that in a dense network of source-destination pairs, the use of directional antennas even blindly and dynamically allows to fight against interference and to enhance the performance in terms of throughput and throughput-connectivity. Considering a communication between a source and a destination with the help of a large number of passive relays, we studied the impact of the number and position of relays on the capacity. It turned out that increasing the number of passive relaying nodes does not increase the system capacity indefinitely, but that saturation occurs and that a few well located relays around source and destination lead to better performance than a larger number of relaying nodes uniformly distributed.

However, so far, nodes were considered as independent units, which did not cooperate to jointly transmit information. In next chapters, we turn to an active cooperation mode, by allowing wireless nodes to interact. In Chapter 4, we consider cooperative systems with small dimensions, as building blocks for larger cooperative systems, and analyze how cooperative protocols can be designed to improve the network performance while making an efficient use of the available resource. In Chapter 5, we consider networks with larger dimensions and see how virtual MIMO affects the capacity when the dimensions of the network grow large.

APPENDIX

3.A Proof of CIR expression

In this appendix, the proof of the CIR expression 3.17 is given.

The channel impulse response is defined by the coefficients:

$$h_l = h(lT) = \sum_{i=0}^{N_{clus}-1} \sum_{k \in \Delta_i} a_{k,i} \int g(\tau - \tau_{k,i} + \tau_0) g^*(\tau - lT) d\tau \tag{3.24}$$

By definition with a rectangle transmitting filter, the integral in (3.24) is non null for a finite set of values of l, more precisely $l \in [0, N_{clus}]$, as shown hereunder:

$$\begin{aligned} \int g(\tau - \tau_{k,i} + \tau_0)\, g^*(\tau - lT) d\tau &= \int g(\tau)\, g^*(\tau + \tau'_{k,i} - (l-i)T) d\tau \\ &= f^{(1)}_{k,l}\, \delta_{i-l} + f^{(2)}_{k,l-1}\, \delta_{i-l-1} \end{aligned} \tag{3.25}$$

with $\tau'_{k,i} \in [0, T[$, and $f^{(1)}$ and $f^{(2)}$ are defined by:

$$\text{for } i \in [0, N_{clus}-1] \quad \text{and} \quad k \in \Delta_i,$$

$$\begin{aligned} f^{(1)}_{k,i} &= \int g(\tau)\, g^*(\tau + \tau'_{k,i}) d\tau \\ &= 1 - \tau'_{k,i}/T \end{aligned} \tag{3.26}$$

$$\begin{aligned} f^{(2)}_{k,i} &= \int g(\tau)\, g^*(\tau - T + \tau'_{k,i}) d\tau \\ &= \tau'_{k,i}/T \end{aligned} \tag{3.27}$$

Then combining (3.24) and (3.25) leads to the coefficients:

$$h_0 = h^{(1)}_0 = \sum_{k \in \Delta_0} a_{k,0}\, f^{(1)}_{k,0} \tag{3.28}$$

$$\begin{aligned} h_l &= h^{(1)}_l + h^{(2)}_{l-1} \quad \text{, for } l \in [1, N_{clus}-1] \\ &= \sum_{k \in \Delta_l} a_{k,l}\, f^{(1)}_{k,l} + \sum_{k \in \Delta_{l-1}} a_{k,l-1}\, f^{(2)}_{k,l-1} \end{aligned}$$

$$h_{N_{clus}} = h^{(2)}_{N_{clus}-1} = \sum_{k \in \Delta_{N_{clus}-1}} a_{k,N_{clus}-1}\, f^{(2)}_{k,N_{clus}-1}$$

Indeed, h is the superposition of two FIR (Finite Impulse Response) $h^{(1)}$ and $h^{(2)}$, of length $N_{clus}-1$, shifted by T with respect to each other. Thus, h is of length N_{clus}. □

3.B Proof of Proposition 1

In this section, we provide the proof of the asymptotic capacity expression (3.21) in *Proposition 1*. $\{\mathbf{H}_N\}_{N\in\mathbb{N}}$ forms a sequence of banded Toeplitz matrices of order $N_{clus}+1$ (non null coefficients). We study their asymptotic behavior, i.e. for $N \gg N_{clus}$. As in [61] we define:

- the circulant matrix $\mathbf{G}_N$ associated to $\mathbf{H}_N$:

$$\mathbf{G}_N = \begin{bmatrix} h_0 & 0 & \dots & 0 & h_{N_{clus}} & \dots & h_1 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ h_{N_{clus}} & & \ddots & \ddots & & \ddots & h_{N_{clus}} \\ 0 & \ddots & & \ddots & \ddots & & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & h_{N_{clus}} & \dots & h_0 \end{bmatrix}_{N\times N}$$

- the sequence $\{\mathbf{A}_N\}_{N\in\mathbb{N}}$ of hermitian matrices $\mathbf{A}_N = \mathbf{H}_N\mathbf{H}_N^H$ with non negative eigenvalues sets $\{\alpha_{k,N}\}_{k\in[0,N-1]}$
- the sequence $\{\mathbf{B}_N\}_{N\in\mathbb{N}}$ of hermitian matrices $\mathbf{B}_N = \mathbf{G}_N\mathbf{G}_N^H$ with non negative eigenvalues sets $\{\beta_{k,N}\}_{k\in[0,N-1]}$

According to lemma 4.2 in [61], $\mathbf{H}_N$ and $\mathbf{G}_N$ are asymptotically equivalent, as well as $\mathbf{H}_N^H$ and $\mathbf{G}_N^H$. Thus, by theorem 2.1.(3) in [61], their products are asymptotically equivalent: $\mathbf{A}_N = \mathbf{H}_N\mathbf{H}_N^H \sim \mathbf{G}_N\mathbf{G}_N^H = \mathbf{B}_N$ and by theorem 2.1.(6), there are finite constant m and M such that $m \le \alpha_{k,N}, \beta_{k,N} \le M$. In particular $\mathbf{A}_N$ and $\mathbf{B}_N$ are nonnegative definite, so $0 \le \alpha_{k,N}, \beta_{k,N} \le M$.

The capacity (3.20) can be written:

$$C = \lim_{N\to+\infty} \frac{W}{N} \sum_{k=0}^{N-1} \log_2(1+\rho\alpha_{k,N}) \tag{3.29}$$

From (3.29) we define the function $F(u) = \log_2 (1 + \frac{\varepsilon}{W\sigma^2} u)$, continuous on $]-\frac{W\sigma^2}{\varepsilon}, +\infty[$ and thus on the interval $[0, M]$ bounding the eigenvalues $\alpha_{k,N}$ and $\beta_{k,N}$. $\{\mathbf{A}_N\}_{N\in\mathbb{N}}$ and $\{\mathbf{B}_N\}_{N\in\mathbb{N}}$ being asymptotically equivalent sequences of hermitian matrices, then *Theorem 4* in [61] allows to conclude that

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{k=0}^{N-1} F(\alpha_{k,N}) = \lim_{N\to+\infty} \frac{1}{N} \sum_{k=0}^{N-1} F(\beta_{k,N}) \tag{3.30}$$

and to rewrite the capacity (3.29) as

$$C = \lim_{N\to+\infty} \frac{W}{N} \sum_{k=0}^{N-1} \log_2(1 + \rho\beta_{k,N}) \tag{3.31}$$

Now $\mathbf{G}_N$ is a circulant matrix, thus diagonalizable in the Fourier basis, leading to the diagonal matrix $\mathbf{C}_N = \mathbf{F}_N \, \mathbf{G}_N \, \mathbf{F}_N^{-1} = diag(\lambda_{0,N}, \ldots, \lambda_{N-1,N})$ and the eigenvalues $\{\lambda_{k,N}\}_{k\in[0,N-1]}$ given by the DFT of the first column of $\mathbf{G}_N$:

$$[\lambda_{0,N}, ..., \lambda_{N-1,N}]^T = \mathbf{F}_N \, [h_0, ..., h_{N_{clus}-1}, 0, ..., 0]^T_{(N\times 1)}$$

where $\mathbf{F}_N = (e^{-\frac{j2\pi(n-1)(\nu-1)}{N}})_{n,\nu\in\{1\ldots N\}}$ is the N-point DFT matrix. Expressing $\mathbf{B}_N$ in function of $\mathbf{C}_N$ gives

$$\mathbf{B}_N = \mathbf{F}_N^{-1} \, \mathbf{C}_N \mathbf{C}_N^H \, \mathbf{F}_N = \mathbf{F}_N^{-1} diag(|\lambda_{0,N}|^2, \ldots, |\lambda_{N-1,N}|^2) \mathbf{F}_N$$

so that $\beta_{k,N} = |\lambda_{N,k}|^2$, which substituted in (3.31) leads (3.21). □

Chapter 4

Cooperation in Small Dimension Networks

4.1 Introduction

4.1.1 Motivation

In wireless communications, multi-path propagation leads to channel fading which can impair communications. Multiple-antenna systems can turn multi-path fading into a benefit for users, by sending a data stream over independent fading channels and recombining the multiple copies properly at the receiver. Indeed the probability of losing a signal decreases when the number of independent random fadings it experiences increases: this is know as spatial diversity. Nevertheless, achieving spatial diversity requires the multiple antennas to be uncorrelated, thus sufficiently spaced— in the order of half the wavelength. As an example, UMTS (Universal Mobile Telecommunications System) transmissions occur in the 2GHZ frequency band, which corresponds to a wavelength of 15 cm, and the antennas need to be spaced by at least 7.5 cm, which is hard to build in small mobile devices.

When a signal is transmitted wirelessly, the broadcast nature of the wireless link allows different wireless terminals to overhear the signal. A natural idea to exploit spatial diversity is to capitalize on the broadcast nature of the wireless link, by allowing several terminals to pool their antennas together to build a distributed multiple antenna systems. Cooperative communications occur when distributed wireless nodes interact to jointly transmit information. Several radio terminals relaying signals for each other form a virtual antenna array and their cooperation enables the exploitation of spatial diversity in fading channels, which is then called cooperative diversity.

The use of distributed antennas allows:

- to provide spatial diversity benefits without the need for physical arrays, though at a loss of spectral efficiency due to the practical half-duplex mode;

- to largely enhance performance in terms of decreased transmit power for a given reliability or increased reliability for a given transmit power.

Several cooperative strategies already exist [23, 24]. The simplest and most famous ones exploiting cooperative diversity are Amplify and Forward (AF), Decode and Forward (DF) and Compress and Forward (CF) [25]. These strategies differ by the processing that the relay performs on its received signal before forwarding it:

- Amplify and Forward (AF): the relay amplifies the noisy signal it received, i.e. multiplies its received signal by a constant subject to its power constraint. The main advantage of AF is the simplicity of the signal processing operated by the relay, while its main drawback lies in the amplification and retransmission of noise;

- Decode and Forward (DF): the relay detects and decodes the received signal according to a given algorithm and re-encodes the information into its transmitted signals. Several algorithms can be used: repetition coding (RDF), parallel channel coding (PDF), space-time coding (STC)...;

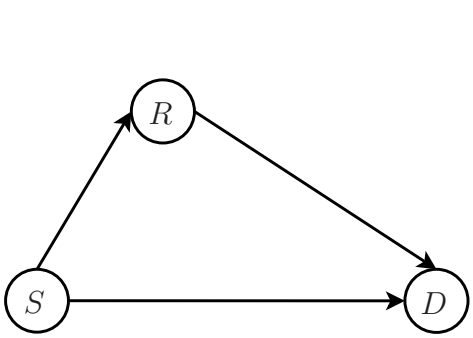
- Compress and Forward (CF): the relay forwards a compressed or quantized version of its received symbols.

The performance of the previous cooperative strategies can be further improved when coupled with the following techniques:

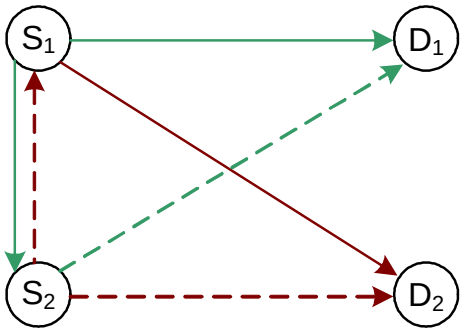
- Selection Relaying: the relay adapts its retransmission according to the realized value of the fading between source and relay. The relay forwards its received signal only if the source-relay channel is above a certain threshold. In the case of AF and CF, selection relaying allows to avoid forwarding too noisy symbols, while in the case of DF it allows to avoid forwarding incorrectly decoded symbols.

- Incremental Relaying: In classical ARQ the source retransmits if it receives a negative acknowledgment via feedback. Incremental relaying is an extension of ARQ to relay-based communications, where limited feedback from the destination is exploited: the relay forward its received signal only if the destination could not decode. By avoiding useless repetitions, incremental relaying allows the degrees of freedom of the channel to be used more efficiently.

Since radio terminals cannot transmit and receive simultaneously in the same frequency band, most cooperative strategies are based on the half-duplex mode. When considering the relay channel S-R-D in Fig 4.1(a), with a source S, a relay R and a destination D, each transmission is divided into two blocks: in the first block, the source transmits and the relay and the destination receive; in the second block the relay retransmits and the destination receives.



(a) Relay channel: one source, one relay, one destination



(b) Cooperative Interference Channel: two cooperating sources and two destinations

Figure 4.1: Small dimension cooperative networks

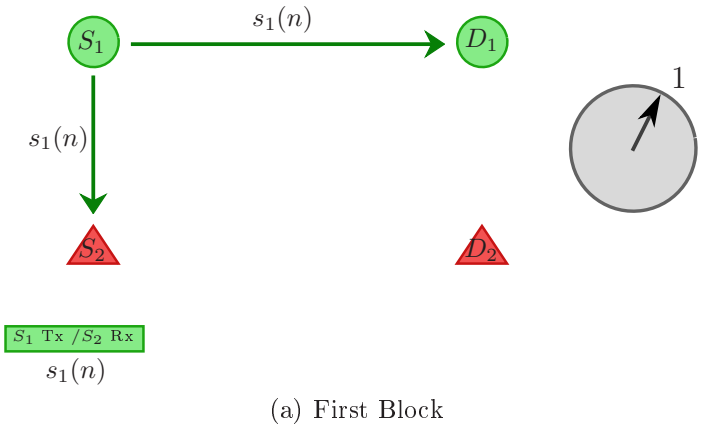
Now let us consider the four-node network in Fig. 4.1(b) with two sources S_1 and S_2 transmitting in a cooperative fashion to two destinations D_1 and D_2 as in [25]. The previous transmission scheme is repeated twice, first for the relay channel $S_1 - S_2 - D_1$ and second for the relay channel $S_2 - S_1 - D_2$ as described in Fig. 4.2 and Fig. 4.3 (b), resulting in a four-block transmission. The use of orthogonal interference free channels for sources and relays transmissions simplifies receivers algorithms but results in a loss of bandwidth. Relaxing the orthogonality constraint may help improving the spectral efficiency of the network, but it will also lead to the introduction of interference in the system.

Given those observations, we examine hereafter the following questions:

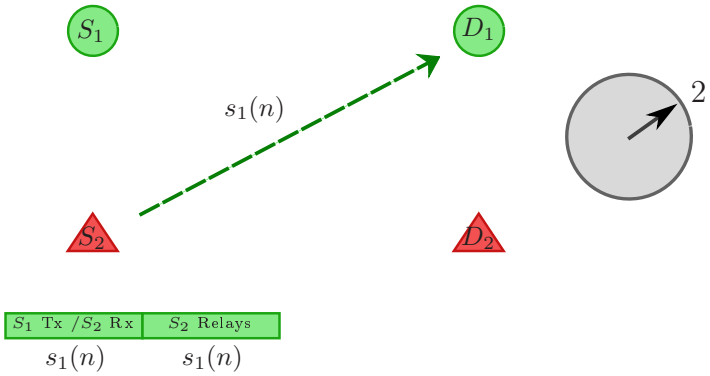
- Can we improve the spectral efficiency of cooperative strategies by relaxing the orthogonality constraint while still meeting the half-duplex constraint?
- How can we mitigate the interference due to the relaxation of the orthogonality constraint?

4.1.2 Contribution

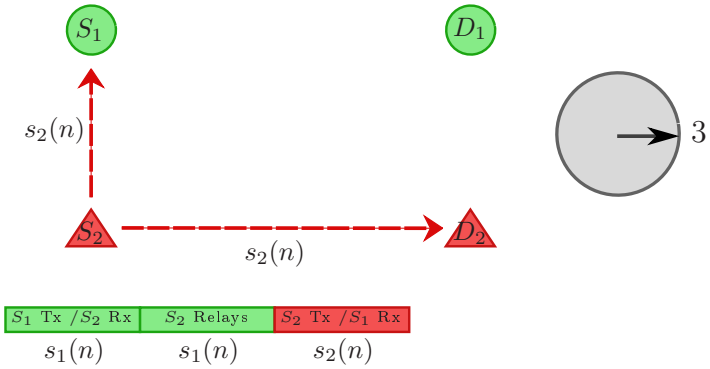
Orthogonality in classical cooperative strategies is two-fold, as illustrated in Fig. 4.3 (b):



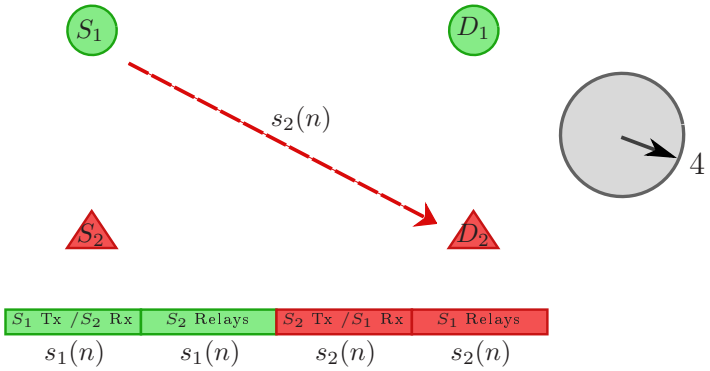
(a) First Block



(b) Second Block



(c) Third Block



(d) Fourth Block

Figure 4.2: Orthogonal Strategies

- inter-device orthogonality: S_1 and S_2 do not transmit signals simultaneously. S_1 and S_2 use Time Division Multiple Access (TDMA) to access the medium in orthogonal subspaces, and thus avoid interference. Moreover this orthogonality allows to respect the half-duplex constraint. Indeed, if S_2 transmitted at the same time as S_1, S_2 would not be able to receive the symbols from S_1, and thus could not relay them later.

- intra-device orthogonality: S_1 transmits different signals— own source signal and relayed signal— in orthogonal subspaces.

Relaxation of the orthogonality constraint can thus be two-fold, and we discuss each possibility of relaxing orthogonality hereafter.

Non-Orthogonal strategies relax the inter-device orthogonality: Acting on the first type of orthogonality, inter-device orthogonality, consists in making two devices access the link at the same time. For example, S_1 transmits a new message while S_2 is relaying a former message from S_1. This leads to cooperative strategies known as non-orthogonal strategies. However, non-orthogonal strategies cannot both meet the half-duplex constraint and make all source symbols benefit from cooperative diversity. Indeed if the half-duplex constraint is met, when S_2 is relaying a message from S_1, S_2 cannot receive at the same time the new message transmitted by S_1. Consequently, half the messages transmitted by S_1 cannot be relayed. Moreover, the orthogonality constraint leads to interference between the signals transmitted simultaneously.

Combination strategies relax the intra-device orthogonality: Acting on the second type of orthogonality, intra-device orthogonality, consists in superposing signals from different origins in a single signal transmitted by a single device. For example, S_1 produces a signal combining its own message with a former message received from S_2. This leads to strategies that we call combination strategies, that recall techniques used at the network layer in order to use efficiently the degrees of freedom of a system, namely Network Coding (NC).

Indeed, loss of bandwidth issues have been tackled at higher layers thanks to network coding [62–67]. Packets arriving at a node on any edge of a network are stocked into a single buffer. At each transmission opportunity, an

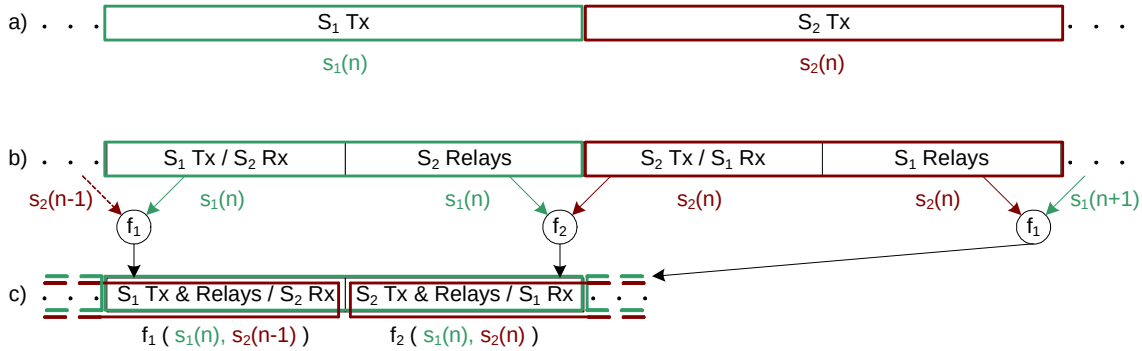


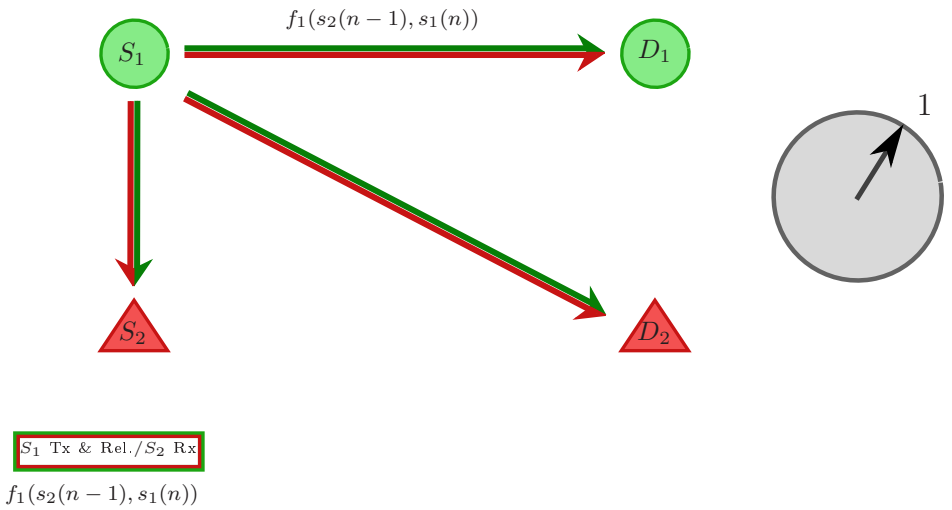
Figure 4.3: Time division channel allocations for (a) orthogonal direct transmissions, (b) usual orthogonal cooperative transmissions (c) proposed scheme: analog network coding cooperative transmissions

output packet is generated as a random linear combination of packets in the buffer within "current" generation [68]. Inspired by network coding, consider a four-node cooperative network using "network precoding" in a two-block transmission scheme, where in each single block one source simultaneously transmits and relays as in Fig. 4.3 (c) and Fig. 4.4:

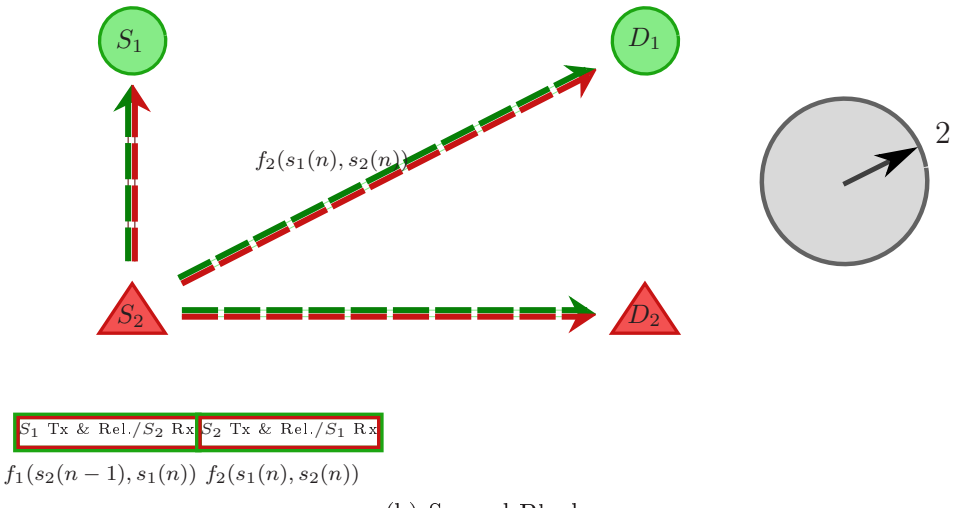
- **first block**: S_1 sends a single signal $f_1(s_1(n), s_2(n-1))$ which is a function of both its own message $s_1(n)$ and a message $s_2(n-1)$ received, decoded and re-encoded by S_1 in the second block of previous transmission (repetition of the codeword - RDF - or use of an independent codeword -PDF), now relayed for S_2. S_2, D_1 and D_2 receive. Since S_2 knows the message in $s_2(n-1)$, it can extract $s_1(n)$, if it also knows the mixing function f_1.

- **second block**: S_2 sends a single signal $f_2(s_2(n), s_1(n))$ which is a function of both its own message $s_2(n)$ and a message $s_1(n)$ received, decoded and re-encoded by S_2 in the first block of the current transmission, now relayed for S_1. S_1, D_1 and D_2 receive. Since S_1 knows the message in $s_1(n)$, it can extract $s_2(n)$, if it also knows f_2.

Functions f_1 and f_2 are the network precoding functions which help improving communication in terms of bandwidth. Knowing f_1 and f_2 allows sources S_2 and S_1 to easily cancel interference and extract the message they will have to relay in next block. But unfortunately, improving the bandwidth utilization has a cost: the introduction of interference at destinations D_1 and



(a) First Block



(b) Second Block

Figure 4.4: Combination strategies

D_2. In the first block, $s_2(n-1)$ is intended to D_2 as relayed signal and acts as interference for D_1, which is only interested in $s_1(n)$; reciprocally, $s_1(n)$, intended to D_1, generates interference for D_2 interested in $s_2(n-1)$. A similar interference problem occurs in the second block. Nevertheless, interference is known at transmitter, thus one can design the precoding functions to take into account this issue. In particular Dirty Paper Coding (DPC) [69], a well-known coding technique to mitigate interference known at transmitter, may help NC. We may expect DPC network precoding to help improving bandwidth efficiency in a cooperative network as well as mitigating interference, thus enhancing performance with respect to usual cooperative schemes.

Our main contributions are to bring network coding in the analog domain at the physical layer, to provide novel cooperative protocols with better spectral efficiency, and to analyze their performance in terms of network throughput and throughput per node. Thanks to Analog Network Coding combined with Dirty Paper precoding, the orthogonality constraint is relaxed, allowing to save time compared to classical DF protocols, and interference resulting from non-orthogonality is mitigated. As a result, bandwidth resource is used more efficiently and the spectral efficiency of the system is improved. Analysis show that our cooperative strategies significantly outperform classical orthogonal DF protocols [70–72].

4.1.3 Related work

As mentioned in the previous section, non-orthogonal strategies allow several devices, sources and relays, to transmit at the same time. In [73] non-orthogonal Amplify and Forward (NAF) protocols - yet preserving the half-duplex constraint - are proposed. In NAF, orthogonality constraint is relaxed by letting a source transmit new symbols when a relay is retransmitting former symbols it received from the source. NAF turns out to improve performances with respect to classical AF. Nevertheless with NAF, because of the half-duplex constraint, only half of the symbols can be relayed and benefit from cooperative diversity. In the scheme we proposed, orthogonality between source and relayed signals is also relaxed, half-duplex preserved, but all symbols benefit from cooperative transmission.

In combination strategies, several messages from different origins are bundled in a single transmitted signal. In [74] a cooperation strategy was proposed for two transmitters and one destination. Each source transmits both information of its own and of its partner, orthogonally superposed using or-

thogonal spreading codes, and this scheme is shown to improve user capacity. Nevertheless, a common destination is assumed for the cooperating pair, the half-duplex constraint is not taken into account, and cooperative periods are divided into two parts: slots where sources transmit only their own signal and slots where they send a cooperative signal. Our proposed scheme is more efficient, because no orthogonality constraint is imposed for source and relayed signal separation. In [75] coded cooperation (CC) is introduced in a system with two sources and one destination and is shown to outperform AF and RDF. However this scheme relies on the frame separation of the source signal and the relayed signal, leading again to bandwidth loss, and a common destination is assumed, a particular case of cooperative system.

