Lattices and codes for wireless communications and security

Laura Luzzi

Mémoire présenté pour obtenir
l'Habilitation à Diriger des Recherches
de CY Cergy Paris Université

LATTICES AND CODES FOR WIRELESS COMMUNICATIONS AND SECURITY

LAURA LUZZI

ETIS UMR8051, CY Cergy Paris Université / ENSEA / CNRS
6 avenue du Ponceau, 95014 Cergy-Pontoise Cedex, France

Soutenue le 12 octobre 2023 devant le jury composé de:

Rapporteur	Jean-Pierre TILLICH	Directeur de Recherche à l'INRIA
Rapporteure	Michèle WIGGER	Professeur à Télécom-Paris
Rapporteur	Ram ZAMIR	Professor at Tel-Aviv University
Présidente du jury	Caroline FONTAINE	Directrice de Recherche au CNRS
Examinateur	Vincent H. POOR	Professor at Princeton University
Garante	Inbar FIJALKOW	Professeure à l'ENSEA