All these works considered a common destination and did not address interference mitigation issues arising in multi-source multi-destination cooperative ad hoc system. In [70–72], we addressed this issue by proposing more spectrally efficient cooperative strategies, obtained by both relaxing the orthogonality constraint and mitigating the consequent interference. Recently [76] studied AF with analog network coding and showed that joint relaying and network coding can enhance the network throughput.

Introduced in [69], DPC is a well-known coding technique to mitigate interference known at transmitter. DPC was also considered in relay networks, eg. in [77–80], as joint coding between cooperating pairs, or to mitigate interference at relay. In [77] DPC transmit cooperation scheme suffers from loss of bandwidth due to the orthogonal cooperation channel used to exchange transmit messages between the two sources and whose cost is not taken into account. In [78], an orthogonal cooperation channel is also used, and DPC is jointly performed by the two transmitters acting like a MIMO Broadcast. In [79], a full duplex S-R-D network is considered, in which the source S sends a signal consisting of two components, one intended to the relay R and one intended to the destination D. In this relay network, Dirty Paper precoding is used at source to mitigate the interference caused at the relay by the second component. On the contrary, in our cooperation scheme, NC takes care of interference at the relay, whereas DPC is used at source and at relay to mitigate interference caused at destinations. In [80] DPC is considered for full-duplex transmit cooperation, with the sources jointly deciding the codewords both will combine in their transmit signals, which needs some signaling to agree on the codewords, not taken into account in the resource expenses. Besides the DPC-ordering is fixed before power allocation optimization, which impacts the individual rates and makes one destination

use forward-decoding and the other backward-decoding. On the contrary, as in [25] we consider a TDMA scheme, but with a time shift between the decoding of received signals at destinations, allowing to respect the half-duplex constraint, while NC allows to maintain a continuous flow of information interesting both destinations. Therefore our strategies are the first to manage combining the half-duplex constraint in the [25]-fashion and the continuous transmission of data interesting all destinations in the [80]-way. Moreover in our schemes, each source chooses its codewords alone, without needing to know what the other chose and both sources select the best DPC-orderings as part of the optimization, which they can achieve alone as long as channel information is available. Finally both destinations can use forward-decoding and do not to need to wait until the end of a frame of codewords to decode backward the first codeword sent.

4.2 System Model

To capture the gain resulting from the NC approach, we consider that all terminals are equipped with a single antenna. Consider the four node network in fig. 4.1(b). Each source S_i , $i \in \{1, 2\}$ generates a sequence $s_i(n)$, $n \in \{1, .., N\}$. These symbols are modeled by independent identically distributed (i.i.d.) circularly-symmetric complex gaussian random variables, with zero mean and variance $\varepsilon_s = \mathbb{E}[|s_i(n)|^2]$. With a transmission bandwidth W, there are W complex symbols per second. At time $t = k/W$, $k \in \mathbb{N}$, the signal transmitted by S_i is denoted $x_i(k)$ whereas $y_{S_i}(k)$ and $y_{D_j}(k)$ represent the signals received by source S_i and destination D_j respectively, with $i, j \in \{1, 2\}$. Finally f_i represents the network coding function performed at S_i. Those functions can be of any kind, not necessarily linear. We first focus on affine functions fi and then we develop a general network coding approach ... To simplify the analysis and the detection at destinations, we first focus on functions f_i performing a linear operation on symbols s_1 and s_2. Then a DPC approach is considered and shown to outperform the other strategies.

As described in section 4.1 and Fig. 4.3 (c), NC cooperative communication divides each transmission into two blocks.

- **First block** at even time indexes $k = 2n$, signals transmitted by S_1

and received by other terminals are:

$$\begin{aligned} x_1(2n) &= f_1(s_1(n), s_2(n-1)) \\ y_{S_2}(2n) &= h_{S_2S_1}\, x_1(2n) + z_{S_2}(2n) \\ y_{D_j}(2n) &= h_{D_jS_1}\, x_1(2n) + z_{D_j}(2n)\ ,\ j \in \{1,2\} \end{aligned}$$

- **Second block** at odd time indexes $k = 2n+1$, signals transmitted by S_2 and received by other terminals are:

$$\begin{aligned} x_2(2n+1) &= f_2(s_1(n), s_2(n)) \\ y_{S_1}(2n+1) &= h_{S_1S_2}\, x_2(2n+1) + z_{S_1}(2n+1) \\ y_{D_j}(2n+1) &= h_{D_jS_2}\, x_2(2n+1) + z_{D_j}(2n+1)\ ,\ j \in \{1,2\} \end{aligned}$$

The channel between transmitter $u \in \{S_1, S_2\}$ and receiver $v \in \{S_1, S_2, D_1, D_2\}$ is represented by h_{vu} which includes the effects of path-loss, shadowing and slow flat fading. These channel coefficients are modeled by independent circularly-symmetric complex gaussian random variables with zero-mean and variance σ^2_{vu}, i.e. Rayleigh fading. $z_v(k)$ are i.i.d. circularly-symmetric complex gaussian noises at receivers, with variance σ^2. Block-fading is assumed: channel gains are constant during a coherence block, and vary independently from one block to another. Each source has a power constraint in the continuous time-channel of P Joules/s and transmits only half of the time, both in the orthogonal interference-free cooperation schemes and in the proposed NC cooperation schemes. Thus the power constraint translates into $P_i = \mathbb{E}[|x_i(n)|^2] \leq \frac{2P}{W}$. Since a source transmits only part of the time, it can increase its transmit power in its transmission block and still meet its average power constraint for the whole transmission.

Finally, each destination is assumed to have perfect CSI of its two incoming channels from sources, whereas sources are assumed to have knowledge of the amplitudes of all channels and perfect CSI of the source-source channel. The knowledge of source-destination channel amplitudes can be obtained at sources thanks to feedback from destinations (the rate of CSI is usually lower than the data rate) once destination nodes have determined the channel gains on their respective incoming links. Note that block-fading is considered and we assume that CSI is available at transmitting nodes. Thus sources can perform rate-adaptation: at each block, a transmitter can select a code of appropriate rate depending on the current block channel condition. This

variable-rate coding scheme [81] will have an average throughput equal to the ergodic rate that would be obtained if the channel varied fast enough for a codeword of infinite length to experience enough independent fading states.

4.3 Precoding Method

4.3.1 Linear Precoding

In Linear Network Coding for RDF, S_1 detects $s_2(n-1)$ in the signal transmitted by S_2 and re-encodes it using the same codeword. Then S_1 forms its transmitted signal $x_1(n)$ as a linear combination of its own codeword $s_1(n)$ and the repeated $s_2(n-1)$. The same process happens at S_2. Therefore function f_i can be represented by a matrix $\mathbf{F}_i$ of size $N_t \times N_s$, i.e. (number of transmit antennas at source) times (number of symbols on which f_i acts). In the single antenna scenario, $\mathbf{F}_i = [f_{i1}, f_{i2}]$ is a row of size 2. Transmitted signals are thus:

$$x_1(2n) = \mathbf{F}_1 \left[s_1(n), s_2(n-1)\right]^T = f_{11}s_1(n) + f_{12}s_2(n-1)$$
$$x_2(2n+1) = \mathbf{F}_2 \left[s_1(n), s_2(n)\right]^T = f_{21}s_1(n) + f_{22}s_2(n)$$

In Linear NC cooperation scheme, the power constraint becomes $P_i = \varepsilon_s \|\mathbf{F}_i\|_F^2 \leq \frac{2P}{W}$. We will consider precoding functions such that $\|\mathbf{F}_i\|_F^2 = 1$, i.e. f_i does not increase the power transmitted by source S_i but shares it between the source message and the relayed message.

Remark : orthogonal TDMA transmissions without relaying can be seen as a particular case of network coding where $\mathbf{F}_1 = [1, 0]$ and $\mathbf{F}_2 = [0, 1]$. Orthogonal interference-free cooperation [25] is also a particular case of our scheme where $\mathbf{F}_1 = [1, 0]$ and $\mathbf{F}_2 = [1, 0]$ during two blocks, and then $\mathbf{F}_2 = [0, 1]$ and $\mathbf{F}_1 = [0, 1]$ during the next two blocks.

4.3.2 Dirty Paper Precoding

Since interference resulting from NC approach is known at the transmitter, more advanced NC functions can include decoding and re-encoding with DPC of messages intended to different destinations [82]. In Dirty Paper NC for PDF, S_1 decodes the message carried by $s_2(n-1)$ and re-encodes it using an independent Gaussian codebook. More precisely, in order to use dirty paper coding, S_1 first orders destinations based on channel knowledge. Then S_1

picks a codeword for the first destination, before choosing a codeword for the second destination, with full non-causal knowledge of the codeword intended to first destination. Thus the second destination does not see interference caused by the codeword for the first destination, whereas the first destination will see the signal intended to the second destination as interference. The signal transmitted by S_1 is the sum of the two codewords, with power sharing across the two codewords taking into account channel knowledge. S_2 will proceed the same way in the following block. The ordering of destinations chosen at each source affects performance. Transmitted signals thus become:

$$x_1(2n) = f_{11}s_1(n) + f_{12}s_2'(n-1)$$
$$x_2(2n+1) = f_{21}s_1'(n) + f_{22}s_2(n)$$

where f_{ij}^2 stands for the power allocated by source S_i to the codeword intended to destination D_j, and s_j' is the independent codeword produced by a source acting as relay after decoding the message carried by s_j. Destinations are assumed to know the orderings (each source can with a single bit indicate the ordering it selected).

4.4 Performance Analysis

In this section, the average rates, as well as the network throughput and the throughput per user are analyzed.

4.4.1 Orthogonal Interference-Free RDF and PDF

For cooperative channels in Fig. 4.3 (b), using RDF the mutual information between input s_1 and output y_{D_1} at D_1 is [25]:

$$I_{RDF}(s_1; y_{D_1}) = \frac{1}{2}\min\{\log(1+\rho|h_{S_2S_1}|^2), \log\left(1+\rho|h_{D_1S_1}|^2+\rho|h_{D_1S_2}|^2\right)\} \tag{4.1}$$

where the input SNR is $\rho = \varepsilon_s/\sigma^2 = 2P/(W\sigma^2)$, and the factor $1/2$ is due to the two channel-uses to send a message. Mutual information $I_{RDF}(s_2; y_{D_2})$ between input s_2 and output y_{D_2} at D_2 is given similarly. Half the degrees of freedom are allocated for transmission to a destination - each destination

is passive half of the time when the signals transmitted do not contain information intended to that destination- therefore the throughput of the first user is $\frac{1}{2}I_{RDF}(s_1; y_{D_1})$ and the total network throughput using RDF is:

$$C_{RDF} = \frac{1}{2}I_{RDF}(s_1; y_{D_1}) + \frac{1}{2}I_{RDF}(s_2; y_{D_2}) \tag{4.2}$$

Using PDF, mutual information between s_1 and y_{D_1} is [83]:

$$\begin{aligned} I_{PDF}(s_1; y_{D_1}) =& \frac{1}{2}\min\{\log(1+\rho|h_{S_2S_1}|^2), \\ & \log(1+\rho|h_{D_1S_1}|^2) + \log(1+\rho|h_{D_1S_2}|^2)\} \end{aligned} \tag{4.3}$$

Mutual information $I_{PDF}(s_2; y_{D_2})$ at D_2 is also given by a similar formula [83]. The total network throughput of PDF is given by:

$$C_{PDF} = \frac{1}{2}I_{PDF}(s_1; y_{D_1}) + \frac{1}{2}I_{PDF}(s_2; y_{D_2}) \tag{4.4}$$

4.4.2 Linear NC RDF

For our proposed network coding cooperative scheme in Fig. 4.3 (c), when the network coding functions are linear transformations, mutual information between input s_1 and output y_{D_1} at destination D_1 can be shown to be:

$$\begin{aligned} I_{LNC}(s_1; y_{D_1}) = \frac{1}{2}\min\Big\{&\log\left(1+\rho|h_{S_2S_1}f_{11}|^2\right), \\ &\log\left(1+\rho\frac{|h_{D_1S_1}f_{11}|^2}{1+\rho|h_{D_1S_1}f_{12}|^2} + \rho\frac{|h_{D_1S_2}f_{21}|^2}{1+\rho|h_{D_1S_2}f_{22}|^2}\right)\Big\} \end{aligned} \tag{4.5}$$

In the minimum in equation (4.5), the first term represents the maximum rate at which relay S_2 can decode the source message s_1 after canceling the interference known at the relay (interference is due to the symbol s_2 the relay emitted previously), whereas the second term represents the maximum rate at which destination D_1 can decode given the transmissions from source S_1 and relay S_2.

A similar formula gives the mutual information between input s_2 and output y_{D_2} at destination D_2, with appropriate changes:

$$\begin{aligned} I_{LNC}(s_2; y_{D_2}) = \frac{1}{2}\min\Big\{&\log\left(1+\rho|h_{S_1S_2}f_{22}|^2\right), \\ &\log\left(1+\rho\frac{|h_{D_2S_2}f_{22}|^2}{1+\rho|h_{D_2S_2}f_{21}|^2} + \rho\frac{|h_{D_2S_1}f_{12}|^2}{1+\rho|h_{D_2S_1}f_{11}|^2}\right)\Big\} \end{aligned} \tag{4.6}$$

In the NC approach, all degrees of freedom are used for transmission to each destination. No time is wasted from the destination point of view, thus the throughput of the first user is $I_{LNC}(s_1; y_{D_1})$ and the total network throughput for this strategy is :

$$\begin{aligned} C_{LNC} = \max_{\{f_{ij}\}_{i,j\in\{1,2\}}} \quad & I_{LNC}(s_1; y_{D_1}) + I_{LNC}(s_2; y_{D_2}) \\ |f_{11}|^2 + |f_{12}|^2 \leq 1 & \\ |f_{21}|^2 + |f_{22}|^2 \leq 1 & \end{aligned} \tag{4.7}$$

The optimization problem turns out to be a non-convex problem, both for Linear NC and for DPC in next section, so that classical convex optimization techniques cannot be used to find a closed-form expression of the power allocation scheme. Moreover, because of limitations due to the quality of the source-relay link, the MAC-BC duality [84] cannot be used to solve the optimization problem as in non-cooperative systems. Finding the optimal power allocation scheme between transmitted and relayed signals at each source is different from BC power allocation problem, because power terms f_{11}^2 and f_{22}^2 appear in the capacity of the links between the two sources, first terms in the minimums in formulas (4.5), (4.6), (4.8), so that the power allocation scheme maximizing the sum-rates of the two BC channels between a source and the two destinations may not be the same as the one maximizing the sum-rate of the cooperative system.

4.4.3 Dirty Paper NC PDF

The mutual information between a source message and the received signals at the intended destination depends on the two orderings Π_1, Π_2 of destinations for DPC chosen by both sources. Knowing all channel amplitudes, each source can compute alone the DPC orderings maximizing the network throughput. Since a relay uses an independent codeword to re-encode the signal it received from the previous source, the total network throughput for this cooperation scheme belonging to the family of PDF can be written :

$$\begin{aligned} C_{DPC} = \max_{\Pi_1, \Pi_2, \{f_{ij}\}_{i,j\in\{1,2\}}} \quad & I_{DPC}(s_1; y_{D_1}) + I_{DPC}(s_2; y_{D_2}) \\ |f_{11}|^2 + |f_{12}|^2 \leq 1 & \\ |f_{21}|^2 + |f_{22}|^2 \leq 1 & \end{aligned}$$

with :

$$\begin{aligned} I_{DPC}(s_1; y_{D_1}) = \frac{1}{2} \min \big\{ & \log\left(1 + \rho |h_{S_2 S_1} f_{11}|^2\right), \\ & \log(1 + SINR_{11}) + \log(1 + SINR_{21}) \big\} \\ I_{DPC}(s_2; y_{D_2}) = \frac{1}{2} \min \big\{ & \log\left(1 + \rho |h_{S_1 S_2} f_{22}|^2\right), \\ & \log(1 + SINR_{12}) + \log(1 + SINR_{22}) \big\} \end{aligned} \tag{4.8}$$

where $SINR_{ij}$ is the Signal-to-Interference plus Noise ratio resulting from the signal transmitted by S_i at D_j:

$$SINR_{ij} = \begin{cases} \rho |h_{D_j S_i} f_{ij}|^2 \text{ , if } S_i \text{ does DPC in favor of } D_j \\ \frac{\rho |h_{D_j S_i} f_{ij}|^2}{1 + \rho |h_{D_j S_i} f_{i\bar{j}}|^2} \text{ , if } S_i \text{ does DPC in favor of } D_{\bar{j}} \end{cases}$$

4.5 Numerical Results

In this section, numerical results are presented to compare the different cooperation strategies. Fig. (4.5) and (4.6) illustrate average per user throughput and total network throughput obtained through Monte Carlo Simulations (1000 channel realizations), in the case of symmetric networks, i.e. where the fading variances are identical $\sigma^2_{vu} = 1$. Optimal power allocations and orderings Π_i were obtained numerically by exhaustive search. Average individual throughput and outage probability are the same for both users, since they are assumed to have the same power constraints and the network is symmetric.

4.5.1 Average Throuhputs

Fig. (4.5) compares RDF [25] and Linear NC for RDF that we propose, and shows that our technique based on Linear Network coding performs much better in terms of per user throughput, thanks to a more efficient use of spectral resources as well as power resources. Fig. (4.6) plots the per user throughputs for PDF [25] and our DPC-NC for PDF. Once again, the NC based strategy enhances performance in terms of individual throughput.

Finally fig. (4.5) and (4.6) also allow to compare the total network throughput of all techniques, and show neat improvements in the network performance thanks to NC methods. Thanks to smart power sharing between

the source and relayed signals, even with repetition coding, and increased spectral efficiency, Linear NC enhances considerably performance compared to classical RDF and PDF. Using a more advanced coding technique, DPC, to mitigate interferences generated at destination by the NC methods leads to even better results.

4.6 Conclusion

Inspired by network coding, we proposed new cooperative strategies for ad hoc networks, which improve the spectral efficiency of the cooperative system by relaxing the orthogonality constraint, while preserving the practical half-duplex constraint. The introduction of interference between source and relayed messages, when considering non-orthogonal transmission scheme, is mitigated thanks to precoding at transmitter. We presented two precoding approaches, linear NC with RDF and Dirty-Paper NC with PDF, relevant technique since the transmitter knows the interference. Thanks to precoding, linear or Dirty Paper based, the cost of the NC approach - introduction of interference - is less than the resulting gain in terms of spectral efficiency, and the performance analysis shows significant improvements in terms of throughput over classical RDF/PDF cooperative strategies.

Future work may include solving the optimization of the power allocation and DPC ordering in particular scenarios of relative channel gains, developing of a selective strategy to circumvent limitations due to the source-relay link, extending the strategies to multiple-antenna nodes, in particular assessing how beamforming can improve performance, and last but not least extending the cooperative strategies to a large network with several source-destination pairs.

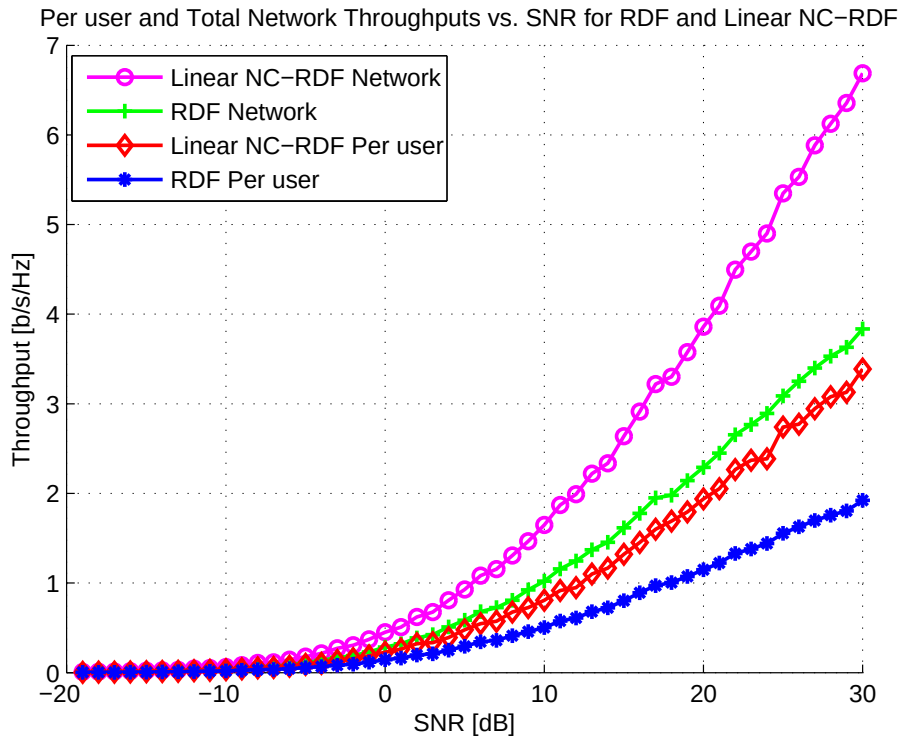


Figure 4.5: Comparison of Per user and Network Throughputs of classical RDF and LNC cooperative methods

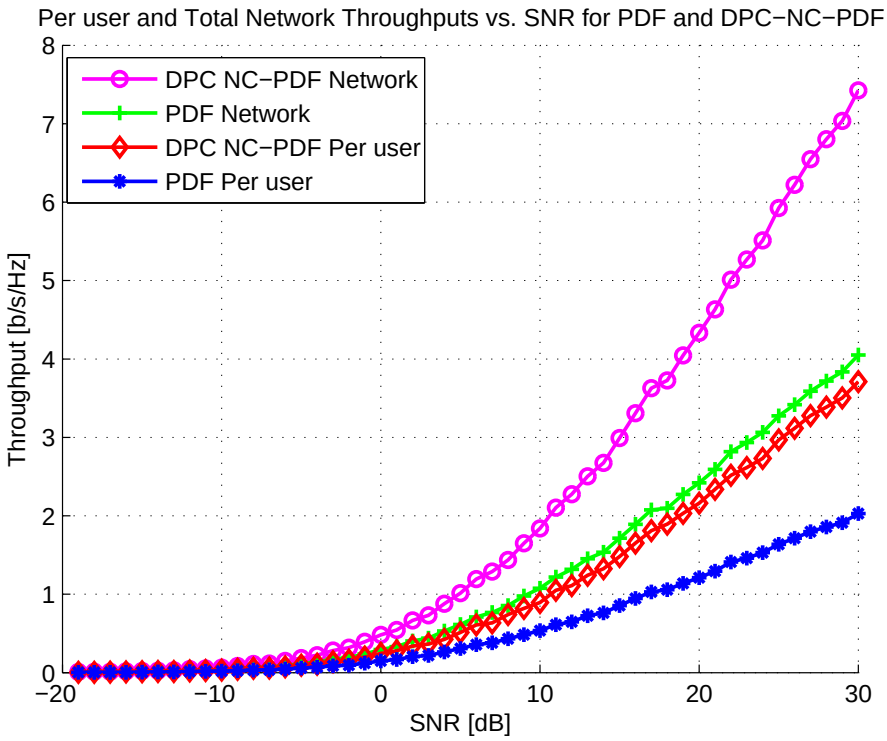


Figure 4.6: Comparison of Per user and Network Throughputs of classical PDF and NC-DPC cooperative methods

Chapter 5

Cooperation in Large Dimension Networks

5.1 Introduction

5.1.1 Motivation

When considering a large ad hoc network, where multiple source-destination pairs communicate simultaneously, the performance scalability with respect to the number of nodes, size or density of the network is a crucial issue [85]. Due to the increasing interference generated by an increasing number of communicating pairs, it may turn out that ad hoc networks are only appropriate for short range communication systems. In order to avoid the interference-limited behavior when the network density increases, some hybrid schemes combining cellular and ad hoc networks have been considered as a convenient alternative [3–5]. However, following the development of MIMO systems, another alternative to combat the interference-limited behavior started emerging: cooperative virtual MIMO networks, where wireless nodes interact to form a virtual antenna array.

In the following, we first review the performance results of ad hoc networks [1, 86] without cooperation between nodes. Then, we address the potential gains resulting from the cooperative approach in ad hoc networks.

Interference-limited dense networks

The pioneering work by Gupta and Kumar [1] on the capacity of wireless ad hoc networks paved the way to many research works on the scaling laws of large networks. Considering a dense ad hoc network with n nodes capable of transmitting at W bits/second, located in a disk of unit area, they showed that the performance of the network was limited by interference and decreased dramatically as the density of the network increased. More precisely, they showed that

- when the n nodes are randomly located, the transport capacity scales as $\Theta(W\sqrt{n})$, and the throughput per node scales as $\Theta\left(\frac{W}{\sqrt{n\log n}}\right)$ bits/second;
- when the n nodes are optimally placed, and the traffic pattern and each transmission range are optimally chosen, the throughput per node scales as $\Theta\left(\frac{W}{\sqrt{n}}\right)$.

Based on those results, the deployment of very dense ad hoc networks appeared to be not desirable and the authors recommended to consider ad hoc

networks with a small number of nodes, or networks where nodes communicate via multi-hopping only with nearby nodes. Nevertheless, their analysis relied on the assumption that transmissions happened in a point-to-point multi-hop path between a source and its destination and that signals transmitted by different nodes and received simultaneously at a node represented interference impairing the communication. This assumption modeled each node as an isolated signal-processing unit, to which all surrounding nodes were potential interferers, and thus ignored the potential gains that could arise, at the physical layer, from cooperation between nodes and virtual MIMO.

Transmission-Range-limited wide networks

In [86], Xie and Kumar considered an ad hoc network where n nodes were located on a plane with a minimum separation distance between nodes, resulting in network with a fixed density but growing size when the number of nodes n increased. They showed that under the assumption of equal per-node power constraints, the transport capacity scaled linearly with the number of nodes for pathloss exponents greater than 6, and that the linear scaling could be achieved by decode-and-forward point-to-point multi-hopping, treating other simultaneous transmissions as interference. Their work initiated a series of works [87–90], each work improving progressively the range of pathloss exponents for which a linear scaling of the transport capacity was possible, until pathloss exponents greater than 4 in [89].

In networks with a fixed density and an increasing area when the number of nodes increases, the distance between source and destination tends to increase. When the pathloss attenuation is fast enough, communication is impaired less by interference than by the transmission range that the transmitted power allows to reach. Optimal scaling being possible with simple multi-hopping, considering more complex transmission scheme may appear irrelevant. Nevertheless, in the regime of slow pathloss attenuation—pathloss exponent between 2 and 4 — where interference may play a larger role as in dense networks, the scaling law of the transport capacity is still an open problem, and gains at the physical layer may be possible through the use of more advanced transmission techniques.

Cooperative virtual MIMO

In a ad hoc network, when distributed wireless nodes interact to jointly transmit or receive information, they form a virtual antenna array and their cooperation enables the exploitation of MIMO gains [91]. Indeed, in regular point-to-point communications, MIMO techniques are known to be able to improve

- spectral efficiency: a spatial multiplexing gain is provided by the simultaneous transmission of independent data streams over different antennas, and results in a linear increase of the capacity with the minimum number of transmit and receive antennas;
- reliability: a spatial diversity gain is provided by making a data stream experience several independent fadings, e.g. with Space-Time-Codes.

In a ideal ad hoc network, where all nodes could perfectly cooperate both at transmission and reception, one could expect the network to behave like a perfect MIMO system. The total capacity would then grow linearly with the number of node pairs as in the point-to-point MIMO channel [38]. However perfect transmit/receive cooperation has a cost, in particular it requires perfect synchronization between cooperating nodes and extensive channel knowledge, which is impossible in practice. Making such a perfect cooperation assumption over the whole network is unrealistic, yet a reasonable alternative consist in dividing the network in several cooperative clusters. Each cluster would gather nodes cooperating in a local area to form a virtual antenna array and virtual MIMO transmissions would then occur in multiple stages between these cooperative clusters. This approach offers a way to leverage the broadcasting nature of the wireless link, and differs from [1] essentially in that simultaneous transmissions are not treated as interference in a point-to-point communication setting [1], but as useful signals that are jointly and cooperatively processed at nodes within a cluster. Note that in cellular networks, current research efforts to improve the performance of cellular systems suggest a tendency towards clusters of base stations cooperating to form virtual antenna arrays serving mobile users.

In this chapter, we consider a dense cooperative ad hoc network, where nodes are grouped in three types of cooperative clusters:

- a source cluster, consisting of source nodes cooperating to form a virtual transmit antenna array;

- a destination cluster, gathering destination nodes cooperating to form a virtual receive antenna array;
- a series of intermediary relaying clusters, between source and destination clusters, each gathering relay nodes cooperating to form a virtual relay antenna array.

This system forms a virtual MIMO multi-hop relay network. Deriving the capacity of such a complex network involves many parameters and variables and is intricate. Fortunately, random matrix theory and free probability theory provide useful mathematical results on the eigenvalues and eigenvectors of large random matrices, that can be applied to the field of wireless communications in order to analyze complex communication systems [39]. Indeed, the transfer of information in a communication system can often be modeled by an equation $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$, where $\mathbf{x}$,$\mathbf{y}$, and $\mathbf{z}$ are respectively the input, output, and noise random vectors, and $\mathbf{H}$ is the random transfer matrix of the system. For such a system, most information-theoretic performance metrics can be shown to depend only on the eigenvalues and eigenvectors of the transfer matrix $\mathbf{H}$, therefore RMT and FPT appear relevant theories to give insight on the scaling behavior of these performance metrics when the system dimensions increase. Moreover, the self-averaging effect of random matrices with large dimensions has be shown to be able to reveal the relevant parameters impacting the performance communication schemes. As a first step towards a future more complete analysis of the performance of ad hoc networks with cooperative clusters, we start the analysis by making some simplifying assumptions. We assume that nodes in a cluster can perfectly exchange information, and that channel knowledge intrinsic to the cluster— channels between nodes within a cluster— is available at all cluster nodes. Thus in each cluster, nodes form a perfect virtual antenna array. We also assume non-noisy communications between relaying clusters. Under those simplifying assumptions and using tools from random matrix theory and free probability theory, we address, in this chapter, the following questions:

- What is the asymptotic capacity of the cooperative-cluster system when the number of nodes in all clusters grow large?
- What are the relevant parameters impacting the system capacity?
- How should nodes in a cooperative cluster process and transmit cooperatively their wireless signals to maximize the system capacity?

5.1.2 Contribution

We consider an N-hop MIMO relay communication system wherein data transmission from k_0 source antennas to k_N destination antennas is made possible through $N-1$ relay levels, each of which are equipped with $k_i,\ i = 1,\ldots,N-1$ antennas. In this transmission chain with $N+1$ levels, it is assumed that each relay receives a faded version of the multi-dimensional signal transmitted from the previous level and, after precoding, retransmits it to the next level. We consider the case where all communication links undergo block Rayleigh flat fading and the fading channels at each hop (between two adjacent levels) may be correlated while the fading channels of any two different hops are independent.

Using tools from the free probability theory and assuming that the noise power at the relay levels, but not at the destination, is negligible, we first derive a closed-form expression of the asymptotic instantaneous end-to-end mutual information between the source input and the destination output, as the number of antennas at all levels $k_i,\ i = 1,\ldots,N-1$ grows large. This asymptotic expression is shown to be independent from the channel realizations and to only depend on the channel statistics. Therefore, as long as the statistical properties of the channel matrices at all hops do not change, the instantaneous mutual information asymptotically converges to the same deterministic expression for any arbitrary channel realization, with two major consequences. First, in the asymptotic regime the mutual information is not a random variable any more but a deterministic value representing an achievable rate. This means that when the channel is random but fixed at the beginning of the transmission block, and the system size is large enough, the capacity in the sense of Shannon is not zero, on the contrary to the capacity of small size systems [38, Section 5.1]. Second, given the stationarity of channel statistical properties, the asymptotic instantaneous mutual information obtained in the non-ergodic block-fading regime also serves as the asymptotic value of the average end-to-end mutual information between the source and the destination. The expression of the asymptotic average mutual information is the same as the asymptotic ergodic mutual information that would be obtained if the channel was an ergodic process. Intuitively, the time-domain ergodicity is recovered in the spatial domain when the dimensions of the system grow large.

We also obtain the singular vectors of the optimal precoding matrices that maximize the average mutual information of the system with a finite

number of antennas at all levels. It is shown that the singular vectors of the optimal precoding matrices are also independent from the channel realizations and can be determined only using statistical knowledge of channel matrices at source and relays. The so-obtained singular vectors turn out to be also optimal in the asymptotic regime of concern. Therefore, combining asymptotic mutual information expression and optimal precoding singular vectors will pave the way for future work on optimal power allocation, i.e. finding the optimal precoding singular values.

Finally, we apply the aforementioned results on the asymptotic mutual information and the structure of the optimal precoding matrices to several communications scenarios, with different number of hops, and types of channel correlation. These examples illustrates the gains resulting from the cooperative cluster approach and reveal the few relevant parameters impacting the capacity of cooperative dense ad hoc networks.

5.1.3 Related works

MIMO relay communication systems have recently attracted much attention due to their potential to substantially improve the signal reception quality, in particular when the direct communication link between the source and the destination is not reliable. Due to its major practical importance as well as its significant technical challenge, deriving the capacity - or bounds on the capacity - of various relay communication schemes is growing to an entire field of research. Of particular interest is the derivation of capacity bounds for systems in which the source, the destination, and the relays are equipped with multiple antennas. Recent works using random matrix theory and free probability theory to derive scaling laws in multi-hop MIMO relay networks with large dimensions (nodes, antennas...) started emerging.

Several works focused on the capacity of two-hop relay networks, such as [26–31, 92]. Assuming fixed channel conditions, lower and upper bounds on the capacity of the two-hop MIMO relay channel were derived in [26]. In the same paper, bounds on the ergodic capacity were also obtained when the communication links undergo i.i.d. Rayleigh fading. The capacity of a two-hop MIMO relay system was studied in [27] in the asymptotic case where the number of relay nodes K grows large while the number of transmit and receive antennas M remains constant. In this setting, the capacity was shown to scale as $C = (M/2)\log K + O(1)$ when CSI is available at relays. The scaling behavior of the capacity in two-hop amplify-and-forward

(AF) networks was analyzed in [28–30] when the numbers of single-antenna sources, relays and destinations grow large. The achievable rates of a two-hop code-division multiple-access (CDMA) decode-and-forward (DF) relay system were derived in [93] when the numbers of transmit antennas and relays grow large. In [31], an ad hoc network with several source-destination pairs communicating through multiple AF-relays was studied and an upper-bound on the asymptotic capacity in the low Signal-to-Noise Ratio (SNR) regime was obtained in the case where the numbers of source, relay and destination nodes grow large. The scaling behavior of the capacity of a two-hop MIMO relay channel was also studied in [92] for bi-directional transmissions. In [94] the optimal relay precoding matrix was derived for a two-hop relay system with perfect knowledge of the source-relay and relay-destination channel matrices at the relay.

Following the work in [44] on the asymptotic eigenvalue distribution of concatenated fading channels, several analysis were proposed for more general multi-hop relay networks, including [32–37]. In particular, considering multi-hop MIMO AF networks, the tradeoffs between rate, diversity, and network size were analyzed in [32], and the diversity-multiplexing tradeoff (DMT) was derived in [33]. The asymptotic capacity of multi-hop MIMO AF relay systems was obtained in [34] when all channel links experience i.i.d. Rayleigh fading while the number of transmit and receive antennas, as well as the number of relays at each hop go to infinity with the same rate. Finally, recent works considering cooperative clusters in ad hoc networks started emerging [36,37]. In [36] a collaborative scheme was proposed for networks with n-nodes in the fixed received SNR regime—which includes dense networks — with a scaling of the per-node throughput in $\Omega(\frac{1}{n^{1/3}})$. In [37] hierarchical virtual MIMO networks were studied, the scaling laws of capacity were derived when the network density increases and the proposed hierarchical protocol was shown to achieve linear scaling of capacity in the number of single-antenna nodes randomly distributed in an area.

5.2 System Model

5.2.1 Multi-Hop MIMO Relay Network

Consider Fig. 5.1 that shows a multi-hop relaying system with k_0 source antennas, k_N destination antennas and $N-1$ relaying levels. The $i-$th

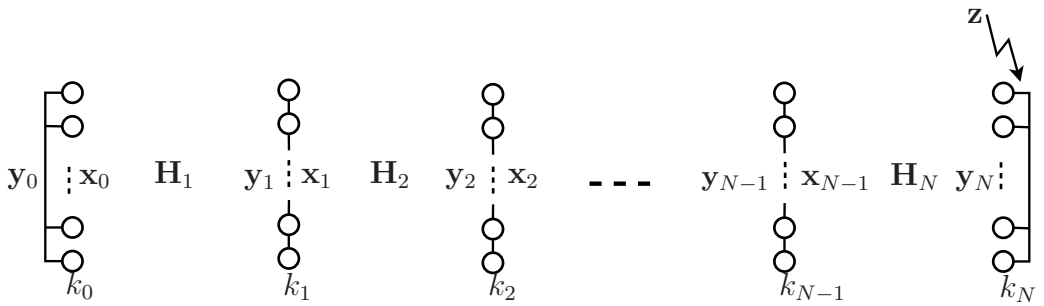


Figure 5.1: Multi-hop Relaying System

relaying level is equipped with k_i antennas. We assume that the noise power is negligible at all relays while at the destination the noise power is such that

$$\mathrm{E}[\mathbf{z}\mathbf{z}^H] = \sigma^2\mathbf{I} = \frac{1}{\eta}\mathbf{I} \tag{5.1}$$

where $\mathbf{z}$ is the circularly-symmetric zero-mean i.i.d. Gaussian noise vector at the destination. The simplifying noise-free relay assumption is a first step towards the future information-theoretic study of the more complex noisy relay scenario. Note that several other authors have implicitly used a similar noise-free relay assumption. For instance, in [33] a multi-hop AF relay network is analyzed and it is proved that the resulting colored noise at the destination can be well-approximated by white noise in the high SNR regime. In a multi-hop MIMO relay system, it can be shown that the white-noise assumption would be equivalent to assuming negligible noise at relays, but non-negligible noise at the destination.

Throughout the section, we assume that the correlated channel matrix at hop $i \in \{1, \ldots, N\}$ can be represented by the Kronecker model

$$\mathbf{H}_i \triangleq \mathbf{C}_{r,i}^{1/2}\mathbf{\Theta}_i\mathbf{C}_{t,i}^{1/2} \tag{5.2}$$

where $\mathbf{C}_{t,i}, \mathbf{C}_{r,i}$ are respectively the transmit and receive correlation matrices, $\mathbf{\Theta}_i$ are zero-mean i.i.d. Gaussian matrices independent over index i, with variance of the (k,l)-th entry

$$\mathrm{E}[|\theta_{kl}^{(i)}|^2] = \frac{a_i}{k_{i-1}} \qquad i = 1, \ldots, N \tag{5.3}$$

where $a_i = d_i^{-\beta}$ represents the pathloss attenuation with β and d_i denoting the pathloss exponent and the length of the i-th hop respectively. We also

assume that channels matrices $\mathbf{H}_i,\ i = 1, \ldots, N$ remain constant during a coherence block of length L and vary independently from a coherence block to another.

Note that the i.i.d. Rayleigh fading channel is obtained from the above Kronecker model when matrices $\mathbf{C}_{t,i}$ and $\mathbf{C}_{r,i}$ are set to identity.

Within a coherence block, the signal transmitted by the k_0 source antennas at time $l \in \{0, \ldots, L-1\}$ is given by the vector $\mathbf{x}_0(l) = \mathbf{P}_0\mathbf{y}_0(l-1)$, where $\mathbf{P}_0$ is the source precoding matrix and $\mathbf{y}_0$ is a zero-mean random vector with

$$\mathrm{E}[\mathbf{y}_0\mathbf{y}_0^H] = \mathbf{I}_{k_0} \qquad \text{and thus} \qquad \mathrm{E}[\mathbf{x}_0\mathbf{x}_0^H] = \mathbf{P}_0\mathbf{P}_0^H \tag{5.4}$$

Assuming that relays work in full-duplex mode, at time $l \in \{0, \ldots, L-1\}$ the relay at level i uses a precoding matrix $\mathbf{P}_i$ to linearly precode its received signal $\mathbf{y}_i(l-1) = \mathbf{H}_i\mathbf{x}_{i-1}(l-1)$ and form its transmitted signal

$$\mathbf{x}_i(l) = \mathbf{P}_i\mathbf{y}_i(l-1) \qquad i = 0, \ldots, N-1 \tag{5.5}$$

The precoding matrices at source and relays $\mathbf{P}_i,\ i = 0, \ldots, N-1$ are subject to the per-node long-term average power constraints[1]

$$\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H]) \leq k_i\mathcal{P}_i \qquad i = 0, \ldots, N-1 \tag{5.6}$$

It should be noticed that choosing diagonal precoding matrices would reduce the above scheme to the simpler AF relaying strategy.

As can be observed from Fig. 5.1, the signal received at the destination at time l is given by

$$\begin{aligned} \mathbf{y}_N(l) &= \mathbf{H}_N\mathbf{P}_{N-1}\mathbf{H}_{N-1}\mathbf{P}_{N-2}\ldots\mathbf{H}_2\mathbf{P}_1\mathbf{H}_1\mathbf{P}_0\mathbf{y}_0(l-N) + \mathbf{z} \\ &= \mathbf{G}_N\mathbf{y}_0(l-N) + \mathbf{z} \end{aligned} \tag{5.7}$$

[1]Recall that $\mathbf{y}_i = \mathbf{H}_i\mathbf{x}_{i-1}$. The power constraint on $\mathbf{x}_{i-1}$, given by $\mathrm{tr}(\mathrm{E}[\mathbf{x}_{i-1}\mathbf{x}_{i-1}^H]) \leq k_{i-1}\mathcal{P}_{i-1}$, along with the variance of the elements $\mathbf{H}_i$, given by $\mathrm{E}[|\theta_{kl}^{(i)}|^2] = \frac{a_i}{k_{i-1}}$, are such that the system is equivalent to a system where random channel elements $\theta_{kl}^{(i)}$ would be i.i.d. with variance a_i and the power constraint on the transmitted signal at level $i-1$ would be finite and equal to $\mathcal{P}_{i-1}$. That equivalent system, with finite transmit power at each level, makes sense from a physical point of view and shows that adding antennas, i.e. increasing the system dimension, does not mean increasing the transmit power from a physical point of view. Nonetheless, in order to derive the asymptotic instantaneous mutual information in Section 5.3, using random matrix theory tools, the variance of random channel elements is required to be normalized by the size of the channel matrix. That is the reason why the normalized model— channel variance and power constraint— was adopted.

where the end-to-end equivalent channel is

$$\begin{aligned}\mathbf{G}_N &\triangleq \mathbf{H}_N\mathbf{P}_{N-1}\mathbf{H}_{N-1}\mathbf{P}_{N-2}\ldots\mathbf{H}_2\mathbf{P}_1\mathbf{H}_1\mathbf{P}_0 \\ &= \mathbf{C}_{r,N}^{1/2}\mathbf{\Theta}_N\mathbf{C}_{t,N}^{1/2}\mathbf{P}_{N-1}\mathbf{C}_{r,N-1}^{1/2}\mathbf{\Theta}_{N-1}\mathbf{C}_{t,N-1}^{1/2}\ldots\mathbf{P}_1\mathbf{C}_{r,1}^{1/2}\mathbf{\Theta}_1\mathbf{C}_{t,1}^{1/2}\mathbf{P}_0 \quad (5.8)\end{aligned}$$

Let us introduce the matrices

$$\begin{aligned}\mathbf{M}_0 &= \mathbf{C}_{t,1}^{1/2}\mathbf{P}_0 \\ \mathbf{M}_i &= \mathbf{C}_{t,i+1}^{1/2}\mathbf{P}_i\mathbf{C}_{r,i}^{1/2} \qquad i = 1,\ldots,N-1 \\ \mathbf{M}_N &= \mathbf{C}_{r,N}^{1/2} \end{aligned} \quad (5.9)$$

Then (5.8) can be rewritten as

$$\mathbf{G}_N = \mathbf{M}_N\mathbf{\Theta}_N\mathbf{M}_{N-1}\mathbf{\Theta}_{N-1}\ldots\mathbf{M}_2\mathbf{\Theta}_2\mathbf{M}_1\mathbf{\Theta}_1\mathbf{M}_0 \quad (5.10)$$

For the sake of clarity, the dimensions of the matrices/vectors involved in our analysis are given below.

$\mathbf{x}_i : k_i \times 1$	$\mathbf{y}_i : k_i \times 1$	$\mathbf{P}_i : k_i \times k_i$
$\mathbf{H}_i : k_i \times k_{i-1}$	$\mathbf{C}_{r,i} : k_i \times k_i$	$\mathbf{C}_{t,i} : k_{i-1} \times k_{i-1}$
$\mathbf{\Theta}_i : k_i \times k_{i-1}$	$\mathbf{M}_i : k_i \times k_i$	

In the sequel, we assume that the channel coherence time is large enough to consider the non-ergodic case and consequently, time index l can be dropped. Finally, we define three channel-knowledge assumptions:

- Assumption $\mathbf{A_s}$, local statistical knowledge at source: the source has only statistical CSI of its forward channel $\mathbf{H}_1$, i.e. the source knows the transmit correlation matrix $\mathbf{C}_{t,1}$.
- Assumption $\mathbf{A_r}$, local statistical knowledge at relay: at the i^{th} relaying level, $i \in \{1,\ldots,N-1\}$, only statistical CSI of the backward channel $\mathbf{H}_i$ and forward channel $\mathbf{H}_{i+1}$ are available, i.e. relay i knows the receive correlation matrix $\mathbf{C}_{r,i}$ and the transmit correlation matrix $\mathbf{C}_{t,i+1}$.
- Assumption $\mathbf{A_d}$, end-to-end perfect knowledge at destination: the destination perfectly knows the end-to-end equivalent channel $\mathbf{G}_N$

Throughout the section, assumption $\mathbf{A_d}$ is always made. Assumption $\mathbf{A_d}$ is the single assumption on channel-knowledge necessary to derive the asymptotic mutual information in Section 5.3, while the two extra assumptions $\mathbf{A_s}$ and $\mathbf{A_r}$ are also necessary in Section 5.4 to obtain the singular vectors of the optimal precoding matrices.

5.2.2 Mutual Information

Under Assumption $\mathbf{A_d}$, the end-to-end mutual information between channel input $\mathbf{y}_0$ and channel output $(\mathbf{y}_N, \mathbf{G}_N)$ is [38]

$$\begin{aligned}\mathcal{I}(y_0;(y_N,G_N)) &= \mathcal{I}(y_0;y_N|G_N) + \underbrace{\mathcal{I}(y_0;G_N)}_{0} = \mathcal{I}(y_0;y_N|G_N) \\ &= \mathcal{H}(y_N|G_N) - \underbrace{\mathcal{H}(y_N|y_0,G_N)}_{\mathcal{H}(z)} \\ &= \mathrm{E}_{G_N}[\mathcal{H}(y_N|G_N=\mathbf{G}_N)] - \mathcal{H}(z)\end{aligned} \tag{5.11}$$

The entropy of the noise vector is known to be $\mathcal{H}(z) = \log\det(\frac{\pi e}{\eta}\mathbf{I}_{k_N})$. Besides, $\mathbf{y}_0$ is zero-mean with variance $\mathrm{E}[\mathbf{y}_0\mathbf{y}_0^H] = \mathbf{I}_{k_0}$, thus given $\mathbf{G}_N$, the received signal $\mathbf{y}_N$ is zero-mean with variance $\mathbf{G}_N\mathbf{G}_N^H + \frac{1}{\eta}\mathbf{I}_{k_N}$. By [38, Lemma 2] the inequality $\mathcal{H}(y_N|G_N=\mathbf{G}_N) \leq \log\det(\pi e\mathbf{G}_N\mathbf{G}_N^H + \frac{\pi e}{\eta}\mathbf{I}_{k_N})$ holds, and the entropy is maximized when equality is met for $\mathbf{y}_N$ circularly-symmetric complex Gaussian, which is the case when $\mathbf{y}_0$ is circularly-symmetric complex Gaussian. Therefore in the sequel, $\mathbf{y}_0$ is considered to be zero-mean circularly-symmetric complex Gaussian and the mutual information (5.11) can be rewritten

$$\begin{aligned}\mathcal{I}(y_0;(y_N,G_N)) &= \mathrm{E}_{G_N}[\log\det(\mathbf{I}_{k_N} + \eta\mathbf{G}_N\mathbf{G}_N^H)] \\ &= \mathrm{E}_{G_N}[\mathcal{I}(y_0;y_N|G_N=\mathbf{G}_N)]\end{aligned} \tag{5.12}$$

where $\mathcal{I}(y_0;y_N|G_N=\mathbf{G}_N) = \log\det(\mathbf{I}_{k_N} + \eta\mathbf{G}_N\mathbf{G}_N^H)$ is the instantaneous mutual end-to-end information for a channel realization $\mathbf{G}_N$.

To optimize the system, we are left with finding the precoders $\mathbf{P}_i$ maximizing the end-to-end mutual information (5.12) subject to power constraints (5.6), i.e. obtaining the maximum average end-to-end mutual information

$$\boldsymbol{C} \triangleq \max_{\{\mathbf{P}_i/\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H])\leq k_i\mathcal{P}_i\}_{i\in\{0,\ldots,N-1\}}} \mathrm{E}\left[\log\det(\mathbf{I}_{k_N} + \eta\,\mathbf{G}_N\mathbf{G}_N^H)\right] \tag{5.13}$$

Note that the non-ergodic regime is considered, therefore (5.12) represents only an average mutual information over channel realizations, and maximizing (5.12) does not a priori have the meaning of channel capacity in the Shannon sense when the system size is finite.

5.3 Asymptotic Mutual Information

In this section, we consider the instantaneous mutual information per transmit antenna between the source and the destination

$$\boldsymbol{I} \triangleq \frac{1}{k_0} \log\det(\mathbf{I}_{k_N} + \eta \mathbf{G}_N \mathbf{G}_N^H) \tag{5.14}$$

and derive its asymptotic value as the number of antennas $k_0, k_1, \ldots, k_N$ go to infinity. The following theorem holds.

Theorem 1. *For the system described in section 5.2, assume that*

- *perfect knowledge of the end-to-end channel* $\mathbf{G}_N$ *is available at the destination (Assumption* $\mathbf{A_d}$*)*
- $k_0, k_1, \ldots, k_N$ *go to infinity while* $\frac{k_i}{k_N} \to \rho_i, \qquad i = 0, \ldots, N$
- $\forall i \in \{0, \ldots, N\}$*, as* k_i *goes to infinity,* $\mathbf{M}_i^H \mathbf{M}_i$ *has a limit eigenvalue distribution with a compact support.*

Then the instantaneous mutual information per transmit antenna $\boldsymbol{I}$ *converges almost surely to*

$$\boldsymbol{I}_\infty = \frac{1}{\rho_0} \sum_{i=0}^{N} \rho_i \mathrm{E}\left[\log\left(1 + \eta \frac{a_{i+1}}{\rho_i} h_i^N \Lambda_i\right)\right] - N \frac{\log e}{\rho_0} \eta \prod_{i=0}^{N} h_i \tag{5.15}$$

where $a_{N+1} = 1$ *by convention,* $h_0, h_1, \ldots, h_N$ *are the solutions of the system of* $N+1$ *equations*

$$\prod_{j=0}^{N} h_j = \rho_i \mathrm{E}\left[\frac{h_i^N \Lambda_i}{\frac{\rho_i}{a_{i+1}} + \eta h_i^N \Lambda_i}\right] \qquad i = 0, \ldots, N \tag{5.16}$$

and the expectation $\mathrm{E}[\cdot]$ *in (5.15) and (5.16) is over* Λ_i *whose distribution is given by the asymptotic eigenvalue distribution* $F_{\mathbf{M}_i^H \mathbf{M}_i}(\lambda)$ *of* $\mathbf{M}_i^H \mathbf{M}_i$*.*

The detailed proof of *Theorem 1* is presented in Appendix 5.B.

Note that the above expression of the asymptotic instantaneous mutual information is valid for any arbitrary set of precoding matrices $\mathbf{P}_i$, $i =$

$0, \ldots, N-1$, such that $\mathbf{M}_i^H \mathbf{M}_i$ has a compactly supported asymptotic eigenvalue distribution when the system dimensions get large.[2]

Given a set of precoding matrices, it can be observed from (5.15) and (5.16) that the asymptotic expression is a deterministic value depending only on channel statistics and not on a particular channel realization. In other words, for a given set of precoding matrices, as long as the statistical properties of the channel matrices do not change, the instantaneous mutual information always converges to the same deterministic achievable rate, regardless of the channel realization. Thus, as the numbers of antennas at all levels grow large, the instantaneous mutual information is not a random variable anymore and the precoding matrices maximizing the asymptotic instantaneous mutual information can be found based only on knowledge of the channel statistics, without requiring any information regarding the instantaneous channel realizations. This further means that when the channel is random but fixed at the beginning of the transmission and the system size grows large enough, the capacity in the sense of Shannon is not zero any more, on the contrary to the capacity of small-size systems [38, Section 5.1].

Moreover, given the stationarity of channel statistical properties, the instantaneous mutual information converges to the same deterministic expression for any arbitrary channel realization. Therefore, the asymptotic instantaneous mutual information (5.15) obtained in the non-ergodic regime also represents the asymptotic value of the average mutual information, whose expression is the same as the asymptotic ergodic end-to-end mutual information that would be obtained if the channel was an ergodic process.

It should also be mentioned that, according to the experimental results illustrated in Section 5.6, the system under consideration behaves like in the asymptotic regime even when it is equipped with a reasonable finite number of antennas at each level. Therefore, (5.15) can also be efficiently used to

[2]Recall that $\mathbf{M}_i^H \mathbf{M}_i = \mathbf{C}_{r,i}^{H/2} \mathbf{P}_i^H \mathbf{C}_{t,i+1} \mathbf{P}_i \mathbf{C}_{r,i}^{1/2}$. The power constraints on signals transmitted by the source or relays are not sufficient to guarantee the boundedness of the eigenvalues of $\mathbf{M}_i^H \mathbf{M}_i$. Indeed, as (5.131) shows, in the asymptotic regime the power constraints represent constraints on the product of first-order moment of the eigenvalues of matrices $\mathbf{P}_i \mathbf{C}_{r,i} \mathbf{P}_i^H$ and $\mathbf{M}_k^H \mathbf{M}_k$— indeed $\lim_{k_i \to \infty} \frac{1}{k_i} \mathrm{tr}(\mathbf{P}_i \mathbf{C}_{r,i} \mathbf{P}_i^H) = \mathrm{E}[\lambda_{\mathbf{P}_i \mathbf{C}_{r,i} \mathbf{P}_i^H}]$ and $\lim_{k_k \to \infty} \frac{1}{k_k} \mathrm{tr}(\mathbf{C}_{t,k+1} \mathbf{P}_k \mathbf{C}_{r,k} \mathbf{P}_k^H) = \mathrm{E}[\Lambda_k]$, which a priori does not prevent the eigenvalue distribution of $\mathbf{M}_i^H \mathbf{M}_i$ from having a non-bounded support. Thus, the assumption that matrices $\mathbf{M}_i^H \mathbf{M}_i$ have a compactly supported asymptotic eigenvalue distribution is a priori not an intrinsic property of the system model, and it was necessary to make that assumption to use *Lemma 2* to prove *Theorem 1*.

maximize the instantaneous mutual information of a finite-size system.

5.4 Optimal Transmission Strategy at Source and Relays

In previous section, the asymptotic instantaneous mutual information (5.15), (5.16) was derived considering arbitrary precoding matrices $\mathbf{P}_i, i \in \{0, \ldots, N-1\}$.

In this section, we analyze the optimal linear precoding strategies $\mathbf{P}_i, i \in \{0, \ldots, N-1\}$ at source and relays that allow to maximize the average mutual information. We characterize the optimal transmit directions, meaning the singular vectors of the precoding matrices at source and relays, for a finite-size system i.e. when $k_0, k_1, \ldots, k_N$ are finite. It turns out that those transmit direction are also the ones maximizing the asymptotic average mutual information. As explained in Section 5.3, in the asymptotic regime, the average mutual information and the instantaneous mutual information have the same asymptotic value, therefore the singular vectors of the precoding matrices maximizing the asymptotic average mutual information are also optimal for the asymptotic instantaneous mutual information (5.15).

In future work, using the results on the optimal directions of transmission (singular vectors of $\mathbf{P}_i$) and the asymptotic mutual information (5.15)–(5.16), we intend to work out the optimal power allocation (singular values of $\mathbf{P}_i$) maximizing the asymptotic mutual information (5.15).

The main result of this section is given by the following theorem:

Theorem 2. *Consider the system described in Section 5.2. For $i \in \{1, \ldots, N\}$ let $\mathbf{C}_{t,i} = \mathbf{U}_{t,i}\mathbf{\Lambda}_{t,i}\mathbf{U}_{t,i}^H$ and $\mathbf{C}_{r,i} = \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{U}_{r,i}^H$ be the eigenvalue decompositions of the correlation matrices $\mathbf{C}_{t,i}$ and $\mathbf{C}_{r,i}$, where $\mathbf{U}_{t,i}$ and $\mathbf{U}_{r,i}$ are unitary and $\mathbf{\Lambda}_{t,i}$ and $\mathbf{\Lambda}_{r,i}$ are diagonal, with their respective eigenvalues ordered in decreasing order. Then, under channel-knowledge assumptions $\mathbf{A_s}$, $\mathbf{A_r}$ and $\mathbf{A_d}$, the optimal linear precoding matrices, that maximize the average mutual information under power constraints (5.6) can be written*

$$\begin{aligned} \mathbf{P}_0 &= \mathbf{U}_{t,1}\mathbf{\Lambda}_{P_0} \\ \mathbf{P}_i &= \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{P_i}\mathbf{U}_{r,i}^H \text{ , for } i \in \{1, \ldots, N-1\} \end{aligned} \tag{5.17}$$

where $\mathbf{\Lambda}_{P_i}$ are diagonal matrices with non-negative real diagonal elements. In other words, the singular vectors of the optimal precoding matrices are

given by the eigenvectors of the channel correlation matrices. Moreover, the singular vectors of the precoding matrices (5.17) are also the ones which maximize the asymptotic average mutual information. Since the asymptotic average mutual information has the same value as the asymptotic instantaneous mutual information, the singular vectors of the precoding matrices (5.17) are eventually also optimal for the asymptotic instantaneous mutual information.

For the proof of *Theorem 2*, the reader is referred to Appendix 5.C.

Theorem 2 means that the transmit directions at source and relays maximizing the average mutual information are such that:

- the source should align the eigenvectors of the transmit covariance matrix $\mathbf{Q} = \mathbf{P}_0\mathbf{P}_0^H$ to the eigenvectors of the transmit correlation matrix $\mathbf{C}_{t,1}$ of the first-hop channel $\mathbf{H}_1$, which requires only local statistical channel knowledge $\mathbf{A_s}$. Note that a similar result was previously obtained in the single-hop MIMO system—without relays— with covariance knowledge at the source in [95] for the single-user case, and in [96] for the multi-user case.

- relay i should align the singular vectors of its precoding matrix $\mathbf{P}_i$ to the eigenvectors $\mathbf{U}_{r,i}$ of the receive correlation matrix of channel $\mathbf{H}_i$ on the right, and to the eigenvectors $\mathbf{U}_{t,i+1}$ of the transmit correlation matrix of channel $\mathbf{H}_{i+1}$ on the left, which requires only local statistical knowledge $\mathbf{A_r}$.

- the problem of optimizing $\mathbf{P}_i$ can be divided into two decoupled problems: optimizing the transmit directions—singular vectors— on one hand, and optimizing the transmit powers—singular values— on the other hand.

We would like to draw the attention of the reader on the following point: the proof of this theorem does not rely on the expression of the asymptotic mutual information given in (5.15) and is independent from *Theorem 1*. On the contrary, *Theorem 2* is first proved in the non-asymptotic regime for any system size, i.e. for any set $\{k_i\}_{i\in\{0,\ldots,N\}}$. The so-obtained singular vectors of the precoding matrices maximizing the average mutual information are aligned to the eigenvectors of channel correlation matrices for any system size, thus the result still holds when the system size increases. Consequently,

in the asymptotic regime, the singular vectors of the precoding matrices maximizing the asymptotic average mutual information are also aligned to the eigenvectors of channel correlation matrices as in (5.17). As explained in Section 5.3, in the asymptotic regime, instantaneous and average mutual informations have the same value. Therefore, the singular vectors given in (5.17) are also the ones maximizing the asymptotic instantaneous mutual information. Eventually, by combining *Theorem 1* and *Theorem 2*, the ultimate objective is to find the optimal precoding matrices maximizing the asymptotic instantaneous/average mutual information using only statistical knowledge of the channel at transmitting nodes.

5.5 Application to MIMO Communication Scenarios

In this section, *Theorem* 1 and *Theorem* 2 are applied to four different communication scenarios. In the first two scenarios, the special case of non-relay assisted MIMO (N=1) without path-loss ($a_1 = 1$) is considered, and we show how (5.15) boils down to known results for the MIMO channel with or without correlation. In the third and fourth scenarios, a multi-hop MIMO system is considered and the asymptotic mutual information is developed in the uncorrelated and exponential correlation cases respectively.

5.5.1 Uncorrelated Single-Hop MIMO with Statistical CSI at Source

Consider a simple single-hop uncorrelated MIMO system with the same number of antennas at source and destination i.e. $\rho_0 = \rho_1 = 1$, and an i.i.d. Rayleigh fading channel i.e. $\mathbf{C}_{t,1} = \mathbf{C}_{r,1} = \mathbf{I}$. Assuming equal power allocation at source antennas, the source precoder is $\mathbf{P}_0 = \sqrt{\mathcal{P}_0}\mathbf{I}$. As $\mathbf{M}_0 = \mathbf{C}_{t,1}^{1/2}\mathbf{P}_0 = \sqrt{\mathcal{P}_0}\mathbf{I}$ and $\mathbf{M}_1 = \mathbf{C}_{r,1}^{1/2} = \mathbf{I}$, their empirical eigenvalue distributions are given by

$$\begin{aligned} dF_{\mathbf{M}_0^H\mathbf{M}_0}(\lambda) &= \delta\left(\lambda - \mathcal{P}_0\right) d\lambda \\ dF_{\mathbf{M}_1^H\mathbf{M}_1}(\lambda) &= \delta(\lambda - 1) d\lambda. \end{aligned} \tag{5.18}$$

Using the distributions in (5.18) to compute the expectations in (5.15)

yields

$$\begin{aligned} \boldsymbol{I}_\infty &= \frac{1}{\rho_0}\sum_{i=0}^{N}\rho_i \mathrm{E}\left[\log\left(1+\frac{\eta}{\rho_i}h_i^N\Lambda_i\right)\right] - N\frac{\log e}{\rho_0}\eta\prod_{i=0}^{N}h_i \\ &= \log\left(1+\eta h_0\mathcal{P}_0\right)+\log(1+\eta h_1)-\log e\ \eta\, h_0\, h_1 \end{aligned} \tag{5.19}$$

where, according to (5.16), h_0 and h_1 are the solutions of the system of 2 equations

$$\begin{aligned} h_0 &= \frac{1}{1+\eta h_1} \\ h_1 &= \frac{\mathcal{P}_0}{1+\eta h_0\mathcal{P}_0} \end{aligned} \tag{5.20}$$

that are given by

$$\begin{aligned} h_0 &= \frac{2}{1+\sqrt{1+4\eta\mathcal{P}_0}} \\ h_1 &= \frac{-1+\sqrt{1+4\eta\mathcal{P}_0}}{2\eta}. \end{aligned} \tag{5.21}$$

Inserting the expression of h_0 and h_1 (5.21) into (5.19), we obtain

$$\boldsymbol{I}_\infty = 2\log\left(\frac{1+\sqrt{1+4\eta\mathcal{P}_0}}{2}\right) - \frac{\log e}{4\eta\mathcal{P}_0}\left(\sqrt{1+4\eta\mathcal{P}_0}-1\right)^2. \tag{5.22}$$

It can be observed that the deterministic expression (5.22) depends only on the system characteristics and is independent from the channel realizations. Moreover equal power allocation is known to be the capacity-achieving power allocation for a MIMO i.i.d. Rayleigh channel with statistical CSI at source [39, Section 3.3.2], [38]. As such, the asymptotic mutual information expression obtained in (5.22) also represents the asymptotic capacity of the system. Finally (5.22) is similar to the expression of the asymptotic capacity per dimension previously derived in [39, Section 3.3.2] for the MIMO Rayleigh Channel with equal number of transmit and receive antennas and statistical CSI at the transmitter.

5.5.2 Correlated Single-Hop MIMO with Statistical CSI at Source

In this example, we consider the more general case of correlated MIMO channel with separable correlation: $\mathbf{H}_1 = \mathbf{C}_{r,1}^{1/2}\mathbf{\Theta}_1\mathbf{C}_{t,1}^{1/2}$. The eigenvalue decomposition of $\mathbf{C}_{t,1}$ is

$$\mathbf{C}_{t,1} = \mathbf{U}_{t,1}\mathbf{\Lambda}_{t,1}\mathbf{U}_{t,1}^H \tag{5.23}$$

where $\mathbf{\Lambda}_{t,1}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{C}_{t,1}$ in the non-increasing order and the unitary matrix $\mathbf{U}_{t,1}$ contains the corresponding eigenvectors. Let us define the transmit covariance matrix

$$\mathbf{Q} \triangleq \mathrm{E}\left[\mathbf{x}_0\mathbf{x}_0^H\right] = \mathbf{P}_0\mathbf{P}_0^H \tag{5.24}$$

It has been shown [95] that the capacity-achieving matrix $\mathbf{Q}^\star$ is given by

$$\mathbf{Q}^\star = \mathbf{U}_{t,1}\mathbf{\Lambda}_{\mathbf{Q}^\star}\mathbf{U}_{t,1}^H \tag{5.25}$$

where $\mathbf{\Lambda}_{\mathbf{Q}^\star}$ is a diagonal matrix containing the capacity-achieving power allocation. Using *Theorem 1* along with (5.23) and (5.25), it can be readily shown that the asymptotic capacity per dimension is equal to

$$C = \mathrm{E}[\log(1 + \frac{\eta}{\rho_0}\Lambda_0 h_0)] + \frac{1}{\rho_0}\mathrm{E}[\log(1 + \eta\Lambda_1 h_1)] - \frac{\log e}{\rho_0}\eta\, h_0 h_1 \tag{5.26}$$

where h_0 and h_1 are the solutions of the system

$$\begin{aligned} h_0 &= \mathrm{E}\left[\frac{\Lambda_1}{1 + \eta\Lambda_1 h_1}\right] \\ h_1 &= \mathrm{E}\left[\frac{\Lambda_0}{1 + \frac{\eta}{\rho_0}\Lambda_0 h_0}\right] \end{aligned} \tag{5.27}$$

and the expectations are over Λ_0 and Λ_1 whose distributions are given by the asymptotic eigenvalue distributions of $\mathbf{\Lambda}_{t,1}\mathbf{\Lambda}_{\mathbf{Q}^\star}$ and $\mathbf{C}_r$, respectively. It should be mentioned that an equivalent expression[3] was obtained in [39, Theorem 3.7] for the capacity of the correlated MIMO channel with statistical CSI at transmitter.

[3] The small differences between (5.26) and the capacity expression in [39, Theorem 3.7] are due to different normalization assumptions in [39]. In particular (5.26) is the mutual information per source antenna while the expression in [39] is the capacity per receive antenna. The equivalence between [39, Theorem 3.7] and (5.26) is obtained according to

5.5.3 Uncorrelated Multi-Hop MIMO with Statistical CSI at Source and Relays

In this example, we consider an uncorrelated multi-hop MIMO system, i.e. all correlation matrices are equal to identity. Then by *Theorem 2* the optimal precoding matrices should be diagonal. Assuming equal power allocation at source and relays, the precoding matrices are of the form $\mathbf{P}_i = \alpha_i \mathbf{I}_{k_i}$, where α_i is real positive and chosen to respect the power constraints.

Using the power constraint expression (5.131), it can be shown by induction on i that the coefficients α_i in the uncorrelated case are given by

$$\begin{aligned} \alpha_0 &= \sqrt{\mathcal{P}_0} \\ \alpha_i &= \sqrt{\frac{\mathcal{P}_i}{a_i \mathcal{P}_{i-1}}} \qquad \forall i \in \{1, \ldots, N-1\} \\ \alpha_N &= 1 \end{aligned} \tag{5.29}$$

Then the asymptotic mutual information for the uncorrelated multi-hop MIMO system with equal power allocation is given by

$$\boldsymbol{I}_\infty = \sum_{i=0}^{N} \frac{\rho_i}{\rho_0} \log\left(1 + \frac{\eta h_i^N a_{i+1} \alpha_i^2}{\rho_i}\right) - N \frac{\log e}{\rho_0} \eta \prod_{i=0}^{N} h_i \tag{5.30}$$

where $h_0, h_1, \ldots, h_N$ are the solutions of the system of $N+1$ multivariate polynomial equations

$$\prod_{j=0}^{N} h_j = \frac{h_i^N \alpha_i^2 a_{i+1}}{1 + \frac{\eta h_i^N a_{i+1} \alpha_i^2}{\rho_i}} \qquad i = 0, \ldots, N \tag{5.31}$$

Note that the asymptotic mutual information is a deterministic value depending only on a few system characteristics: signal power $\mathcal{P}_i$, noise power $1/\eta$, pathloss a_i, number of hops N and ratio of the number of antennas ρ_i.

---

the following notation equivalence ({ [39]-notation} $\sim$ {(5.26)-notation}):

$$C \sim \rho_0 \boldsymbol{I}_\infty \qquad \beta \sim \rho_0 \qquad \mathbf{SNR} \sim \mathcal{P}_0 \eta \qquad \Gamma \sim \frac{h_0}{\rho_0} \qquad \Upsilon \sim \frac{h_1}{\mathcal{P}_0}$$

$\Lambda_R \sim \Lambda_1$, both with distribution given by the eigenvalue distribution of $\mathbf{C}_r$

$\Lambda \sim \dfrac{\Lambda_0}{\mathcal{P}_0}$, both with distribution given by the eigenvalue distribution of $\boldsymbol{\Lambda}_{t,1} \boldsymbol{\Lambda}_{\mathbf{Q}^\star} / \mathcal{P}_0$

(5.28)

5.5.4 Exponentially Correlated Multi-Hop MIMO with Statistical CSI at Source and Relays

In this section, the asymptotic mutual information (5.15) is developed in the case of exponential correlation matrices and precoding matrices with optimal singular vectors.

Optimal precoding directions: For $i \in \{1, \ldots, N\}$, the eigenvalue decompositions of channel correlation matrices $\mathbf{C}_{t,i}$ and $\mathbf{C}_{r,i}$ can be written

$$\begin{aligned} \mathbf{C}_{t,i} &= \mathbf{U}_{t,i}\mathbf{\Lambda}_{t,i}\mathbf{U}_{t,i}^H \\ \mathbf{C}_{r,i} &= \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{U}_{r,i}^H \end{aligned} \tag{5.32}$$

where $\mathbf{U}_{t,i}$ and $\mathbf{U}_{r,i}$ are unitary, and $\mathbf{\Lambda}_{t,i}$ and $\mathbf{\Lambda}_{r,i}$ are diagonal with their respective eigenvalues ordered in decreasing order. Following *Theorem 2*, we consider precoding matrices of the form $\mathbf{P}_i = \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{P_i}\mathbf{U}_{r,i}^H$, i.e. the singular vectors of $\mathbf{P}_i$ are optimally aligned to the eigenvectors of channel correlation matrices.

Consequently, we can rewrite matrices $\mathbf{M}_i^H\mathbf{M}_i$ (5.9) as

$$\begin{aligned} \mathbf{M}_0^H\mathbf{M}_0 &= \mathbf{U}_{t,1}^H\mathbf{\Lambda}_{P_0}^2\mathbf{\Lambda}_{t,1}\mathbf{U}_{t,1} \\ \mathbf{M}_i^H\mathbf{M}_i &= \mathbf{U}_{r,i}^H\mathbf{\Lambda}_{r,i}\mathbf{\Lambda}_{P_i}^2\mathbf{\Lambda}_{t,i+1}\mathbf{U}_{r,i} \qquad i = 1, \ldots, N-1 \\ \mathbf{M}_N^H\mathbf{M}_N &= \mathbf{U}_{r,N}^H\mathbf{\Lambda}_{r,N}\mathbf{U}_{r,N} \end{aligned} \tag{5.33}$$

Thus, the eigenvalues of matrices $\mathbf{M}_i^H\mathbf{M}_i$ are contained in the following diagonal matrices

$$\begin{aligned} \mathbf{\Lambda}_0 &= \mathbf{\Lambda}_{P_0}^2\mathbf{\Lambda}_{t,1} \\ \mathbf{\Lambda}_i &= \mathbf{\Lambda}_{r,i}\mathbf{\Lambda}_{P_i}^2\mathbf{\Lambda}_{t,i+1} \qquad i = 1, \ldots, N-1 \\ \mathbf{\Lambda}_N &= \mathbf{\Lambda}_{r,N} \end{aligned} \tag{5.34}$$

The asymptotic mutual information, given by (5.15) and (5.16), involves expectations of functions of Λ_i whose distribution is given by the asymptotic eigenvalue distribution $F_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda)$ of $\mathbf{M}_i^H\mathbf{M}_i$. Equation (5.34) shows that a function $g_1(\Lambda_i)$ can be written as a function $g_2(\Lambda_{P_i}^2,\ \Lambda_{r,i},\ \Lambda_{t,i+1})$, where the variables $\Lambda_{P_i}^2$, $\Lambda_{r,i}$, and $\Lambda_{t,i+1}$ are respectively characterized by the asymptotic eigenvalue distributions $F_{\mathbf{P}_i^H\mathbf{P}_i}(\lambda)$, $F_{\mathbf{C}_{r,i}}(\lambda)$, and $F_{\mathbf{C}_{t,i+1}}(\lambda)$ of matrices $\mathbf{P}_i^H\mathbf{P}_i$, $\mathbf{C}_{r,i}$ and $\mathbf{C}_{t,i+1}$. Therefore expectations in (5.15) and (5.16) can be computed

using the asymptotic joint distribution of $(\Lambda_{P_i}^2,\ \Lambda_{r,i},\ \Lambda_{t,i+1})$ instead of the distribution $F_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda)$. To simplify notations, we rename the variables as follows

$$X = \Lambda_{P_i}^2 \qquad Y = \Lambda_{r,i} \qquad Z = \Lambda_{t,i+1} \tag{5.35}$$

Then, the expectation of a function $g_1(\Lambda_i)$ can be written

$$\begin{aligned}\mathrm{E}[g_1(\Lambda_i)] = \mathrm{E}[g_2(X,Y,Z)] &= \int_z\int_y\int_x g_2(x,y,z) f_{X,Y,Z}(x,y,z)\,dx\,dy\,dz \\ &= \int_z\int_y\int_x g_2(x,y,z) f_{X|Y,Z}(x|y,z)\ \ f_{Y|Z}(y|z)\ \ f_Z(z)\,dx\,dy\,dz\end{aligned} \tag{5.36}$$

Exponential Correlation Model: So far, general correlation matrices were considered. We now introduce the exponential correlation model and further develop (5.36) for the distributions $f_{Y|Z}(y|z)$ and $f_Z(z)$ resulting from that particular correlation model.

We assume that Level i is equipped with a uniform linear array (ULA) of length L_i, characterized by its antenna spacing $l_i = L_i/k_i$ and its characteristic distances $\Delta_{t,i}$ and $\Delta_{r,i}$ proportional to transmit and receive spatial coherences respectively. Then the receive and transmit correlation matrices at Level i can respectively be modeled by the following Hermitian Wiener-class[4] Toeplitz matrices [97–99]:

$$\mathbf{C}_{r,i} = \begin{bmatrix} 1 & r_{r,i} & r_{r,i}^2 & \dots & r_{r,i}^{k_i-1} \\ r_{r,i} & 1 & \ddots & \ddots & \vdots \\ r_{r,i}^2 & \ddots & \ddots & \ddots & r_{r,i}^2 \\ \vdots & \ddots & \ddots & 1 & r_{r,i} \\ r_{r,i}^{k_i-1} & \dots & r_{r,i}^2 & r_{r,i} & 1 \end{bmatrix}_{k_i\times k_i} \tag{5.37}$$

[4] A sequence of $n\times n$ Toeplitz Matrices $\mathbf{T}_n = [t_{k-j}]_{n\times n}$ is said to be in the Wiener class [61, Section 4.4] if the sequence $\{t_k\}$ of first-column and first-row elements is absolutely summable, i.e. $\lim_{n\to+\infty}\sum_{k=-n}^{n}|t_k| < +\infty$.
If $|r_{r,i}| < 1$, then $\lim_{K_i\to+\infty}(\sum_{k=0}^{K_i-1} r_{r,i}^k + \sum_{k=-K_{i-1}}^{-1} r_{r,i}^{-k}) = \frac{1}{1-r_{r,i}} + \frac{1/r_{r,i}}{1-1/r_{r,i}} < \infty$, and consequently $\mathbf{C}_{r,i}$ is in the Wiener class. $\mathbf{C}_{t,i}$ is obviously also in the Wiener class if $|r_{t,i}| < 1$.

$$\mathbf{C}_{t,i+1} = \begin{bmatrix} 1 & r_{t,i+1} & r_{t,i+1}^2 & \dots & r_{t,i+1}^{k_i-1} \\ r_{t,i+1} & 1 & \ddots & \ddots & \vdots \\ r_{t,i+1}^2 & \ddots & \ddots & \ddots & r_{t,i+1}^2 \\ \vdots & \ddots & \ddots & 1 & r_{t,i+1} \\ r_{t,i+1}^{k_i-1} & \dots & r_{t,i+1}^2 & r_{t,i+1} & 1 \end{bmatrix}_{k_i \times k_i} \tag{5.38}$$

where the antenna correlation at receive (resp. transmit) side $r_{r,i} = e^{-\frac{l_i}{\Delta_{r,i}}} \in [0,1)$ (resp. $r_{t,i+1} = e^{-\frac{l_i}{\Delta_{t,i}}} \in [0,1)$) is an exponential function of antenna spacing l_i and characteristic distance $\Delta_{r,i}$ (resp. $\Delta_{t,i}$) at relaying Level i.

As K_i grows large, the sequence of Toeplitz matrices $\mathbf{C}_{r,i}$ of size $K_i \times K_i$ is fully characterized by the continuous real function $f_{r,i}$, defined for $\lambda \in [0, 2\pi)$ by [61, Section 4.1]

$$\begin{aligned} f_{r,i}(\lambda) &= \lim_{K_i \to +\infty} \left(\sum_{k=0}^{K_i-1} r_{r,i}^k e^{jk\lambda} + \sum_{k=-(K_i-1)}^{-1} r_{r,i}^{-k} e^{jk\lambda} \right) \\ &= \frac{1}{1 - r_{r,i} e^{j\lambda}} + \frac{r_{r,i} e^{-j\lambda}}{1 - r_{r,i} e^{-j\lambda}} \\ &= \frac{1 - r_{r,i}^2}{|1 - r_{r,i} e^{j\lambda}|^2} \end{aligned} \tag{5.39}$$

We also denote the essential infimum and supremum of $f_{r,i}$ by $m_{f_{r,i}}$ and $M_{f_{r,i}}$ respectively [61, Section 4.1]. In a similar way, we can define the continuous real function $f_{t,i+1}$ characterizing the sequence of Toeplitz matrices $\mathbf{C}_{t,i+1}$ by replacing $r_{r,i}$ in (5.39) by $r_{t,i+1}$, and we denote by $m_{f_{t,i+1}}$ and $M_{f_{t,i+1}}$ its essential infimum and supremum respectively.

By Szegö Theorem [61, Theorem 9], recalled hereafter in *Lemma 6*, for any real function $g(\cdot)$ (resp. $h(\cdot)$) continuous on $[m_{f_{r,i}}, M_{f_{r,i}}]$ (resp. $[m_{f_{t,i+1}}, M_{f_{t,i+1}}]$), we have

$$\begin{aligned} \int_y g(y) f_Y(y)\, dy &\triangleq \lim_{K_i \to +\infty} \frac{1}{K_i} \sum_{k=1}^{K_i} g\left(\lambda_{\mathbf{C}_{r,i}}(k)\right) = \frac{1}{2\pi} \int_0^{2\pi} g\left(f_{r,i}(\lambda)\right)\, d\lambda \\ \int_z h(z) f_Z(z)\, dz &\triangleq \lim_{K_i \to +\infty} \frac{1}{K_i} \sum_{k=1}^{K_i} h\left(\lambda_{\mathbf{C}_{t,i+1}}(k)\right) = \frac{1}{2\pi} \int_0^{2\pi} h\left(f_{t,i+1}(\nu)\right)\, d\nu \end{aligned} \tag{5.40}$$

Assuming that variables $Y = \Lambda_{r,i}$ and $Z = \Lambda_{t,i+1}$ are independent, and applying Szegö Theorem to (5.36), we can write

$$\begin{aligned}
\mathrm{E}[g_1(\Lambda_i)] &= \int_z \int_y \underbrace{\left(\int_x g_2(x,y,z) f_{X|Y,Z}(x|y,z)\, dx \right)}_{g_3(y,z)} \; f_Y(y)\; f_Z(z)\, dy\, dz \\
&= \int_z \left(\int_y g_3(y,z) f_Y(y)\, dy \right) \; f_Z(z)\, dz \\
&= \int_z \left(\frac{1}{2\pi} \int_{\lambda=0}^{2\pi} g_3\left(f_{r,i}(\lambda), z\right)\; d\lambda \right) \; f_Z(z)\, dz \text{ , by Szegö Theorem (5.40)} \\
&= \frac{1}{2\pi} \int_{\lambda=0}^{2\pi} \left(\int_z g_3\left(f_{r,i}(\lambda), z\right) \; f_Z(z)\, dz \right) \; d\lambda \\
&= \frac{1}{(2\pi)^2} \int_{\lambda=0}^{2\pi} \int_{\nu=0}^{2\pi} g_3\left(f_{r,i}(\lambda), f_{t,i+1}(\nu)\right)\; d\lambda\, d\nu \text{ , by Szegö Theorem (5.40)}
\end{aligned} \tag{5.41}$$

Equal power allocation over optimal precoding directions: We further assume equal power allocation over the optimal directions, i.e. the singular values of $\mathbf{P}_i$ are chosen to be all equal: $\mathbf{\Lambda}_{P_i} = \alpha_i \mathbf{I}_{k_i}$, where α_i is real positive and chosen to respect the power constraint (5.6). Equal power allocation may not be the optimal power allocation scheme, but it is considered in this example for simplicity.

Using the power constraint expression for general correlation models (5.131) and considering precoding matrices $\mathbf{P}_i$ with optimal singular vectors as in *Theorem 2* and equal singular values α_i, i.e. precoding matrices of the form $\mathbf{P}_i = \mathbf{U}_{r,i}^H (\alpha_i \mathbf{I}_{k_i}) \mathbf{U}_{t,i+1}$, we can show by induction on i that the coefficients α_i respecting the power constraints for any correlation model are given by

$$\begin{aligned}
\alpha_0 &= \sqrt{\mathcal{P}_0} \\
\alpha_i &= \sqrt{\frac{\mathcal{P}_i}{a_i \mathcal{P}_{i-1}} \frac{\mathrm{tr}(\mathbf{\Lambda}_{r,i-1})}{\mathrm{tr}(\mathbf{\Lambda}_{r,i})} \frac{k_i}{\mathrm{tr}(\mathbf{\Lambda}_{t,i} \mathbf{\Lambda}_{r,i-1})}} \qquad \forall i \in \{1, \ldots, N-1\} \\
\alpha_N &= 1
\end{aligned} \tag{5.42}$$

Applying the exponential correlation model to (5.42) and making the dimensions of the system grow large, it can be shown that in the asymptotic regime,

the α_i respecting the power constraint for the exponentially correlated system converge to the same value (5.29) as for the uncorrelated system.

Then $X = \Lambda_{P_i}^2 = \alpha_i^2$ is independent from Y and Z, thus $f_{X|Y,Z}(x|y,z) = f_X(x) = \delta(x - \alpha_i^2)$. Consequently,

$$g_3(y,z) = \int_x g_2(x,y,z)\delta(x-\alpha_i^2)\,dx = g_2(\alpha_i^2, y, z) \tag{5.43}$$

and (5.41) becomes

$$\mathrm{E}[g_1(\Lambda_i)] = \frac{1}{(2\pi)^2}\int_{\lambda=0}^{2\pi}\int_{\nu=0}^{2\pi} g_2\left(\alpha_i^2, \frac{1-r_{r,i}^2}{|1-r_{r,i}e^{j\lambda}|^2}, \frac{1-r_{t,i+1}^2}{|1-r_{t,i+1}e^{j\nu}|^2}\right)\,d\lambda\,d\nu \tag{5.44}$$

Asymptotic Mutual Information: Using (5.44) in (5.15) with $g_2(x,y,z) = \log\left(1+\eta\frac{a_{i+1}}{\rho_i}h_i^N xyz\right)$ gives the expression of the asymptotic mutual information

$$\begin{aligned}\boldsymbol{I}_\infty = &\sum_{i=0}^{N}\frac{\rho_i}{\rho_0(2\pi)^2}\int_{\lambda=0}^{2\pi}\int_{\nu=0}^{2\pi}\log\left(1+h_i^N\frac{\eta a_{i+1}\alpha_i^2(1-r_{r,i}^2)(1-r_{t,i+1}^2)}{\rho_i|1-r_{r,i}e^{j\lambda}|^2|1-r_{t,i+1}e^{j\nu}|^2}\right)\,d\lambda\,d\nu\\ &- N\frac{\log e}{\rho_0}\eta\prod_{i=0}^{N}h_i\end{aligned} \tag{5.45}$$

where $h_0, h_1, \ldots, h_N$ are the solutions of the following system of $N+1$ equations, obtained by using (5.44) in (5.16) with $g_2(x,y,z) = \frac{h_i^N\Lambda_i xyz}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N xyz}$

$$\prod_{j=0}^{N}h_j = \frac{\rho_i}{(2\pi)^2}\int_{\lambda=0}^{2\pi}\int_{\nu=0}^{2\pi}\frac{h_i^N a_{i+1}\alpha_i^2(1-r_{r,i}^2)(1-r_{t,i+1}^2)}{\rho_i|1-r_{r,i}e^{j\lambda}|^2|1-r_{t,i+1}e^{j\nu}|^2+\eta h_i^N a_{i+1}\alpha_i^2(1-r_{r,i}^2)(1-r_{t,i+1}^2)}\,d\lambda\,d\nu$$
for $i = 0, \ldots, N$ (5.46)

(with the convention $r_{r,0} = r_{t,N+1} = 0$). Using the changes of variables

$$\begin{aligned} t &= \tan\left(\frac{\lambda}{2}\right)\ ,\ \text{thus}\ \cos(\lambda) = \frac{1-t^2}{1+t^2} \quad\text{and}\quad d\lambda = \frac{2du}{1+t^2}\\ u &= \tan\left(\frac{\nu}{2}\right)\ ,\ \text{thus}\ \cos(\nu) = \frac{1-u^2}{1+u^2} \quad\text{and}\quad d\nu = \frac{2du}{1+u^2}\end{aligned} \tag{5.47}$$

and performing some algebraic manipulations that are skipped for the sake of conciseness, (5.45) and (5.46) can be rewritten

$$\boldsymbol{I}_\infty = \sum_{i=0}^{N} \frac{\rho_i}{\rho_0 \pi^2} \int_{t=-\infty}^{+\infty} \int_{u=-\infty}^{+\infty} \log\left(1 + c_{r,i}c_{t,i+1} \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i} \frac{(1+t^2)}{(c_{r,i}^2 + t^2)} \frac{(1+u^2)}{(c_{t,i+1}^2 + u^2)}\right) \frac{dt}{1+t^2} \frac{du}{1+u^2} - N\frac{\log e}{\rho_0}\eta \prod_{i=0}^{N} h_i \tag{5.48}$$

where $h_0, h_1, \ldots, h_N$ are the solutions of the system of $N+1$ equations

$$\prod_{j=0}^{N} h_j = \frac{2}{\pi} \frac{h_i^N a_{i+1}\alpha_i^2}{\sqrt{c_{r,i}c_{t,i+1} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}} \sqrt{\frac{1}{c_{r,i}c_{t,i+1}} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}}} K(m_i) \tag{5.49}$$

where $K(\cdot)$ is the complete elliptic integral of the first kind [100], and

$$\begin{aligned} c_{r,i} &= \frac{1 - r_{r,i}}{1 + r_{r,i}} \\ c_{t,i+1} &= \frac{1 - r_{t,i+1}}{1 + r_{t,i+1}} \\ m_i &= 1 - \frac{\left(\frac{c_{t,i+1}}{c_{r,i}} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}\right)\left(\frac{c_{r,i}}{c_{t,i+1}} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}\right)}{\left(\frac{1}{c_{r,i}c_{t,i+1}} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}\right)\left(c_{r,i}c_{t,i+1} + \frac{\eta h_i^N a_{i+1}\alpha_i^2}{\rho_i}\right)} \end{aligned} \tag{5.50}$$

Those expression show that only a few relevant parameters affect the performance of this complex system: signal power $\mathcal{P}_i$, noise power $1/\eta$, pathloss a_i, number of hops N, ratio of the number of antennas ρ_i, and correlation ratios $c_{r,i}$ and $c_{t,i}$.

5.6 Numerical Results

In this section, we present numerical results to validate *Theorem 1* and to show that even for a small number of antennas, the behavior of the system is close to the behavior in the asymptotic regime, making *Theorem 1* a useful tool for optimization of finite-size systems as well as large networks.

5.6.1 Uncorrelated Multi-Hop MIMO

The uncorrelated system described in Section 5.5.3.C is first considered.

Fig. 5.2 plots the asymptotic mutual information from *Theorem 1* as well as the instantaneous mutual information obtained for an arbitrary channel realization ('experimental' curves) for a system with 10 antennas at source, destination and each relay level, and one, two or three hops, the case $N = 1$ hop corresponding to a MIMO channel. Fig. 5.3 plots the same type of curves, for a system with 100 antennas at each level. The distance between source and destination d was kept constant when increasing the number of hops N, and $N - 1$ relays were inserted between source and destination with equal spacing $d_i = d/N$ between each relaying level. Equal power allocation, i.e. matrices P_i proportional to the identity matrix, as well as non-correlated channels, i.e. channel correlation matrices all equal to identity, and pathloss exponent $\beta = 2$ were considered in these simulations, whose purpose is mainly to validate the formula in Theorem 1, not to optimize the system. We would like to point out that plotting the experimental curves for different channel realizations gave similar results, and that for the sake of clarity and conciseness, we exhibit the experimental curves only for one realization.

Fig. 5.3 shows the perfect match between the instantaneous mutual information for an arbitrary channel realization and the asymptotic mutual information, validating the asymptotic formula for large dimensions of the network. On the other hand Fig. 5.2 shows that the instantaneous mutual information of a system with a small number of antennas behaves very closely to the asymptotic regime, justifying the usefulness of the asymptotic formula even for optimizing systems with small size.

Finally, Fig. 5.4 plots the asymptotic mutual information for one, two, and three hops, as well as the value of the instantaneous mutual information for random channel realizations when the number of antennas at all levels increases. The concentration of the instantaneous mutual information values around the asymptotic limit when the system size increases shows the convergence of the instantaneous mutual information towards the asymptotic limit as the number of antennas grows large at all levels with the same rate.

5.6.2 One-Sided Exponentially Correlated Multi-Hop MIMO

Based on the model discussed in Section 5.5.4.D, the one-sided exponentially correlated system is considered in this section. In the case of one-sided corre-

lation, e.g. $r_{r,i} = 0$ and $r_{t,i} \geq 0$ for all $i \in \{0, \ldots, N\}$, the asymptotic mutual information (5.51), (5.52) is reduced to

$$\boldsymbol{I}_\infty = \sum_{i=0}^{N} \frac{\rho_i}{\rho_0 \pi} \int_{-\infty}^{+\infty} \log\left(1 + c_{t,i+1} \frac{\eta h_i^N a_{i+1} \alpha_i^2}{\rho_i} \frac{(1+u^2)}{(c_{t,i+1}^2 + u^2)}\right) \frac{du}{1+u^2} - N \frac{\log e}{\rho_0} \eta \prod_{i=0}^{N} h_i \tag{5.51}$$

where $h_0, h_1, \ldots, h_N$ are the solutions of the system of $N+1$ equations

$$\prod_{j=0}^{N} h_j = \frac{h_i^N a_{i+1} \alpha_i^2}{\sqrt{c_{t,i+1} + \frac{\eta h_i^N a_{i+1} \alpha_i^2}{\rho_i}} \sqrt{\frac{1}{c_{t,i+1}} + \frac{\eta h_i^N a_{i+1} \alpha_i^2}{\rho_i}}} \tag{5.52}$$

One-sided correlation was considered to avoid the involved computation of the elliptic integral $K(m_i)$ in the system of equations (5.52), and therefore to simplify simulations.

Fig. 5.5 and 5.6 plot the asymptotic mutual information for 10 and 100 antennas at each level respectively, and one, two or three hops, as well as the instantaneous mutual information obtained for an arbitrary channel realization ('experimental' curves). As in the uncorrelated case, the perfect match of the experimental and asymptotic curves in Fig. 5.6 with 100 antennas validates the asymptotic formula in *Theorem 1* in the presence of correlation. Fig. 5.5 shows that even for a small number of antennas, the system behaves closely to the asymptotic regime in the correlated case.

Finally, Fig. 5.7 plots the instantaneous mutual information for random channel realizations against the size of the system and shows its convergence towards the asymptotic mutual information when the number of antennas increases. Comparing Fig. 5.7 to the corresponding Fig. 5.4 in the uncorrelated case, it appears that convergence towards the asymptotic limit is slower in the correlated case.

5.7 Conclusion

We studied a MIMO multi-hop relay network, in correlated fading, where relays perform linear precoding on their received signal before retransmission. On one hand, using free probability theory, we derived a closed-form expression of the instantaneous end-to-end mutual information in the asymptotic regime when the number of antennas at all levels grows large. This deterministic expression turned out to depend only on channel statistics and not

on particular channel realizations, and to also serve as the asymptotic value of the average end-to-end mutual information. We also showed that multi-hop networks with finite dimensions behave closely to the asymptotic regime, even for a small number of antennas, making the asymptotic mutual information a powerful tool for optimizing the instantaneous mutual information of finite-size systems with only statistical knowledge of the channel.

On the other hand, we showed that, for any system size, and a fortiori in the asymptotic regime, the precoding matrices maximizing the average mutual information have a particular structure: at each level, the singular vectors of the optimal precoding matrix must be aligned to the eigenvectors of the transmit and receive correlation matrices of the backward and forward channels respectively. Thus, the singular vectors of the optimal precoding matrices can be determined with only local statistical channel knowledge at each level.

In the sequel, the analysis will first be extended to the noisy-relay scenario. Then, combining asymptotic mutual information and optimal directions of transmissions, future work will focus on optimizing the power allocations, i.e. the singular values of the optimal precoding matrices.

Future research directions also include the analysis of the cooperative clustering effect: given a total number of antennas k_i at level i, instead of considering that the level consists of a single node equipped with many antennas (k_i), we can consider that a level contains n_i nodes— sources, relays, or destinations— with (k_i/n_i) antennas each. Clustering has a direct impact on the structure of correlation matrices, which become block-diagonal matrices, where blocks represent the correlation between antennas at a node, while antennas at different relays sufficiently separated in space are supposed uncorrelated. In the limit of a relaying level containing k_i relays equipped with a single antenna, we fall back to the case of uncorrelated fading with correlation matrices equal to identity. The optimal size of clusters in correlated fading is expected to depend on the SNR regime. Finally, in the cooperative cluster setting described in Section 5.1.1, a complete analysis would need to take into account the cost of cooperation since communication within a cluster is actually unperfect, and channel estimation errors may occur within a cluster.

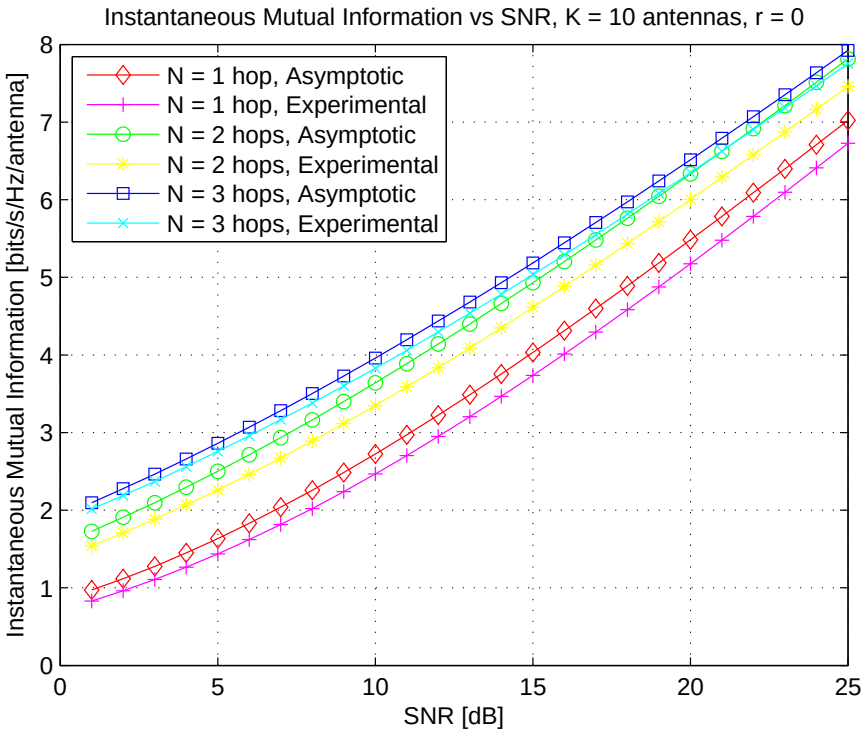


Figure 5.2: Uncorrelated case: Asymptotic Mutual Information and Instantaneous Mutual Information versus SNR, with K = 10 antennas, for single-hop MIMO, 2 hops, and 3 hops

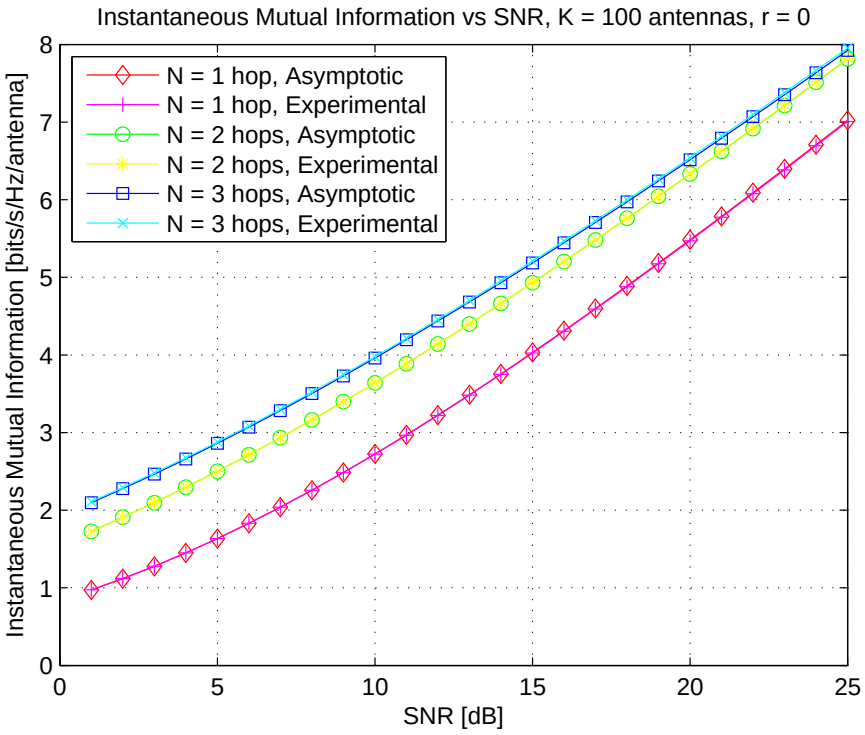


Figure 5.3: Uncorrelated case: Asymptotic Mutual Information and Instantaneous Mutual Information versus SNR, with K = 100 antennas, for single-hop MIMO, 2 hops, and 3 hops

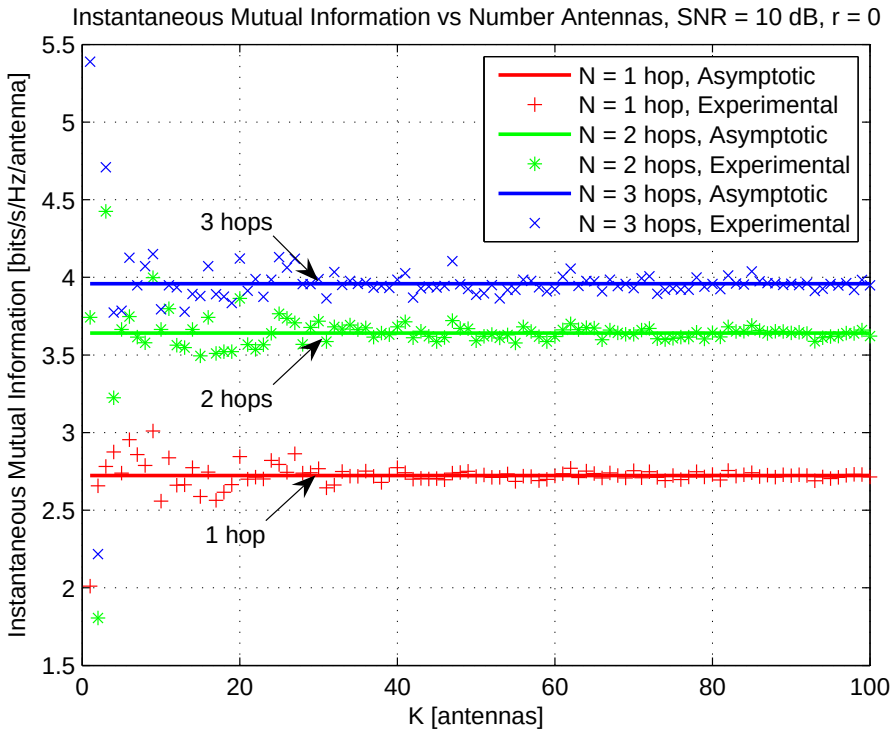


Figure 5.4: Uncorrelated case: Asymptotic Mutual Information and Instantaneous Mutual Information versus K_N, at SNR=10 dB, for single-hop MIMO, 2 hops, and 3 hops

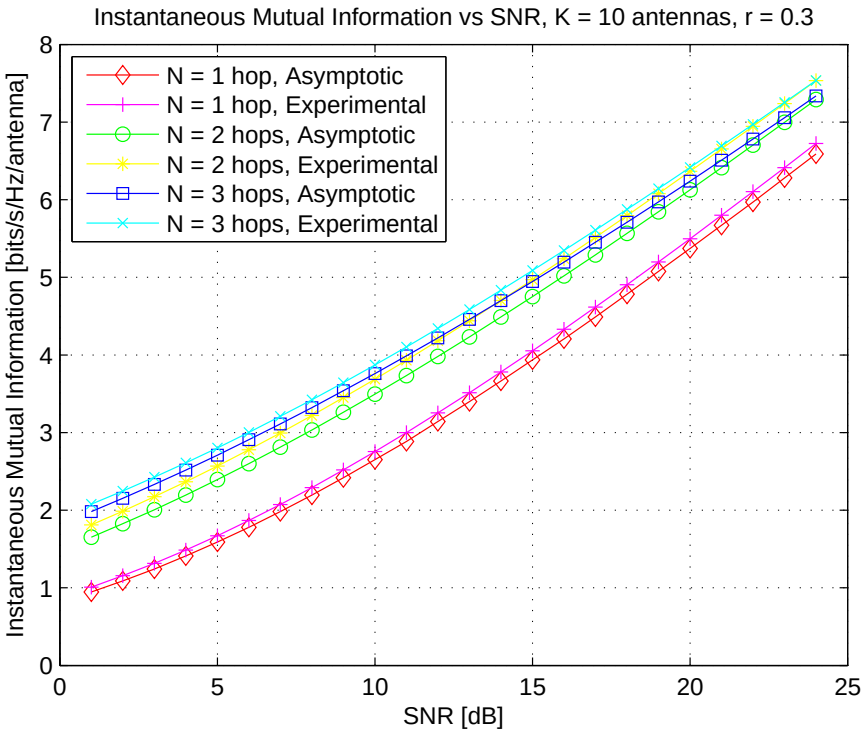


Figure 5.5: One-sided exponential correlation case: Asymptotic Mutual Information and Instantaneous Mutual Information versus SNR, with K = 10 antennas, r=0.3, for single-hop MIMO, 2 hops, and 3 hops

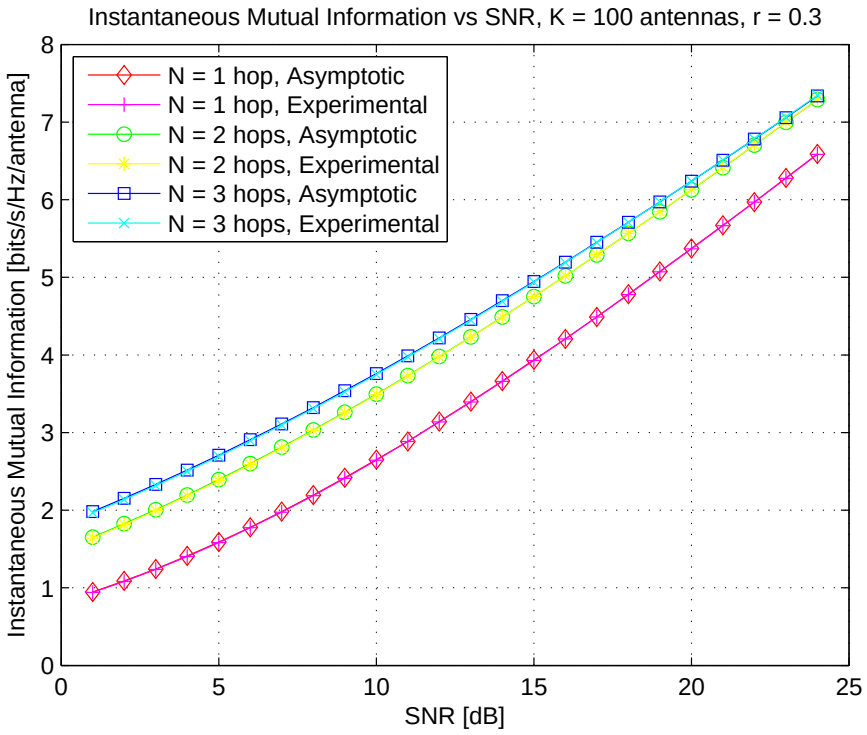


Figure 5.6: One-sided exponential correlation case: Asymptotic Mutual Information and Instantaneous Mutual Information versus SNR, with K = 100 antennas, r=0.3, for single-hop MIMO, 2 hops, and 3 hops

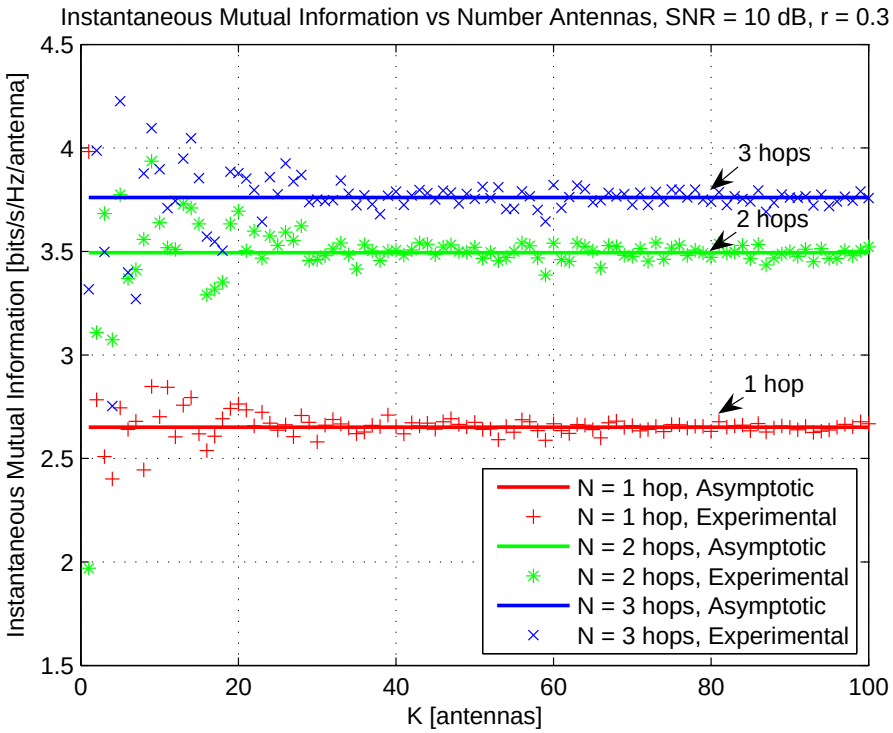


Figure 5.7: One-sided exponential correlation case: Asymptotic Mutual Information and Instantaneous Mutual Information versus K_N, at SNR=10 dB, r=0.3, for single-hop MIMO, 2 hops, and 3 hops

APPENDIX

5.A Useful results from Random Matrix Theory and Free Probability Theory

In this appendix, transforms and lemmas used in the proofs of *Theorems 1* and *2* are provided, while the proofs of *Theorems 1* and *2* are detailed in Appendices 5.B and 5.C, respectively.

5.A.1 Transforms

Let $\mathbf{T}$ be a square matrix of size n with real eigenvalues $\lambda_1, \ldots, \lambda_n$. The empirical eigenvalue distribution $F_{\mathbf{T}}$ of $\mathbf{T}$ is defined by

$$F_{\mathbf{T}}(x) \triangleq \frac{1}{n} \sum_{i=1}^{n} u(x - \lambda_i) \tag{5.53}$$

where $u(x)$ is recalled to be the unit-step function. We define the following transformations [44]

$$\textit{Stieltjes transform}: \quad G_{\mathbf{T}}(s) \triangleq \int \frac{1}{\lambda - s} dF_{\mathbf{T}}(\lambda) \tag{5.54}$$

$$\Upsilon_{\mathbf{T}}(s) \triangleq \int \frac{s\lambda}{1 - s\lambda} dF_{\mathbf{T}}(\lambda) \tag{5.55}$$

$$\textit{S-transform}: \quad S_{\mathbf{T}}(z) \triangleq \frac{z+1}{z} \Upsilon_{\mathbf{T}}^{-1}(z) \tag{5.56}$$

where $\Upsilon^{-1}(\Upsilon(s)) = s$.

5.A.2 Lemmas

We gather here several lemmas used in the proofs of *Theorems 1* and *2*. *Lemmas 1*, *3*, *5* and *7* are proved in Appendix 5.A.3, while *Lemmas 2*, *6*, and *4* come from [101], [61], and [102] respectively.

Lemma 1. *Consider an $n \times p$ matrix $\mathbf{A}$ and a $p \times n$ matrix $\mathbf{B}$, such that their product $\mathbf{AB}$ has non-negative real eigenvalues. Denote $\xi = \frac{p}{n}$. Then*

$$S_{\mathbf{AB}}(z) = \frac{z+1}{z+\xi} S_{\mathbf{BA}}\left(\frac{z}{\xi}\right) \tag{5.57}$$

Note that *Lemma 1* is a generalized form of expressions formerly derived, for instance in [103, Eq. (1.2)], [44, Eq. (15)].

Lemma 2 ([101, Prop. 4.4.9 and 4.4.11])**.** *For $n \in \mathbb{N}$, let $p(n) \in \mathbb{N}$ be such that $\frac{p(n)}{n} \to \xi$ as $n \to \infty$. Let*

- *$\mathbf{\Theta}(n)$ be a $p(n) \times n$ complex Gaussian random matrix with i.i.d. elements with variance $\frac{1}{n}$.*
- *$\mathbf{A}(n)$ be a $n \times n$ constant matrix such that $\sup_n \|\mathbf{A}(n)\| < +\infty$ and $(\mathbf{A}(n), \mathbf{A}(n)^H)$ has the limit eigenvalue distribution μ.*
- *$\mathbf{B}(n)$ be a $p(n) \times p(n)$ Hermitian random matrix, independent from $\mathbf{\Theta}(n)$, with an empirical eigenvalue distribution converging almost surely to a compactly supported probability measure ν.*

Then, as $n \to \infty$,

- *the empirical eigenvalue distribution of $\mathbf{\Theta}(n)^H \mathbf{B}(n) \mathbf{\Theta}(n)$ converges almost surely to the compound free Poisson distribution $\pi_{\nu,\xi}$ [101]*
- *the family $(\{\mathbf{\Theta}(n)^H \mathbf{B}(n) \mathbf{\Theta}(n)\}, \{\mathbf{A}(n), \mathbf{A}(n)^H\})$ is asymptotically free almost everywhere.*

Thus the limiting eigenvalue distribution of $\mathbf{\Theta}(n)\mathbf{B}(n)\mathbf{\Theta}(n)^H \mathbf{A}(n)\mathbf{A}(n)^H$ is the free convolution $\pi_{\nu,\xi} \boxtimes \mu$ and its S-transform *is*

$$S_{\mathbf{\Theta B \Theta}^H \mathbf{A A}^H}(z) = S_{\mathbf{\Theta B \Theta}^H}(z) S_{\mathbf{A A}^H}(z) \tag{5.58}$$

Note that if the elements of $\mathbf{\Theta}(n)$ had variance $\frac{1}{p(n)}$ instead of $\frac{1}{n}$, the conclusion on asymptotic freeness of $(\{\mathbf{\Theta}(n)^H \mathbf{B}(n) \mathbf{\Theta}(n)\}, \{\mathbf{A}(n), \mathbf{A}(n)^H\})$ and equation (5.58) would still hold.

Lemma 3. *Consider an $n \times p$ matrix $\mathbf{A}$ with zero-mean i.i.d. entries with variance $\frac{a}{p}$. Assume that the dimensions go to infinity while $\frac{n}{p} \to \zeta$, then*

$$\begin{aligned} S_{\mathbf{A}\mathbf{A}^H}(z) &= \frac{1}{a}\ \frac{1}{(1+\zeta z)} \\ S_{\mathbf{A}^H\mathbf{A}}(z) &= \frac{1}{a}\ \frac{1}{(z+\zeta)} \end{aligned} \tag{5.59}$$

Lemma 4 ([102, Theorem H.1.h]). *Let $\mathbf{A}$ and $\mathbf{B}$ be two positive semi-definite hermitian matrices of size $n \times n$, $\lambda_{\mathbf{A}}(i)$ and $\lambda_{\mathbf{B}}(i)$ their respective eigenvalues ordered in decreasing order, and $\lambda_{\mathbf{AB}}(i)$ the eigenvalues of $\mathbf{AB}$, then the following inequality holds:*

$$\sum_{i=1}^{n} \lambda_{\mathbf{A}}(i)\lambda_{\mathbf{B}}(n-i+1) \ \leq\ \mathrm{tr}(\mathbf{AB}) = \sum_{i=1}^{n} \lambda_{\mathbf{AB}}(i) \ \leq\ \sum_{i=1}^{n} \lambda_{\mathbf{A}}(i)\lambda_{\mathbf{B}}(i) \tag{5.60}$$

Lemma 5. *For $i \in \{1, \ldots, N\}$, let $\mathbf{A}_i$ be a $n_i \times n_{i-1}$ random matrix. Assume that*

- *$\mathbf{A}_1, \ldots, \mathbf{A}_N$ are mutually independent*
- *n_i goes to infinity while $\frac{n_i}{n_{i-1}} \to \zeta_i$*
- *as n_i goes to infinity, the eigenvalue distribution of $\mathbf{A}_i\mathbf{A}_i^H$ converges almost surely in distribution to a compactly supported measure ν_i,*
- *as $n_1, \ldots, n_N$ go to infinity, the eigenvalue distribution of $(\bigotimes_{i=N}^{1} \mathbf{A}_i)(\bigotimes_{i=N}^{1} \mathbf{A}_i)^H$ converges almost surely in distribution to a measure μ_N*

Then μ_N is compactly supported.

Lemma 6 ([61, Theorem 9]). *Let $\mathbf{T}_n$ be a sequence of Wiener-class Toeplitz matrices, characterized by the function $f(\lambda)$ with essential infimum m_f and essential supremum M_f. Let $\lambda_{T_n}(1), \ldots, \lambda_{T_n}(n)$ be the eigenvalues of $\mathbf{T}_n$ and s be any positive integer. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \lambda_{T_n}^s(k) = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda)^s d\lambda \tag{5.61}$$

Furthermore, if $f(\lambda)$ is real, or equivalently, the matrices $\mathbf{T}_n$ are all Hermitian, then for any function $g(\cdot)$ continuous on $[m_f, M_f]$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g(\lambda_{T_n}(k)) = \frac{1}{2\pi} \int_0^{2\pi} g(f(\lambda)) d\lambda \tag{5.62}$$

Lemma 7. *For $i \geq 1$, given a set of deterministic matrices $\{\mathbf{A}_k\}_{k\in\{0,\ldots,i\}}$ and a set of independent random matrices $\{\mathbf{\Theta}_k\}_{k\in\{1,\ldots,i\}}$, with i.i.d. zero-mean gaussian elements with variance σ_k^2,*

$$\mathrm{tr}\left(\mathrm{E}\left[\bigotimes_{k=i}^{1} \{\mathbf{A}_k \mathbf{\Theta}_k\} \mathbf{A}_0 \mathbf{A}_0^H \bigotimes_{k=1}^{i} \{\mathbf{\Theta}_k^H \mathbf{A}_k^H\} \right] \right) = \mathrm{tr}(\mathbf{A}_0 \mathbf{A}_0^H) \prod_{k=1}^{i} \sigma_k^2 \mathrm{tr}(\mathbf{A}_k \mathbf{A}_k^H) \tag{5.63}$$

5.A.3 Proofs of Lemmas

The proofs of *Lemmas 1*, *3*, *5* and *7* are given hereafter.

Proof of *Lemma 1*

1. Given two complex matrices $\mathbf{A}$ of size $m \times n$, and $\mathbf{B}$ of size $n \times m$ and the characteristic polynomials $\chi_{\mathbf{AB}}(\lambda) \triangleq \det(\lambda \mathbf{I}_m - \mathbf{AB})$ and $\chi_{\mathbf{BA}}(\lambda) \triangleq \det(\lambda \mathbf{I}_n - \mathbf{BA})$, we first show that

$$\forall \lambda \in \mathbb{C} \ , \ \lambda^n \chi_{\mathbf{AB}}(\lambda) = \lambda^m \chi_{\mathbf{BA}}(\lambda) \tag{5.64}$$

For $\lambda = 0$, (5.64) is obviously true.

For $\lambda \neq 0$, we recall the block-determinant formula

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A}) \det(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}) \tag{5.65}$$

$$= \det(\mathbf{D}) \det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) \tag{5.66}$$

and we apply it to the determinant

$$\begin{aligned}
&\det \begin{bmatrix} \mathbf{I}_n & \mathbf{B} \\ \mathbf{A} & \lambda \mathbf{I}_m \end{bmatrix} \\
&= \det(\mathbf{I}_n) \det(\mathbf{I}_m - \mathbf{AB}) = \det(\mathbf{I}_m - \mathbf{AB}) = \chi_{\mathbf{AB}}(\lambda) \\
&= \det(\lambda \mathbf{I}_m) \det(\mathbf{I}_n - \frac{1}{\lambda}\mathbf{BA}) = \lambda^{m-n} \det(\lambda \mathbf{I}_n - \mathbf{BA}) = \lambda^{m-n} \chi_{\mathbf{BA}}(\lambda)
\end{aligned} \tag{5.67}$$

This yields the desired result (5.64)

2. Now we show that **AB** and **BA** have the same non-zero eigenvalues with the same multiplicities.

   Since $\mathbb{C}$ is an algebraically-closed field, the characteristic polynomials are splitted. Denoting by $\lambda_1, \ldots, \lambda_k$ the distinct non-zero roots of $\chi_{\mathbf{AB}}(\lambda)$, by $m_1, ..., m_k > 0$ their multiplicities in $\chi_{\mathbf{AB}}(\lambda)$, and $m_0 \geq 0$ the multiplicity of 0 as root of $\chi_{\mathbf{AB}}(\lambda)$, we can write

$$\chi_{\mathbf{AB}}(\lambda) = \lambda^{m_0} \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i} \tag{5.68}$$

   With similar notations, we can also write

$$\chi_{\mathbf{BA}}(\lambda) = \lambda^{m'_0} \prod_{i=1}^{k'} (\lambda - \lambda'_i)^{m'_i} \tag{5.69}$$

   Using (5.68) and (5.69) in (5.64), we get the polynomial equality

$$\forall \lambda \in \mathbb{C} \ , \quad \lambda^{m_0+n} \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i} = \lambda^{m'_0+m} \prod_{i=1}^{k'} (\lambda - \lambda'_i)^{m'_i} \tag{5.70}$$

   and consequently

$$\begin{aligned}
&m_0 + n = m'_0 + m \\
&k = k' \\
&\lambda_i = \lambda'_i \text{ and } m_i = m'_i \ , \forall i \in \{1, \ldots, k\}
\end{aligned} \tag{5.71}$$

   This means that **AB** and **BA** have the same non-zero eigenvalues with the same multiplicities, and that the multiplicities of their 0-eigenvalues are related.

3. Finally, introducing the assumption that **AB** (and thus **BA**) has real eigenvalues, we show the relation (5.57) on their *S-transforms*.

   Using (5.71), the empirical eigenvalue distributions of **AB** and **BA** are

$$\begin{aligned} F_{\mathbf{AB}}(\lambda) &= \frac{m_0}{m}u(\lambda) + \frac{1}{m}\sum_{i=1}^{k} m_i u(\lambda - \lambda_i) \\ F_{\mathbf{BA}}(\lambda) &= \frac{m'_0}{n}u(\lambda) + \frac{1}{n}\sum_{i=1}^{k} m_i u(\lambda - \lambda_i) \end{aligned} \tag{5.72}$$

   where $u(\lambda)$ is recalled to be the unit-step function. Thus, recalling $m_0 + n = m'_0 + m$, we get

$$F_{\mathbf{AB}}(\lambda) = \frac{n}{m}F_{\mathbf{BA}}(\lambda) + \left(1 - \frac{n}{m}\right)u(\lambda) \tag{5.73}$$

   Using (5.73) to compute their respective *Stieltjes transforms* $G(s) \triangleq \int \frac{1}{\lambda - s} dF(\lambda)$ leads to

$$G_{\mathbf{AB}}(z) = \frac{n}{m}G_{\mathbf{BA}}(z) - \left(1 - \frac{n}{m}\right)\frac{1}{z} \tag{5.74}$$

   Since $\Upsilon(s) = -1 - \frac{1}{s}G(\frac{1}{s})$, from (5.74), we get

$$\Upsilon_{\mathbf{AB}}(s) = \frac{n}{m}\Upsilon_{\mathbf{BA}}(s) \tag{5.75}$$

   and finally, using $\{z = \Upsilon_{\mathbf{AB}}(s) = \frac{n}{m}\Upsilon_{\mathbf{BA}}(s)\} \Leftrightarrow \{\Upsilon^{-1}_{\mathbf{AB}}(z) = s = \Upsilon^{-1}_{\mathbf{BA}}\left(\frac{z}{n/m}\right)\}$ and the definition of the *S-transform* $S(z) \triangleq \frac{z+1}{z}\Upsilon^{-1}(z)$ yields the desired result

$$S_{\mathbf{AB}}(z) = \frac{z+1}{z + \frac{n}{m}} S_{\mathbf{BA}}\left(\frac{z}{n/m}\right) \tag{5.76}$$

   This ends the proof of *Lemma 1*.

□

Proof of *Lemma 3*

Consider an $n \times p$ matrix $\mathbf{A}$ with zero-mean i.i.d. entries with variance $\frac{a}{p}$. It can be written $\mathbf{A} = \sqrt{a}\mathbf{X}$, where $\mathbf{X}$ has zero-mean i.i.d. entries with variance $\frac{1}{p}$.
Define the matrices $\mathbf{Y} = a\mathbf{I}_n$ and $\mathbf{Z} = \mathbf{X}\mathbf{X}^H\mathbf{Y} = \mathbf{A}\mathbf{A}^H$. By noting that the *S-transform* of $\mathbf{Y} = a\mathbf{I}_n$ is $S_{a\mathbf{I}}(z) = \frac{1}{a}$ and applying [44, Theorem 1], we get

$$\begin{aligned} S_{\mathbf{X}\mathbf{X}^H}(z) &= \frac{1}{(1+\zeta z)} \\ S_{\mathbf{A}\mathbf{A}^H}(z) = S_{\mathbf{Z}}(z) &= S_{\mathbf{X}\mathbf{X}^H}(z) S_{\mathbf{Y}}(z) = \frac{1}{(1+\zeta z)} \ \frac{1}{a} \end{aligned} \tag{5.77}$$

Applying *Lemma 1* to $S_{\mathbf{A}^H\mathbf{A}}(z)$ yields

$$S_{\mathbf{A}^H\mathbf{A}}(z) = \frac{z+1}{z+\zeta} S_{\mathbf{A}\mathbf{A}^H}\left(\frac{z}{\zeta}\right) = \frac{1}{a} \ \frac{1}{(z+\zeta)} \tag{5.78}$$

This completes the proof of *Lemma 3*.

□

Proof of *Lemma 5*

The proof of *Lemma 5* is done by induction on N. For $N = 1$, *Lemma 5* obviously holds. Assuming that *Lemma 5* holds for N, we now show that it also holds for $N + 1$.

We first recall that the eigenvalues of Gramian matrices $\mathbf{A}\mathbf{A}^H$ are non-negative. Thus the support of μ_{N+1} is lower-bounded by 0, and we are left with showing that it is also upper-bounded.

Denoting $\mathbf{B}_N = (\bigotimes_{i=N}^{1} \mathbf{A}_i)(\bigotimes_{i=N}^{1} \mathbf{A}_i)^H$, we can write

$$\mathbf{B}_{N+1} = \mathbf{A}_{N+1}\mathbf{B}_N\mathbf{A}_{N+1}^H \tag{5.79}$$

For a matrix $\mathbf{A}$, let $\lambda_{\mathbf{A},\max}$ denote its largest eigenvalue. The largest

eigenvalue of $\mathbf{B}_{N+1}$ is given by

$$\begin{aligned}
\lambda_{\mathbf{B}_{N+1},\max} &= \max_{\mathbf{x}} \ \frac{\mathbf{x}^H \mathbf{B}_{N+1} \mathbf{x}}{\mathbf{x}^H\mathbf{x}} \\
&= \max_{\mathbf{x}} \ \frac{\mathbf{x}^H \mathbf{A}_{N+1}\mathbf{B}_N\mathbf{A}_{N+1}^H \mathbf{x}}{\mathbf{x}^H\mathbf{x}} \\
&= \max_{\mathbf{x}} \ \frac{\mathrm{tr}(\mathbf{B}_N \mathbf{A}_{N+1}^H\mathbf{x}\mathbf{x}^H\mathbf{A}_{N+1})}{\mathbf{x}^H\mathbf{x}} \\
&\leq \max_{\mathbf{x}} \ \frac{\sum_{k=1}^{n_N} \lambda_{\mathbf{B}_N}(k)\, \lambda_{\mathbf{A}_{N+1}^H\mathbf{x}\mathbf{x}^H\mathbf{A}_{N+1}}(k)}{\mathbf{x}^H\mathbf{x}} \text{ , by } \textit{Lemma 4} \\
&\leq \max_{\mathbf{x}} \ \lambda_{\mathbf{B}_N,\max} \ \frac{\sum_{k=1}^{n_N} \lambda_{\mathbf{A}_{N+1}^H\mathbf{x}\mathbf{x}^H\mathbf{A}_{N+1}}(k)}{\mathbf{x}^H\mathbf{x}} \\
&= \lambda_{\mathbf{B}_N,\max} \ \max_{\mathbf{x}} \ \frac{\mathrm{tr}(\mathbf{A}_{N+1}^H\mathbf{x}\mathbf{x}^H\mathbf{A}_{N+1})}{\mathbf{x}^H\mathbf{x}} \\
&= \lambda_{\mathbf{B}_N,\max} \ \max_{\mathbf{x}} \ \frac{\mathbf{x}^H\mathbf{A}_{N+1}\mathbf{A}_{N+1}^H\mathbf{x}}{\mathbf{x}^H\mathbf{x}} \\
&= \lambda_{\mathbf{B}_N,\max} \ \lambda_{\mathbf{A}_{N+1}\mathbf{A}_{N+1}^H,\max}
\end{aligned} \tag{5.80}$$

To simplify notations, we rename the random variables as follows:

$$X = \lambda_{\mathbf{B}_{N+1},\max} \qquad Y = \lambda_{\mathbf{B}_N,\max} \qquad Z = \lambda_{\mathbf{A}_{N+1}\mathbf{A}_{N+1}^H,\max} \tag{5.81}$$

Then (5.80) can be rewritten

$$X \leq YZ \tag{5.82}$$

Let $a \geq 0$, by (5.82) we have

$$F_X(a) = \Pr\{X < a\} \geq \Pr\{YZ < a\} = F_{YZ}(a) \tag{5.83}$$

which still holds for the asymptotic distributions as $n_1, \ldots, n_{N+1}$ go to infinity, while $\frac{n_i}{n_{i-1}} \to \zeta_i$. Denoting the plane region $\mathcal{D}_a = \{x, y \geq 0 / xy < a\}$, we

can write

$$\begin{aligned} F_{YZ}(a) &= \iint_{y,z\in\mathcal{D}_a} f_{Y,Z}(y,z)dydz \\ &= \iint_{y,z\in\mathcal{D}_a} f_Y(y)f_Z(z)dydz \text{ , by independence of } Y \text{ and } Z \\ &= \int_{y=0}^{+\infty}\left(\int_{z=0}^{a/y} f_Z(z)dz\right) f_Y(y)dy \\ &= \int_{y=0}^{+\infty} F_Z\left(\frac{a}{y}\right) f_Y(y)dy \end{aligned} \tag{5.84}$$

By assumption, the distributions of $\mathbf{A}_{N+1}\mathbf{A}_{N+1}^H$ and $\mathbf{B}_N$ converge almost surely to compactly supported measures. Thus, their largest eigenvalues are asymptotically upper-bounded and the support of the asymptotic distributions of Y and Z are upper-bounded, i.e.

$$\begin{aligned} &\exists c_y \geq 0 \text{ such that } \forall y \geq c_y \text{ , } F_Y(y) = 1 \qquad (f_Y(y) = 0) \\ &\exists c_z \geq 0 \text{ such that } \forall z \geq c_z \text{ , } F_Z(z) = 1 \qquad (f_Z(z) = 0) \end{aligned} \tag{5.85}$$

Let $a \geq c_y\, c_z$, then for all $0 < y \leq c_y$, we have $\frac{a}{y} \geq \frac{a}{c_y} \geq c_z$ and $F_Z\left(\frac{a}{y}\right) = 1$, as the dimensions go to infinity with constant rates. Therefore, in the asymptotic regime, we have

$$\begin{aligned} F_{YZ}(a) &= \int_{y=0}^{c_y} F_Z\left(\frac{a}{y}\right) f_Y(y)dy \\ &= \int_{y=0}^{c_y} 1 f_Y(y)dy = F_Y(c_y) = 1 \end{aligned} \tag{5.86}$$

Combining (5.83) and (5.86), we get $F_X(a) = 1$ for $a > c_y\, c_z$. Thus, there exists a constant c_x such that $0 \leq c_x \leq c_y\, c_z$ and $\forall x \geq c_x$, $F_X(x) = 1$, which means that the support of the asymptotic distribution of X is upper-bounded. As a consequence, the support of the asymptotic eigenvalue distribution of $\mathbf{B}_{N+1}$ is also upper-bounded. Therefore, the support of μ_{N+1} is upper-bounded, which concludes the proof.

□

Proof of *Lemma 7*

The proof of *Lemma 7* is done by induction. We first prove that *Lemma 7* holds for $i = 1$. To that purpose, we define the matrix $\mathbf{B} = \mathbf{A}_1\mathbf{\Theta}_1\mathbf{A}_0\mathbf{A}_0^H\mathbf{\Theta}_1^H\mathbf{A}_1^H$. Then

$$\mathrm{tr}(\mathrm{E}[\mathbf{A}_1\mathbf{\Theta}_1\mathbf{A}_0\mathbf{A}_0^H\mathbf{\Theta}_1^H\mathbf{A}_1^H]) = \mathrm{tr}(\mathrm{E}[\mathbf{B}]) = \sum_{j=1}^{k_1}\mathrm{E}[b_{jj}] \tag{5.87}$$

The expectation of the j^{th} diagonal element b_{jj} of matrix $\mathbf{B}$ is

$$\begin{aligned}\mathrm{E}[b_{jj}] &= \sum_{k,l,m,n,p}\mathrm{E}[a_{jk}^{(1)}\theta_{kl}^{(1)}a_{lm}^{(0)}a_{nm}^{(0)*}\theta_{pn}^{(1)*}a_{jp}^{(1)*}] \\ &= \sum_{k,l,m}|a_{jk}^{(1)}|^2|a_{lm}^{(0)}|^2\underbrace{\mathrm{E}[|\theta_{kl}^{(1)}|^2]}_{\sigma_1^2}\ \text{, since } \mathrm{E}[\theta_{kl}^{(1)}\theta_{pn}^{(1)*}] = \sigma_1^2\delta_{k,p}\delta_{l,n} \\ &= \sigma_1^2\sum_k|a_{jk}^{(1)}|^2\sum_{l,m}|a_{lm}^{(0)}|^2\end{aligned} \tag{5.88}$$

Thus (5.87) becomes

$$\mathrm{tr}(\mathrm{E}[\mathbf{B}]) = \sigma_1^2\sum_{j,k}|a_{jk}^{(1)}|^2\sum_{l,m}|a_{lm}^{(0)}|^2 = \sigma_1^2\mathrm{tr}(\mathbf{A}_1\mathbf{A}_1^H)\mathrm{tr}(\mathbf{A}_0\mathbf{A}_0^H) \tag{5.89}$$

which shows that *Lemma 7* holds for $i = 1$.

Now, assuming that *Lemma 7* holds for $i-1$, we show that it also holds for i. We define the matrix $\mathbf{B}_i = \bigotimes_{k=i}^{1}\{\mathbf{A}_k\mathbf{\Theta}_k\}\mathbf{A}_0\mathbf{A}_0^H\bigotimes_{k=1}^{i}\{\mathbf{\Theta}_k^H\mathbf{A}_k^H\}$.

Then

$$\begin{aligned}\mathrm{tr}(\mathrm{E}[\mathbf{B}_i]) &= \mathrm{tr}(\mathrm{E}[\mathbf{A}_i\mathbf{\Theta}_i\mathbf{B}_{i-1}\mathbf{\Theta}_i^H\mathbf{A}_i^H]) \\ &= \sum_{j=1}^{k_1}\mathrm{E}[b_{jj}^{(i)}]\end{aligned} \tag{5.90}$$

The expectation of the j^{th} diagonal element $b_{jj}^{(i)}$ of matrix $\mathbf{B}_i$ is

$$\begin{aligned}\mathrm{E}[b_{jj}^{(i)}] &= \sum_{k,l,m,n}\mathrm{E}[a_{jk}^{(i)}\theta_{kl}^{(i)}b_{lm}^{(i-1)}\theta_{nm}^{(i)*}a_{jn}^{(i)*}] \\ &= \sum_{k,l}|a_{jk}^{(i)}|^2\mathrm{E}[b_{ll}^{(i-1)}]\underbrace{\mathrm{E}[|\theta_{kl}^{(i)}|^2]}_{\sigma_i^2} \\ &= \sigma_i^2\sum_k|a_{jk}^{(i)}|^2\sum_l\mathrm{E}[b_{ll}^{(i-1)}]\end{aligned} \tag{5.91}$$

where the second equality is due to the independence of $\mathbf{\Theta}_i$ and $\mathbf{B}_{i-1}$ and to the fact that $\mathrm{E}[\theta_{kn}^{(i)}\theta_{lm}^{(i)*}] = \sigma_i^2\delta_{k,p}\delta_{l,n}$. Thus (5.90) becomes

$$\begin{aligned}\mathrm{tr}(\mathrm{E}[\mathbf{B}_i]) &= \sigma_i^2 \sum_{j,k} |a_{jk}^{(i)}|^2 \sum_l \mathrm{E}[b_{ll}^{(i-1)}] = \sigma_i^2 \mathrm{tr}(\mathbf{A}_i\mathbf{A}_i^H)\mathrm{tr}(\mathrm{E}[\mathbf{B}_{i-1}]) \\ &= \sigma_i^2 \mathrm{tr}(\mathbf{A}_i\mathbf{A}_i^H)\mathrm{tr}(\mathbf{A}_0\mathbf{A}_0^H) \prod_{k=1}^{i-1} \sigma_k^2 \mathrm{tr}(\mathbf{A}_k\mathbf{A}_k^H) = \mathrm{tr}(\mathbf{A}_0\mathbf{A}_0^H) \prod_{k=1}^{i} \sigma_k^2 \mathrm{tr}(\mathbf{A}_k\mathbf{A}_k^H)\end{aligned} \tag{5.92}$$

which shows that if *Lemma 7* holds for $i-1$, then it holds for i.

Therefore *Lemma 7* holds for any $i \geq 1$, which concludes the proof.

□

5.B Proof of Theorem 1

In this appendix, after listing the main steps of the proof of *Theorem 1*, we provide the detailed proof of each step. Note that the proof of *Theorem 1* uses tools from free probability theory introduced in Appendix 5.A.

The proof of *Theorem 1* goes through four steps as follows:

1. Obtain $S_{\mathbf{G}_N\mathbf{G}_N^H}(z)$.
2. Use $S_{\mathbf{G}_N\mathbf{G}_N^H}(z)$ to find $\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(z)$.
3. Use $\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(z)$ to obtain $d\boldsymbol{I}/d\eta$.
4. Integrate $d\boldsymbol{I}/d\eta$ to obtain $\boldsymbol{I}$ itself.

- **First Step: obtain** $S_{\mathbf{G}_N\mathbf{G}_N^H}(z)$

Theorem 3. *As $k_i, i = 0, \ldots, N$ go to infinity with the same rate, the S-transform of $\mathbf{G}_N\mathbf{G}_N^H$ is given by*

$$S_{\mathbf{G}_N\mathbf{G}_N^H}(z) = S_{\mathbf{M}_N^H\mathbf{M}_N}(z) \prod_{i=1}^{N} \frac{\rho_{i-1}}{a_i} \frac{1}{(z+\rho_{i-1})} S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{z}{\rho_{i-1}}\right) \tag{5.93}$$

Proof. The proof is done by induction using *Lemmas 1, 3, 2.*

First, we prove (5.93) for $N = 1$. Note that

$$\mathbf{G}_1\mathbf{G}_1^H = \mathbf{M}_1\mathbf{\Theta}_1\mathbf{M}_0\mathbf{M}_0^H\mathbf{\Theta}_1^H\mathbf{M}_1^H \tag{5.94}$$

therefore

$$\begin{aligned}
S_{\mathbf{G}_1\mathbf{G}_1^H}(z) &= S_{\mathbf{\Theta}_1\mathbf{M}_0\mathbf{M}_0^H\mathbf{\Theta}_1^H\mathbf{M}_1^H\mathbf{M}_1}(z) && \text{, by } \textit{Lemma 1}\\
&= S_{\mathbf{\Theta}_1\mathbf{M}_0\mathbf{M}_0^H\mathbf{\Theta}_1^H}(z)S_{\mathbf{M}_1^H\mathbf{M}_1}(z) && \text{, by } \textit{Lemma 2}\\
&= \tfrac{z+1}{z+\frac{k_0}{k_1}}S_{\mathbf{M}_0\mathbf{M}_0^H\mathbf{\Theta}_1^H\mathbf{\Theta}_1}\left(\tfrac{z}{\frac{k_0}{k_1}}\right)S_{\mathbf{M}_1^H\mathbf{M}_1}(z) && \text{, by } \textit{Lemma 1}\\
&= \tfrac{z+1}{z+\frac{k_0}{k_1}}S_{\mathbf{M}_0\mathbf{M}_0^H}\left(\tfrac{z}{\frac{k_0}{k_1}}\right)S_{\mathbf{\Theta}_1^H\mathbf{\Theta}_1}\left(\tfrac{z}{\frac{k_0}{k_1}}\right)S_{\mathbf{M}_1^H\mathbf{M}_1}(z) && \text{, by } \textit{Lemma 2}\\
&= \tfrac{z+1}{z+\frac{k_0}{k_1}}S_{\mathbf{M}_0\mathbf{M}_0^H}\left(\tfrac{z}{\frac{k_0}{k_1}}\right)\tfrac{1}{a_1}\ \tfrac{1}{\frac{z}{\frac{k_0}{k_1}}+\frac{k_1}{k_0}}S_{\mathbf{M}_1^H\mathbf{M}_1}(z) && \text{, by } \textit{Lemma 3}\\
&= S_{\mathbf{M}_1^H\mathbf{M}_1}(z)\ \tfrac{\rho_0}{a_1}\ \tfrac{1}{z+\rho_0}S_{\mathbf{M}_0^H\mathbf{M}_0}\left(\tfrac{z}{\rho_0}\right) && \text{, by } \textit{Lemma 5}
\end{aligned} \tag{5.95}$$

Now, we need to prove that if (5.93) holds for $N = q$, it also holds for $N = q + 1$. Note that

$$\mathbf{G}_{q+1}\mathbf{G}_{q+1}^H = \mathbf{M}_{q+1}\mathbf{\Theta}_{q+1}\mathbf{M}_q\mathbf{\Theta}_q\ldots\mathbf{M}_1\mathbf{\Theta}_1\mathbf{M}_0\mathbf{M}_0^H\mathbf{\Theta}_1^H\mathbf{M}_1^H\ldots\mathbf{\Theta}_q^H\mathbf{M}_q^H\mathbf{\Theta}_{q+1}^H\mathbf{M}_{q+1}^H \tag{5.96}$$

therefore,

$$\begin{aligned}
S_{\mathbf{G}_{q+1}\mathbf{G}_{q+1}^H}(z) &= S_{\mathbf{M}_{q+1}\ldots\mathbf{M}_{q+1}^H}(z)\\
&= S_{\mathbf{\Theta}_{q+1}\mathbf{M}_q\ldots\mathbf{M}_q^H\mathbf{\Theta}_{q+1}^H\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z)\ \text{, by } \textit{Lemma 1}
\end{aligned} \tag{5.97}$$

The empirical eigenvalue distribution of Wishart matrices $\mathbf{\Theta}_i\mathbf{\Theta}_i^H$ converges almost surely to the Marčenko-Pastur law whose support is compact, and by assumption, for $i \in \{0, \ldots, N+1\}$ the empirical eigenvalue distribution of $\mathbf{M}_i^H\mathbf{M}_i$ converges to an asymptotic distribution with a compact support. Thus, by *Lemma 5*, the asymptotic eigenvalue distribution of $\mathbf{M}_q\mathbf{\Theta}_q\ldots\mathbf{\Theta}_q^H\mathbf{M}_q^H$ has a compact support. Therefore *Lemma 2* can be applied to (5.97), leading to:

$$
\begin{aligned}
&S_{\mathbf{G}_{q+1}\mathbf{G}_{q+1}^H}(z) = S_{\mathbf{\Theta}_{q+1}\ldots\mathbf{\Theta}_{q+1}^H}(z) S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \text{ , by } \textit{Lemma 2} \\
&= \frac{z+1}{z+\frac{k_q}{k_{q+1}}} S_{\mathbf{M}_q\ldots\mathbf{M}_q^H\mathbf{\Theta}_{q+1}^H\mathbf{\Theta}_{q+1}}\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right) S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \text{ , by } \textit{Lemma 1} \\
&= \frac{z+1}{z+\frac{k_q}{k_{q+1}}} S_{\mathbf{M}_q\ldots\mathbf{M}_q^H}\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right) S_{\mathbf{\Theta}_{q+1}^H\mathbf{\Theta}_{q+1}}\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right) S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \text{ , by } \textit{Lemma 2} \\
&= \frac{z+1}{z+\frac{k_q}{k_{q+1}}} \left(S_{\mathbf{M}_q^H\mathbf{M}_q}\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right) \prod_{i=1}^{q} \frac{\frac{k_{i-1}}{k_q}}{a_i} \frac{1}{\frac{z}{\frac{k_q}{k_{q+1}}}+\frac{k_{i-1}}{k_q}} S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right)}{\frac{k_{i-1}}{k_q}}\right)\right) \times \\
&\quad \frac{1}{a_{q+1}} \frac{1}{\frac{k_{q+1}}{k_q}+\frac{z}{\frac{k_q}{k_{q+1}}}} S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \text{ , by } \textit{Lemma 3} \\
&= \frac{z+1}{z+\frac{k_q}{k_{q+1}}} S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \frac{\frac{k_q}{k_{q+1}}}{a_{q+1}} \frac{1}{z+1} S_{\mathbf{M}_q^H\mathbf{M}_q}\left(\frac{z}{\frac{k_q}{k_{q+1}}}\right) \times \\
&\quad \prod_{i=1}^{q} \frac{\frac{k_{i-1}}{k_{q+1}}}{a_i} \frac{1}{z+\frac{k_{i-1}}{k_{q+1}}} S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{z}{\frac{k_{i-1}}{k_{q+1}}}\right) \\
&= S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \prod_{i=1}^{q+1} \frac{\frac{k_{i-1}}{k_{q+1}}}{a_i} \frac{1}{z+\frac{k_{i-1}}{k_{q+1}}} S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{z}{\frac{k_{i-1}}{k_{q+1}}}\right) \\
&= S_{\mathbf{M}_{q+1}^H\mathbf{M}_{q+1}}(z) \prod_{i=1}^{q+1} \frac{\rho_{i-1}}{a_i} \frac{1}{(z+\rho_{i-1})} S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{z}{\rho_{i-1}}\right). \qquad (5.98)
\end{aligned}
$$

The proof is complete. □

- **Second Step: use $S_{\mathbf{G}_N\mathbf{G}_N^H}(z)$ to find $\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(z)$**

Theorem 4. *Let us define $a_{N+1} = 1$. We have*

$$
s\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}^N(s) = \prod_{i=0}^{N} \frac{\rho_i}{a_{i+1}} \Upsilon_{M_i^H M_i}^{-1}\left(\frac{\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H(s)}}{\rho_i}\right) \qquad (5.99)
$$

Proof. From (5.93) it follows that

$$\frac{z}{z+1}S_{\mathbf{G}_N\mathbf{G}_N^H}(z) = \frac{z}{z+1}S_{\mathbf{M}_N^H\mathbf{M}_N}(z)\prod_{i=1}^{N}\ \frac{\rho_{i-1}}{a_i}\ \frac{1}{z+\rho_{i-1}}\ \frac{\frac{z}{\rho_{i-1}}+1}{\frac{z}{\rho_{i-1}}}\ \frac{\frac{z}{\rho_{i-1}}}{\frac{z}{\rho_{i-1}}+1}\ S_{\mathbf{M}_{i-1}^H\mathbf{M}_{i-1}}\left(\frac{z}{\rho_{i-1}}\right). \tag{5.100}$$

Using (5.56) in (5.100), we obtain

$$\Upsilon^{-1}_{\mathbf{G}_N\mathbf{G}_N^H}(z) = \frac{1}{z^N}\Upsilon^{-1}_{M_N^H M_N}(z)\prod_{i=1}^{N}\ \frac{\rho_{i-1}}{a_i}\ \Upsilon^{-1}_{M_{i-1}^H M_{i-1}}\left(\frac{z}{\rho_{i-1}}\right) \tag{5.101}$$

or, equivalently,

$$\Upsilon^{-1}_{\mathbf{G}_N\mathbf{G}_N^H}(z) = \frac{1}{z^N}\prod_{i=0}^{N}\ \frac{\rho_i}{a_{i+1}}\ \Upsilon^{-1}_{M_i^H M_i}\left(\frac{z}{\rho_i}\right) \tag{5.102}$$

with the convention $a_{N+1} = 1$. Substituting $z = \Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(s)$ in (5.102), Equation (5.99) follows. This completes the proof. □

- **Third Step: use $\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(z)$ to obtain $d\boldsymbol{I}/d\eta$**

Theorem 5. *In the asymptotic regime, as $k_0, k_1, \ldots, k_N$ go to infinity while $\frac{k_i}{k_N} \to \rho_i, i = 0, \ldots, N$, the derivative of the instantaneous mutual information is given by*

$$\frac{d\boldsymbol{I}_\infty}{d\eta} = \frac{1}{\rho_0 \ln 2}\prod_{i=0}^{N} h_i \tag{5.103}$$

where $h_0, h_1, \ldots, h_N$ are the solutions to the following set of $N+1$ equations

$$\prod_{j=0}^{N} h_j = \rho_i \mathrm{E}\left[\frac{h_i^N \Lambda_i}{\frac{\rho_i}{a_i+1} + \eta h_i^N \Lambda_i}\right] \qquad i = 0, \ldots, N. \tag{5.104}$$

The expectation in (5.104) is over Λ_i whose probability distribution function is given by $F_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda)$ (convention: $a_{N+1} = 1$).

Proof. First, we note that

$$\begin{aligned}
\boldsymbol{I} &= \frac{1}{k_0}\log\det(\mathbf{I}+\eta\mathbf{G}_N\mathbf{G}_N^H) \\
&= \frac{1}{k_0}\log\prod_{i=1}^{k_N}(1+\eta\lambda_i(\mathbf{G}_N\mathbf{G}_N^H)) \\
&= \frac{1}{k_0}\sum_{i=1}^{k_N}\log(1+\eta\lambda_i(\mathbf{G}_N\mathbf{G}_N^H)) \\
&= \frac{k_N}{k_0}\frac{1}{k_N}\sum_{i=1}^{k_N}\log(1+\eta\lambda_i(\mathbf{G}_N\mathbf{G}_N^H)) \\
&= \frac{k_N}{k_0}\int\log(1+\eta\lambda)dF^{k_N}_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda) \\
&\overset{a.s.}{\to} \frac{1}{\rho_0}\int\log(1+\eta\lambda)dF_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda) \\
&= \frac{1}{\rho_0\ln 2}\int\ln(1+\eta\lambda)dF_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda) \qquad (5.105)
\end{aligned}$$

where $F^{k_N}_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda)$ is the (non-asymptotic) empirical eigenvalue distribution of $\mathbf{G}_N\mathbf{G}_N^H$, that converges almost-surely to the asymptotic empirical eigenvalue distribution $F_{\mathbf{G}_N\mathbf{G}_N^H}$, whose support is compact. Indeed, the empirical eigenvalue distribution of Wishart matrices $\mathbf{\Theta}_i\mathbf{\Theta}_i^H$ converges almost surely to the Marčenko-Pastur law whose support is compact, and by assumption, for $i\in\{0,\ldots,N+1\}$ the empirical eigenvalue distribution of $\mathbf{M}_i^H\mathbf{M}_i$ converges to an asymptotic distribution with a compact support, thus by *Lemma 5*, the asymptotic eigenvalue distribution of $\mathbf{G}_N\mathbf{G}_N^H$ has a compact support. The function log is a continuous function, thus bounded on the compact support of the asymptotic eigenvalue distribution of $\mathbf{G}_N\mathbf{G}_N^H$, which enabled the application of the bounded convergence theorem to obtain the almost-sure convergence in (5.105).

Due to (5.105) in the asymptotic regime, the derivative of the mutual

information with respect to η is linked to $\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(z)$:

$$\begin{aligned}\frac{d\boldsymbol{I}_\infty}{d\eta} &= \frac{1}{\rho_0 \ln 2}\int \frac{\lambda}{1+\eta\lambda} dF_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda) \\ &= \frac{1}{-\rho_0\eta \ln 2}\int \frac{-\eta\lambda}{1-(-\eta)\lambda} dF_{\mathbf{G}_N\mathbf{G}_N^H}(\lambda) \\ &= \frac{1}{-\rho_0\eta \ln 2}\Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(-\eta). \end{aligned} \tag{5.106}$$

Let us denote

$$t = \Upsilon_{\mathbf{G}_N\mathbf{G}_N^H}(-\eta) \tag{5.107}$$

$$g_i = \Upsilon^{-1}_{\mathbf{M}_i^H\mathbf{M}_i}\left(\frac{t}{\rho_i}\right) \qquad i = 0, \ldots, N \tag{5.108}$$

and, for the sake of simplicity, let $\alpha = \rho_0 \ln 2$. From (5.106), we have

$$t = -\eta\alpha\frac{d\boldsymbol{I}_\infty}{d\eta} \tag{5.109}$$

Substituting $s = -\eta$ in (5.99) and using (5.107) and (5.108), it follows that

$$-\eta t^N = \prod_{i=0}^{N} \frac{\rho_i}{a_{i+1}}\ g_i. \tag{5.110}$$

Finally, from (5.108) and the very definition of Υ in (5.55), we obtain

$$t = \rho_i \int \frac{g_i\lambda}{1-g_i\lambda} dF_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda) \qquad i = 0, \ldots, N. \tag{5.111}$$

Substituting (5.109) in (5.110) and (5.111) yields

$$(-\eta)^{N+1}\left(\alpha\frac{d\boldsymbol{I}}{d\eta}\right)^N = \prod_{i=0}^{N} \frac{\rho_i}{a_{i+1}}\ g_i. \tag{5.112}$$

and

$$-\eta\left(\alpha\frac{d\boldsymbol{I}_\infty}{d\eta}\right) = \rho_i \int \frac{g_i\lambda}{1-g_i\lambda} dF_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda) \qquad i = 0, \ldots, N. \tag{5.113}$$

Letting

$$h_i \;=\; \left(\frac{\rho_i}{a_{i+1}}\right)^{\frac{1}{N}} \left(\frac{g_i}{-\eta}\right)^{\frac{1}{N}} \tag{5.114}$$

it follows from (5.112) that

$$\alpha\frac{d\boldsymbol{I}_\infty}{d\eta} = \prod_{i=0}^{N} h_i. \tag{5.115}$$

Using (5.114) and (5.115) in (5.113), we obtain

$$-\eta\prod_{j=0}^{N} h_j = \rho_i \int \frac{-\eta h_i^N \,\frac{a_{i+1}}{\rho_i}\,\lambda}{1-(-\eta)h_i^N\frac{a_{i+1}}{\rho_i}\lambda} dF_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda) \qquad i=0,\ldots,N \tag{5.116}$$

or, equivalently,

$$\begin{aligned}\prod_{j=0}^{N} h_j &= \rho_i \int \frac{h_i^N\lambda}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N\lambda} dF_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda) \\ &= \rho_i \mathrm{E}\left[\frac{h_i^N\Lambda_i}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N\Lambda_i}\right] \qquad i=0,\ldots,N. \end{aligned} \tag{5.117}$$

This, along with equation (5.115), complete the proof. □

- **Fourth Step: integrate $d\boldsymbol{I}/d\eta$ to obtain $\boldsymbol{I}$ itself**

The fourth step leads to our main *Theorem 1* that we briefly recall hereunder.

Theorem 6. *For the system described in section 5.2, under the assumptions listed in* Theorem 1*, as $k_0, k_1, \ldots, k_N$ go to infinity while $\frac{k_i}{k_N} \to \rho_i, i = 0,\ldots,N$ the end-to-end instantaneous mutual information per transmit antenna $\boldsymbol{I}$ converges almost surely to*

$$\boldsymbol{I}_\infty = \frac{1}{\rho_0}\sum_{i=0}^{N}\rho_i \mathrm{E}\left[\log\left(1+\eta\frac{a_{i+1}}{\rho_i}h_i^N\Lambda_i\right)\right] - N\frac{\log e}{\rho_0}\eta\prod_{i=0}^{N}h_i \tag{5.118}$$

where $h_0, h_1, \ldots, h_N$ are the solutions of the system of $N+1$ equations

$$\prod_{j=0}^{N} h_j = \rho_i \mathrm{E}\left[\frac{h_i^N\Lambda_i}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N\Lambda_i}\right] \qquad i=0,\ldots,N \tag{5.119}$$

and where E *is over Λ_i with probability distribution given by $F_{\mathbf{M}_i^H\mathbf{M}_i}(\lambda)$.*

Proof. Proof is accomplished by computing the derivative of $\boldsymbol{I}_\infty$ in (5.15) in terms of η and showing that the derivative matches with (5.103). This shows that (5.15) is one primitive function of $\frac{d\boldsymbol{I}_\infty}{d\eta}$. Since primitive functions of $\frac{d\boldsymbol{I}_\infty}{d\eta}$ differ by a constant, the constant was chosen such that the mutual information (5.15) is null when SNR η goes to zero: $\lim_{\eta\to 0}\boldsymbol{I}_\infty(\eta)=0$.

We now proceed with computing the derivative of $\boldsymbol{I}_\infty$. If (5.15) holds, then we have (recall $\alpha=\rho_0\ln 2$)

$$\alpha\boldsymbol{I}_\infty=\sum_{i=0}^{N}\rho_i\mathrm{E}\left[\ln\left(1+\frac{\eta a_{i+1}}{\rho_i}h_i^N\Lambda_i\right)\right]-N\eta\prod_{i=0}^{N}h_i. \tag{5.120}$$

From (5.120) we have

$$\begin{aligned}
\alpha\frac{d\boldsymbol{I}_\infty}{d\eta}&=\sum_{i=0}^{N}\rho_i\mathrm{E}\left[\frac{\Lambda_i\left(h_i^N+N\eta h_i^{N-1}h_i'\right)}{\frac{\rho_i}{a_{i+1}}(1+\frac{\eta a_{i+1}}{\rho_i}h_i^N\Lambda_i)}\right]-N\prod_{i=0}^{N}h_i-N\eta\left(\sum_{i=0}^{N}h_i'\prod_{\substack{j=0\\j\neq i}}^{N}h_j\right)\\
&=\sum_{i=0}^{N}\rho_i\mathrm{E}\left[\frac{\Lambda_i h_i^N}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N\Lambda_i}\right]+N\eta\sum_{i=0}^{N}\frac{h_i'}{h_i}\rho_i\mathrm{E}\left[\frac{\Lambda_i h_i^N}{\frac{\rho_i}{a_{i+1}}+\eta h_i^N\Lambda_i}\right]\\
&\quad-N\prod_{i=0}^{N}h_i-N\eta\left(\sum_{i=0}^{N}\frac{h_i'}{h_i}\prod_{j=0}^{N}h_j\right)\\
&=\sum_{i=0}^{N}\prod_{j=0}^{N}h_j+N\eta\left(\sum_{i=0}^{N}\frac{h_i'}{h_i}\prod_{j=0}^{N}h_j\right)-N\prod_{i=0}^{N}h_i-N\eta\left(\sum_{i=0}^{N}\frac{h_i'}{h_i}\prod_{j=0}^{N}h_j\right)\\
&=(N+1)\prod_{j=0}^{N}h_j-N\prod_{j=0}^{N}h_j\\
&=\prod_{j=0}^{N}h_j
\end{aligned}\tag{5.121}$$

where $h_i'\triangleq\frac{dh_i}{d\eta}$ and the third line is due to (5.16). Equation (5.103) immediately follows from (5.121). This completes the proof. □

5.C Proof of Theorem 2

In this appendix, we provide the proof of *Theorem 2*. As in the proofs in [95] for the average mutual information in the single-user MIMO case with covariance knowledge at source, or in [96] for the average mutual information in the multi-user MIMO case also with covariance knowledge at source, both without relaying, or in [94] for the MIMO two-hop relay system with full CSI at the relay, our proof of Theorem 2 is based on [102, Theorem H.1.h] recalled in *Lemma 4*. We extend and develop those proofs here to suit the MIMO multi-hop relaying model described in Fig. 5.1.

The proof for our system goes through three steps that we detail hereafter:

- **Step 1:** Use the singular value decomposition (SVD) $\mathbf{U}_i\mathbf{D}_i\mathbf{V}_i^H = \mathbf{\Lambda}_{t,i+1}^{1/2}\mathbf{U}_{t,i+1}^H\mathbf{P}_i\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}^{1/2}$ and show that unitary matrices $\mathbf{U}_i$ and $\mathbf{V}_i$ impact the maximization of the average mutual information through the power constraints only, whereas diagonal matrices $\mathbf{D}_i$ play a role both in mutual information expression and in power constraints.

- **Step 2:** Give the expression of the power constraints in function of $\mathbf{D}_i, \mathbf{U}_i, \mathbf{V}_i$ and channel correlation matrices only.

- **Step 3:** Show that the directions minimizing the trace in the power constraint for all choices of singular values of $\mathbf{P}_i$, are the directions given in *Theorem 2*.

Before detailing each step, we recall that the maximum average mutual information is given by

$$\boldsymbol{C} \triangleq \max_{\{\mathbf{P}_i/\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H])\leq k_i\mathcal{P}_i\}_{i\in\{0,\ldots,N-1\}}} \mathrm{E}\left[\log\det(\mathbf{I}_{k_N} + \eta\ \mathbf{G}_N\mathbf{G}_N^H)\right] \tag{5.122}$$

- **Step 1: clarify how the average mutual information depends on the transmit directions and the transmit powers**

For $i \in \{1, \ldots, N\}$ we define

$$\mathbf{\Theta}_i' = \mathbf{U}_{r,i}^H\mathbf{\Theta}_i\mathbf{U}_{t,i} \tag{5.123}$$

Since $\mathbf{\Theta}_i$ is zero-mean i.i.d. complex Gaussian, thus unitarily invariant, and $\mathbf{U}_{r,i}$ and $\mathbf{U}_{t,i}$ are unitary matrices, $\mathbf{\Theta}_i'$ has the same distribution as $\mathbf{\Theta}_i$.

For $i \in \{0, \ldots, N-1\}$, we consider the following singular value decomposition (SVD)

$$\mathbf{U}_i \mathbf{D}_i \mathbf{V}_i^H = \mathbf{\Lambda}_{t,i+1}^{1/2} \mathbf{U}_{t,i+1}^H \mathbf{P}_i \mathbf{U}_{r,i} \mathbf{\Lambda}_{r,i}^{1/2} \tag{5.124}$$

where $\mathbf{U}_i$, $\mathbf{V}_i$ are unitary matrices, $\mathbf{D}_i$ is a real diagonal matrix with non-negative diagonal elements ordered in non-increasing order of amplitude, and the convention $\mathbf{C}_{r,0} = \mathbf{U}_{r,0} = \mathbf{\Lambda}_{r,0}^{1/2} = \mathbf{I}_{k_0}$ is used.

We now rewrite the average mutual information in function of matrices $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{D}_i$, in order to take the maximization in (5.13) over $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{D}_i$ instead of $\mathbf{P}_i$. Using (5.123) and (5.124) the average mutual information $\mathcal{I}$ can be expressed in function of matrices $\mathbf{\Theta}'_i$, $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{D}_i$ as

$$\begin{aligned}
\mathcal{I} &\triangleq \mathrm{E}\left[\log\det(\mathbf{I}_{k_N} + \eta\, \mathbf{G}_N \mathbf{G}_N^H)\right] \\
&= \mathrm{E}\Big[\log\det(\mathbf{I}_{k_N} + \eta\, \mathbf{U}_{r,N} \mathbf{\Lambda}_{r,N}^{1/2}\, \mathbf{\Theta}'_N\, \mathbf{U}_{N-1} \mathbf{D}_{N-1} \mathbf{V}_{N-1}^H\, \mathbf{\Theta}'_{N-1} \ldots \mathbf{U}_1 D_1 \mathbf{V}_1^H\, \mathbf{\Theta}'_1\, \mathbf{U}_0 \mathbf{D}_0 \mathbf{V}_0^H \\
&\qquad \mathbf{V}_0 \mathbf{D}_0^H \mathbf{U}_0^H\, \mathbf{\Theta}_1'^H\, \mathbf{V}_1 D_1^H \mathbf{U}_1^H \ldots \mathbf{\Theta}_{N-1}'^H\, \mathbf{V}_{N-1} \mathbf{D}_{N-1}^H \mathbf{U}_{N-1}^H\, \mathbf{\Theta}_N'^H \mathbf{\Lambda}_{r,N}^{1/2} \mathbf{U}_{r,N}^H)\Big]
\end{aligned} \tag{5.125}$$

$\mathbf{\Theta}'_i$ being zero-mean i.i.d. complex Gaussian, multiplying it by unitary matrices does not change its distribution, thus $\mathbf{\Theta}''_i = \mathbf{V}_i^H \mathbf{\Theta}'_i \mathbf{U}_{i-1}$ is equal in distribution to $\mathbf{\Theta}'_i$ and the average mutual information can be rewritten

$$\begin{aligned}
\mathcal{I} &= \mathrm{E}\Big[\log\det(\mathbf{I}_{k_N} + \eta\, \mathbf{\Lambda}_{r,N}^{1/2} \mathbf{\Theta}''_N \mathbf{D}_{N-1} \mathbf{\Theta}''_{N-1} \ldots \mathbf{D}_1 \mathbf{\Theta}''_1 \mathbf{D}_0 \\
&\qquad \mathbf{D}_0^H \mathbf{\Theta}_1''^H \mathbf{D}_1^H \ldots \mathbf{\Theta}_{N-1}''^H \mathbf{D}_{N-1}^H \mathbf{\Theta}_N''^H \mathbf{\Lambda}_{r,N}^{1/2})\Big] \\
&= \mathrm{E}\left[\log\det(\mathbf{I}_{k_N} + \eta\, \mathbf{\Lambda}_{r,N}^{1/2} \bigotimes_{i=N}^{1} \{\mathbf{\Theta}''_i \mathbf{D}_{i-1}\} \bigotimes_{i=1}^{N} \{\mathbf{D}_{i-1}^H \mathbf{\Theta}_i''^H\}\, \mathbf{\Lambda}_{r,N}^{1/2})\right]
\end{aligned} \tag{5.126}$$

Therefore, the maximum average mutual information can then be written

$$\boldsymbol{C} = \max_{\substack{\mathbf{D}_i, \mathbf{U}_i, \mathbf{V}_i \\ \mathrm{tr}(\mathrm{E}[\mathbf{x}_i \mathbf{x}_i^H]) \leq k_i \mathcal{P}_i \\ \forall i \in \{0, \ldots, N-1\}}} \mathrm{E}\left[\log\det(\mathbf{I}_{k_N} + \eta\, \mathbf{\Lambda}_{r,N}^{1/2} \bigotimes_{i=N}^{1} \{\mathbf{\Theta}''_i \mathbf{D}_{i-1}\} \bigotimes_{i=1}^{N} \{\mathbf{D}_{i-1}^H \mathbf{\Theta}_i''^H\}\, \mathbf{\Lambda}_{r,N}^{1/2})\right] \tag{5.127}$$

Expression (5.126) shows that the average mutual information $\mathcal{I}$ does not depend on the matrices $\mathbf{U}_i$ and $\mathbf{V}_i$, which determine the transmit directions

at source and relays, but only on the singular values contained in matrices $\mathbf{D}_i$. Nevertheless, as shown by (5.127), the maximum average mutual information $\boldsymbol{C}$ depends on the matrices $\mathbf{U}_i, \mathbf{V}_i$—and thus on the transmit directions—through the power constraints.

- **Step 2: give the expression of the power constraints in function of $\mathbf{D}_i, \mathbf{U}_i, \mathbf{V}_i$ and channel correlation matrices**

We show hereunder that the average power of transmitted signal $\mathbf{x}_i$ at i-th relaying level is given by

$$\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H]) = a_i\mathrm{tr}(\mathbf{P}_i\mathbf{C}_{r,i}\mathbf{P}_i^H) \prod_{k=0}^{i-1}\frac{a_k}{k_k}\mathrm{tr}(\mathbf{C}_{t,k+1}\mathbf{P}_k\mathbf{C}_{r,k}\mathbf{P}_k^H) \tag{5.128}$$

(with the convention: $\mathbf{C}_{r,0} = \mathbf{I}_{k_0}$ and $a_0 = 1$).

Proof. The average power of transmitted signal $\mathbf{x}_i$ can be written

$$\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H]) = \mathrm{tr}(\mathrm{E}[\bigotimes_{k=i}^{1}\{\mathbf{A}_k\boldsymbol{\Theta}_k\}\mathbf{A}_0\mathbf{A}_0^H\bigotimes_{k=1}^{i}\{\boldsymbol{\Theta}_k^H\mathbf{A}_k^H\}])$$

with

$$\begin{aligned}\mathbf{A}_i &= \mathbf{P}_i\mathbf{C}_{r,i}^{1/2}\\ \mathbf{A}_k &= \mathbf{M}_k = \mathbf{C}_{t,k+1}^{1/2}\mathbf{P}_k\mathbf{C}_{r,k}^{1/2}\ ,\ \forall k\in\{0,\ldots,i-1\}\\ \sigma_k^2 &= \frac{a_k}{k_{k-1}}\end{aligned} \tag{5.129}$$

(recall the convention: $\mathbf{C}_{r,0} = \mathbf{I}_{k_0}$ and $a_0 = 1$). Applying *Lemma 7* to $\mathrm{tr}(\mathrm{E}\{\mathbf{x}_i\mathbf{x}_i^H\})$ yields

$$\begin{aligned}\mathrm{tr}(\mathrm{E}[\mathbf{x}_i\mathbf{x}_i^H]) &= \mathrm{tr}(\mathbf{C}_{t,1}\mathbf{P}_0\mathbf{C}_{r,0}\mathbf{P}_0^H) \prod_{k=1}^{i-1}\frac{a_k}{k_{k-1}}\mathrm{tr}(\mathbf{C}_{t,k+1}\mathbf{P}_k\mathbf{C}_{r,k}\mathbf{P}_k^H) \frac{a_i}{k_{i-1}}\mathrm{tr}(\mathbf{P}_i\mathbf{C}_{r,i}\mathbf{P}_i^H)\\ &= a_i\mathrm{tr}(\mathbf{P}_i\mathbf{C}_{r,i}\mathbf{P}_i^H) \prod_{k=0}^{i-1}\frac{a_k}{k_k}\mathrm{tr}(\mathbf{C}_{t,k+1}\mathbf{P}_k\mathbf{C}_{r,k}\mathbf{P}_k^H)\end{aligned} \tag{5.130}$$

which concludes the proof. □

Using (5.128) in the power constraints (5.6), those constraints can be rewritten as a product of trace-factors:

$$\begin{aligned}
&\mathrm{tr}(\mathbf{P}_0\mathbf{P}_0^H) \leq k_0\mathcal{P}_0 \\
&a_i\mathrm{tr}(\mathbf{P}_i\mathbf{C}_{r,i}\mathbf{P}_i^H) \prod_{k=0}^{i-1}\frac{a_k}{k_k}\mathrm{tr}(\mathbf{C}_{t,k+1}\mathbf{P}_k\mathbf{C}_{r,k}\mathbf{P}_k^H) \leq k_i\mathcal{P}_i \ , \forall i \in \{1,\ldots,N-1\}
\end{aligned} \tag{5.131}$$

In order to express (5.131) in function of matrices $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{D}_i$, we first rewrite (5.124) as

$$\mathbf{P}_i = \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{t,i+1}^{-1/2}\mathbf{U}_i\mathbf{D}_i\mathbf{V}_i^H\mathbf{\Lambda}_{r,i}^{-1/2}\mathbf{U}_{r,i}^H \tag{5.132}$$

and inject (5.132) in (5.131) to obtain

$$\begin{aligned}
\mathrm{tr}(\mathbf{P}_i\mathbf{C}_{r,i}\mathbf{P}_i^H) &= \mathrm{tr}(\mathbf{U}_{t,i+1}\mathbf{\Lambda}_{t,i+1}^{-1/2}\mathbf{U}_i\mathbf{D}_i\mathbf{V}_i^H\mathbf{\Lambda}_{r,i}^{-1/2}\mathbf{U}_{r,i}^H \ \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{U}_{r,i}^H \\
&\qquad \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}^{-1/2}\mathbf{V}_i\mathbf{D}_i^H\mathbf{U}_i^H\mathbf{\Lambda}_{t,i+1}^{-1/2}\mathbf{U}_{t,i+1}^H) \\
&= \mathrm{tr}(\mathbf{\Lambda}_{t,i+1}^{-1}\mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i^H) \\
\mathrm{tr}(\mathbf{C}_{t,k+1}\mathbf{P}_k\mathbf{C}_{r,k}\mathbf{P}_k^H) &= \mathrm{tr}(\mathbf{D}_k\mathbf{D}_k^H) \\
&= \mathrm{tr}(\mathbf{D}_k^2)
\end{aligned} \tag{5.133}$$

where $\mathbf{D}_i^2 = \mathbf{D}_i\mathbf{D}_i^H$ is a real diagonal matrix with non-negative diagonal elements ordered in decreasing order. This leads to the following expression of the power constraints in function of $\mathbf{U}_i, \mathbf{D}_i$

$$\begin{aligned}
&\mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{U}_0\mathbf{D}_0^2\mathbf{U}_0^H) \leq k_0\mathcal{P}_0 \\
&a_i\mathrm{tr}(\mathbf{\Lambda}_{t,i+1}^{-1}\mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i^H) \leq \frac{k_i\mathcal{P}_i}{\prod_{k=0}^{i-1}\frac{a_k}{k_k}\mathrm{tr}(\mathbf{D}_k^2)} \ , \forall i \in \{2,\ldots,N-1\}
\end{aligned} \tag{5.134}$$

Previously in Step 1, it was shown that matrices $\mathbf{V}_i$ do not have an impact on the expression of the average mutual information $\mathcal{I}$ (5.126), and surprisingly (5.134) now shows that matrices $\mathbf{V}_i$ do not have an impact on the power constraints either, only matrices $\mathbf{U}_i$ and $\mathbf{D}_i$ do. It should also be noticed that matrix $\mathbf{U}_i$ has an impact on the power constraint of the i-th relay only.

- **Step 3: give the optimal transmit directions**

To determine the optimal directions of transmission at source, we apply *Lemma 4* to the source power constraint (5.134) $\mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{U}_0\mathbf{D}_0^2\mathbf{U}_0^H) \leq k_0\mathcal{P}_0$ and conclude that for all choices of diagonal elements of $\mathbf{D}_0^2$, the matrix $\mathbf{U}_0$ that minimizes the trace $\mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{U}_0\mathbf{D}_0^2\mathbf{U}_0^H)$ is $\mathbf{U}_0 = I_{k_0}$. Indeed, consider a matrix $\mathbf{D}_0$ and a matrix $\mathbf{U}_0$ that meet the constraint $\mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{U}_0\mathbf{D}_0^2\mathbf{U}_0^H) \leq k_0\mathcal{P}_0$. Then by *Lemma 4*, we have $\mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{D}_0^2) \leq \mathrm{tr}(\mathbf{\Lambda}_{t,1}^{-1}\mathbf{U}_0\mathbf{D}_0^2\mathbf{U}_0^H) \leq k_0\mathcal{P}_0$. Hence, $\mathbf{D}_0$ and $\mathbf{I}_0$ also meet the constraint, which means that the choice $\mathbf{U}_0 = I_{k_0}$ is feasible.

Therefore, the source precoder becomes

$$\begin{aligned}\mathbf{P}_0 &= \mathbf{U}_{t,1}\mathbf{\Lambda}_{t,1}^{-1/2}\mathbf{D}_0\mathbf{V}_0^H\mathbf{\Lambda}_{r,0}^{-1/2}\mathbf{U}_{r,0}^H = \mathbf{U}_{t,1}\mathbf{\Lambda}_{t,1}^{-1/2}\mathbf{D}_0\mathbf{V}_0^H \\ &= \mathbf{U}_{t,1}\mathbf{\Lambda}_{P_0}\mathbf{V}_0^H\end{aligned} \tag{5.135}$$

This recalls the known-result (5.25) in the single-hop MIMO case, where the optimal precoding covariance matrix at source was shown [95, 96] to be

$$\mathbf{Q}^\star \triangleq \mathrm{E}[\mathbf{x}_0\mathbf{x}_0^H] = \mathbf{P}_0\mathbf{P}_0^H = \mathbf{U}_{t,1}\mathbf{\Lambda}_{\mathbf{Q}^\star}\mathbf{U}_{t,1}^H \tag{5.136}$$

Similarly, to determine the optimal direction of transmission at i-th relaying level, we apply *Lemma 4* to the i-th power constraint: for all choices of diagonal elements of $\mathbf{D}_0^2$, the matrix $\mathbf{U}_i$ that minimizes the trace $\mathrm{tr}(\mathbf{\Lambda}_{t,i+1}^{-1}\mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i^H)$ is $\mathbf{U}_i = I_{k_i}$. This leads to the precoding matrix at level i

$$\mathbf{P}_i = \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{t,i+1}^{-1/2}\mathbf{D}_i\mathbf{V}_i^H\mathbf{\Lambda}_{r,i}^{-1/2}\mathbf{U}_{r,i}^H \tag{5.137}$$

Now since matrices $\mathbf{V}_i, i \in \{0, \ldots, N-1\}$ have an impact neither on the expression of the average mutual information nor on the power constraints, they can be chosen to be equal to identity: $\mathbf{V}_i = \mathbf{I}, i \in \{0, \ldots, N-1\}$. This leads to the (non-unique but simple) optimal precoding matrices

$$\begin{aligned}\mathbf{P}_0 &= \mathbf{U}_{t,1}\mathbf{\Lambda}_{P_0} \\ \mathbf{P}_i &= \mathbf{U}_{t,i+1}\mathbf{\Lambda}_{P_i}\mathbf{U}_{r,i}^H\end{aligned} \tag{5.138}$$

with the diagonal matrices $\mathbf{\Lambda}_{P_i} = \mathbf{\Lambda}_{t,i+1}^{-1/2}\mathbf{D}_i\mathbf{\Lambda}_{r,i}^{-1/2}$ containing the singular values of $\mathbf{P}_i$.

This completes the proof of Theorem 2. □

Chapter 6

Conclusion and Perspectives

In this thesis, we investigated physical layer techniques to improve the performance in wireless ad hoc networks. In Chapter 3, we focused on ad hoc networks, in which wireless nodes do not have advanced cooperative capabilities. We first highlighted the physical layer factors that can enhance the link capacity performance in dense ad hoc networks. In such interference-limited networks, the link capacity performance can be improved by using directive antennas, exploiting wireless nodes mobility, or managing the position and number of relaying nodes. In a first part, we focused more precisely on the impact of antenna directivity on the throughput and connectivity of a dense network of source-destination pairs. We proposed a dynamic blind beamforming scheme, that allows to benefit from antenna directivity in wireless ad hoc networks while avoiding heavy feedback to track the position of nodes. With the proposed scheme, the optimization of the network performance was shown to result from a tradeoff between interference-reduction achieved by using narrower transmission beams, and delay-increase caused by the successive blind orientations of the directive antenna in several directions to cover its intended destination without knowing its position. We showed that there exists an optimal beamwidth and number of antenna rotations that maximize the throughput and connectivity, and that their optimal values depend on the network density. In a second part, we considered a communication system where a source-destination pair communicates with the support of a dense network of passive relays. We studied the impact of the position and number of passive relaying nodes on the link capacity. The capacity of the system turned out to saturate when the number of relaying nodes increased. Indeed, the largest contributions to capacity resulted from retransmissions from relaying nodes close to the pair, while contributions of nodes located far from source or destination became more and more negligible as the network size increased. A few well located relays around source and destination were shown to lead to better performance than a larger number of relaying nodes uniformly distributed. This work also revealed that if cooperation between relaying nodes was enabled, gains usually obtained in MIMO systems could be exploited.

In Chapter 4 and 5, we turned our attention to wireless networks, in which relaying nodes had more advanced cooperative capabilities. In Chapter 4, we focused on cooperative networks with a small number of nodes, which represent the building blocks for larger networks. We showed that in classical cooperative strategies, the wireless resource is not efficiently used, in particular due to the orthogonality-constraint imposed to these strategies.

We proposed more spectrally efficient cooperative strategies. The proposed strategies preserved the practical half-duplex assumption, but relaxed the orthogonality constraint, by allowing a transmitter to combine messages from different origins in a single transmitted signal. The introduction of interference due to non-orthogonality was mitigated thanks to precoding, in particular to Dirty Paper Coding. Combined with smart power allocation, our cooperation strategies allowed to save time, led to a more efficient use of the bandwidth, and to an improved throughput with respect to classical Decode and Forward strategies.

Finally, in Chapter 5, we addressed the issue of the link capacity performance in ad hoc networks with a increasing number of communicating pairs. We proposed a cooperative model for ad hoc networks, in which nodes are grouped in cooperative clusters, and transmissions occur from a source cluster to a destination cluster via multi-hopping through intermediate relaying clusters. Each cluster was assumed to perform linear precoding on its received signal before retransmitting. The proposed model lies between the following two extreme models: on one hand, actual non-cooperative ad hoc networks, whose performance is known to be interference-limited, and whose throughput per communicating pair vanishes to zero as the density of the network increases; on the other hand, ideal fully-cooperative ad hoc networks, in which all nodes cooperate perfectly at transmission and reception, and whose capacity scales linearly with the number of communicating pairs as in a point-to-point MIMO system. Under the simplifying assumption that cooperation in the clusters was perfect, and using tools from free probability theory, we derived the asymptotic end-to-end mutual information of the network when the number of nodes in all clusters grew large. The mutual information per communicating pair was shown to converge to a deterministic value, depending only on the system characteristics and channel statistics. We also provided the optimal structure of the linear precoding performed at each cluster to maximize the system capacity: we showed that the singular vectors of the optimal precoding matrices should be aligned to the eigenvectors of channel correlation matrices. Finally, we applied the aforementioned results to several communication scenarios, with different numbers of hops, and different channel correlation models. These examples allowed to illustrate the performance gains resulting from the cooperative-cluster approach in ad hoc networks.

To conclude, we might say that the physical layer cooperative approach is a sound approach to improve the link capacity in dense wireless ad hoc

networks, as long as the cooperative strategies are properly designed to make an efficient use of the wireless resource.

Future Research Perspectives

We present here several areas and open issues that require further research.

First, the physical layer factors and techniques presented in Chapter 3 may not be the only ones that can improve the link capacity performance in dense wireless ad hoc networks by reducing interference. Indeed, Time reversal (TR) has been studied for quite a long time as a method to focus an ultrasonic wave both in time and space, and recent works started considering TR in wireless communications systems, in particular for Ultra Wide Band (UWB) applications. TR focusing capability in dense ad hoc networks could be used to mitigate interference and thus improve the link capacity performance.

Second, most works on cooperative networks mentioned in Chapters 4 and 5 were designed for and analyzed in the high SNR regime. Very few works targeted the low SNR regime of relaying networks. In the wideband regime, alternatively named low SNR regime since the power is shared over a large number of degrees of freedom, the performance is not interference-limited, but energy-limited. Taking advantage from the physical combination of signals in the wireless link, analog network coding appears as a relevant approach in the low SNR regime. Analog network coding could outperform other approaches, including cognitive radio, whose efforts are targeted at avoiding interference.

Finally, the analysis of the link capacity in a dense cooperative ad hoc network presented in Chapter 5 relied on the simplifying assumption of perfect cooperation within cooperative clusters. Further research is necessary to remove this simplifying assumption, and to take into account the cost of cluster cooperation on the network performance, in particular in terms of channel knowledge and feedback. Moreover, a particular topology was considered: sources and destinations could be gathered in single source and destination clusters respectively. A more complete analysis would provide the capacity in the case of more general network topologies.

Bibliography

[1] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[2] J. Li, C. Blake, D.S.J. De Couto, H.I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Proc. ACM MobiComŠ01, Rome, Italy*, July 2001.

[3] Y.D. Lin and Y.C. Hsu, "Multihop cellular: a new architecture for wireless communications," in *Proc. IEEE INFOCOM 2000, Annual Joint Conference of the IEEE Computer and Communications Societies*, Mar. 2000.

[4] H. Wu, C. Qiao, and S. De; O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 2105–2115, Oct. 2001.

[5] B. Pickholtz R. Vojcic B. Zadeh, A.N. Jabbari, "Self-organizing packet radio ad hoc networks with overlay SOPRANO," *IEEE Commun. Mag.*, vol. 40, no. 6, pp. 149–157, June 2002.

[6] D.S.J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A highthroughput path metric for multihop wireless routing," in *Proc. ACM MobiComŠ03, San Diego, USA*, Sept. 2003.

[7] A. Spyropoulos and C.S. Raghavendra, "Capacity bounds for ad-hoc networks using directional antennas," in *Proc. IEEE ICC'03*, May 2003, vol. 1, pp. 348–352.

[8] S. Yi, Y. Pei, and S. Kalyanaraman, "On the capacity improvement of ad hoc wireless networks using directional antennas," in *Proc. ACM MobiHoc'03*, June 2003, pp. 108–116.

[9] I. Jawhar and J. Wu, "Resource allocation in wireless networks using directional antennas," in *Proc. IEEE PERCOM'06*, Mar. 2006.

[10] S. Aeron, S. Venkatesh, "Scaling laws and operation of wireless ad-hoc and sensor networks," in *Proc. IEEE IWSSP'03*, Sept./Oct. 2003, pp. 367–370.

[11] J. Zhang and S. C. Liew, "Capacity improvement of wireless ad hoc networks with directional antennae," *ACM Mobile Computing and Communications Review*, 2006.

[12] T. Ueda, S. Tanaka, D. Saha, S. Roy, and S. Bandyopadhyay, "A rotational sector-based, receiver-oriented mechanism for location tracking and medium access control in ad hoc networks using directional antenna," in *Proc. IFIP PWC'03*, Sept. 2003.

[13] S. Bandyopadhyay, M. N. Pal, D. Saha, T. Ueda, K. Hasuike, and R. Pal, "Improving system performance of ad hoc wireless network with directional antenna," in *Proc. IEEE ICC'03*, May 2003, vol. 2, pp. 1146–1150.

[14] J.-L. Hsu and I. Rubin, "Performance analysis of directional random access scheme for multiple access mobile ad-hoc wireless networks," in *Proc. IEEE MILCOM'05*, Oct. 2005, vol. 1, pp. 45–51.

[15] M. Grossglauser and D.N.C. Tse, "Mobility increases the capacity of ad-hoc wireless networks," *IEEE/ACM Trans. Networking*, vol. 10, no. 4, pp. 477Ű486, Aug. 2002.

[16] S.N. Diggavi, M. Grossglauser, and D.N.C. Tse, "Even one-dimensional mobility increases the capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3947–3954, Nov. 2005.

[17] A.El. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *Proc. INFOCOM 2004*, Mar. 2004, vol. 1, pp. 464–475.

[18] N. Fawaz, M. Debbah, and D. Gesbert, "Capacity and positioning in dense scattering environments," in *Proc. SPAWC 2007, 8th IEEE Workshop on Signal Processing Advances for Wireless Communications, Helsinki, Finland*, June 2007.

[19] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[20] Z. Dawy and H. Kamoun, "The general gaussian relay channel: analysis and insights," in *Proc. 5th Int. ITG Conf. on Source and Channel Coding*, 2004, pp. 469–476.

[21] Z. Dawy and P. Leelapornchai, "Optimal number of relay nodes in wireless ad hoc networks with nonŰcooperative accessing schemes," in *Proc. ISITA 2002*, 2002.

[22] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. INFOCOM 2002*, 2002, pp. 1577–1586.

[23] I. Maric G. Kramer and R. D. Yates, *Cooperative Communications*, vol. 1, NOW Publishers Inc., Foundations and Trends in Networking, Hanover, 2006.

[24] F.H.P. Fitzek and M.D. Katz, *Cooperation in Wireless Networks: Principles and Applications-Real Egoistic Behavior is to Cooperate!*, Springer, Netherlands, 2006.

[25] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[26] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inform. Theory*, vol. 53, pp. 29–43, Jan. 2005.

[27] H. Bölcskei, R.U. Nabar, O. Oyman, and A.J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[28] V. Morgenshtern and H. Bölcskei, "Capacity of large amplify-and-forward relay networks," in *Proc. IEEE Communication Theory Workshop 2006*, May 2006.

[29] V. Morgenshtern and H. Bölcskei, "Large random matrix analysis of relay networks," in *Proc. 2006 Allerton Conference*, 2006.

[30] V.I. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3319–3349, Oct. 2007.

[31] H. Li, Z. Han, and H.V. Poor, "Asymptotic analysis of large cooperative relay networks using random matrix theory," *EURASIP Journal on Advances in Signal Processing*, Feb. 2008.

[32] S. Borade, L. Zheng, and R. Gallager, "Amplify-and-forward in wireless relay networks: Rate, diversity, and network size," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3302–3318, Oct. 2007.

[33] S. Yang and J.-C. Belfiore, "Diversity of MIMO multihop relay channels," submitted to IEEE Trans. Inform. Theory.

[34] S. Yeh and O. Leveque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE ISIT'07*, June 2007.

[35] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding strategy of multi-level precode & forward in correlated channels," in *Proc. IEEE ITW'08, Information Theory Workshop, Porto, Portugal*, May 2008.

[36] S. Aeron and V. Saligrama, "Wireless ad hoc networks: Strategies and scaling laws for the fixed snr regime," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2044–2059, June 2007.

[37] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3549–3572, Oct. 2007.

[38] I. E. Telatar, "Capacity of multi-antenna gaussian channels," Technical Memorandum, Oct. 1995.

[39] A.M. Tulino and S. Verdu, "Random matrix theory and wireless communications," in *Foundations and Trends in Communications and Information Theory*, vol. 1. NOW, the essence of knowledge, 2004.

[40] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20A, pp. 32Ű52, 1928.

[41] E. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Annals of Mathematics*, vol. 62, no. 3, pp. 548Ű564, Nov. 1955.

[42] V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457Ű483, 1967.

[43] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, pp. 41Ű59, May 1996.

[44] R.R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 2086–2091, July 2002.

[45] R.R. Müller, "A random matrix model of communication via antenna array," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2495–2506, Sept. 2002.

[46] D.N.C. Tse and S.V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[47] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[48] R.R. Müller, "Multiuser receivers for randomly spread signals: Fundamental limits with and without decision-feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 268Ű283, Jan. 2001.

[49] D.B. Johnson and D.A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*. 1996, pp. 153–181, Kluwer Academic Publishers.

[50] C. Perkins and E.M. Royer, "Ad-hoc on-demand distance vector routing," in *Proc. of WMCSAŠ99*, 1999.

[51] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, A. Qayyum, and L. Viennot, "Optimized link state routing protocol," in *Proc. of INMICŠ01*, 2001.

[52] Z.J. Haas and M.R. Pearlman, *Ad hoc networking*, chapter ZRP: a hybrid framework for routing in Ad Hoc networks.

[53] M. Sharif and B. Hassibi, "On the capacity of mimo broadcast channels with partial side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[54] C. Bettstetter, C. Hartmann, and C. Moser, "How does randomized beamforming improve the connectivity of ad hoc networks?," in *Proc. IEEE ICC'05*, May 2005, vol. 5, pp. 3380–3385.

[55] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*, Wiley-Interscience, New York, July 2006.

[56] X. Liu and R. Srikant, "An information-theoretic view of connectivity in wireless sensor networks," in *Proc. IEEE SECON'04*, Oct. 2004.

[57] H. Bolcskei, R.U. Nabar, O. Oyman, and A.J. Paulraj, "Capacity scaling laws in mimo relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[58] N. Fawaz, M. Debbah, and D. Gesbert, "Capacity of dense scattering environments," in *Proc. ARC IRAMUS Workshop 2007, Val Thorens, France*, Jan. 2007.

[59] N. Ida, *Engineering Electromagnetics, 2nd Edition*, Springer, May 2004.

[60] Theodore S.Rappaport, *Wireless Communications Principles and Practice, 2nd Edition*, Communications Engineering and Emerging Technologies. Prentice Hall, 2002.

[61] R.M. Gray, "Toeplitz and circulant matrices: A review," in *Foundations and Trends in Communications and Information Theory*, vol. 2, pp. 155–239. NOW, the essence of knowledge, Stanford University, 2006.

[62] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[63] S.Y.R. Li, R.W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[64] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[65] D. S. Lun, M. Médard, and R. Koetter, "Efficient operation of wireless packet networks using network coding," in *Proc. IWCT 2005, Oulu, Finland*, 2005.

[66] S. Deb, M. Effros, T. Ho, D.R. Karger, R. Koetter, D.S. Lun, M. Médard, and N. Ratnakar, "Network coding for wireless applications: A brief tutorial," in *Proc. IWWAN 2005, London, UK*, 2005.

[67] C. Fragouli and E. Soljanin, *Network Coding Fundamentals*, vol. 2, NOW Publishers Inc., Foundations and Trends in Networking, 2007.

[68] P. Chou, "Network coding for the internet and wireless networks," Tutorial, Mar. 2006.

[69] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, Mar. 1983.

[70] N. Fawaz, D. Gesbert, and M. Debbah, "When network coding and dirty paper coding meet in a cooperative ad hoc network," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, May 2008.

[71] N. Fawaz, D. Gesbert, and M. Debbah, "When network coding and dirty paper coding cooperate," in *Proc. IEEE Winter School on Coding and Information Theory 2007, La Colle sur Loup, France*, Mar. 2007.

[72] N. Fawaz, D. Gesbert, and M. Debbah, "When network coding and dirty paper coding meet in a cooperative ad hoc network," Tech. Rep. EURECOM+2279, EURECOM, France, Sept. 2006.

[73] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[74] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, part i, ii," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1948, Nov. 2003.

[75] T. E. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 375–391, Feb. 2006.

[76] S. Katti, I. Marić, A. Goldsmith, D. Katabi, and M. Médard, "Joint relaying and network coding in wireless networks," in *Proc. IEEE ISIT '07*, June 2007.

[77] C. T. K. Ng and A. J. Goldsmith, "Transmitter cooperation in ad-hoc wireless networks: Does dirty-paper coding beat relaying?," in *Proc. IEEE Information Theory Workshop 2004*, Oct. 2004, pp. 277–282.

[78] C. T. K. Ng, N. Jindal, A. J. Goldsmith, and U. Mitra, "Capacity gain from two-transmitter and two receiver cooperation," *IEEE Trans. Inform. Theory*, 2007.

[79] C. K. Lo, S. Vishwanath, and R. W. Jr. Heath, "Rate bounds for MIMO relay channels using precoding," in *Proc. IEEE GLOBECOM '05*, Nov./Dec. 2005, vol. 3, pp. 277–282.

[80] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[81] D.N.C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[82] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. IEEE GLOBECOM '01*, Nov. 2001, vol. 2, pp. 1344–1348.

[83] J. N. Laneman, *Cooperation in Wireless Networks: Principles and Applications*, chapter Cooperative Diversity: Models, Algorithms, and Architectures, Springer, 2006.

[84] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658– 2668, Oct. 2003.

[85] F. Xue and P.R. Kumar, *Scaling Laws for Ad Hoc Wireless Networks: An Information Theoretic Approach*, vol. 1, NOW Publishers Inc., Foundations and Trends in Networking, Hanover, 2006.

[86] L.L. Xie and P.R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[87] A. Jovicic, P. Viswanath, and S.R. Kulkarni, "Upper bounds to transport capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 50, no. 11, pp. 2555– 2565, Nov. 2004.

[88] F. Xue, L.L. Xie, and P.R. Kumar, "The transport capacity of wireless networks over fading channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 834–847, Mar. 2005.

[89] L.L. Xie and P.R. Kumar, "On the path-loss attenuation regime for positive cost and linear scaling of transport capacity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2313–2328, June 2006.

[90] S.H.A. Ahmad, A. Jovicic, and P. Viswanath, "On outer bounds to the capacity region of wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2770–2776, June 2006.

[91] D. Gesbert, M. Shafi, D.S. Shiu, P.J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[92] R. Vaze and R. W. Jr. Heath, "Capacity scaling for MIMO two-way relaying," submitted to IEEE Trans. Inform. Theory.

[93] L. Cottatellucci, T. Chan, and N. Fawaz, "Large system design and analysis of protocols for decode-forward relay networks," in *Proc. ICST WiOpt/PhysComNet 2008, Berlin, Germany*, Apr. 2008.

[94] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398–1407, Apr. 2007.

[95] S.A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, July 2004.

[96] A. Soysal and S. Ulukus, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI," *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1402–1412, Sept. 2007.

[97] S.L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 1350–1359, Sept. 2001.

[98] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, July 2004.

[99] C. Oestges, B. Clerckx, M. Guillaud, and M. Debbah, "Dual-polarized wireless communications: From propagation models to system performance evaluation," *IEEE Trans. Wireless Commun.*, May 2007, submitted.

[100] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, 1964.

[101] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, American Mathematical Society, 2000.

[102] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[103] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, no. 2, pp. 331–339, Nov. 1995.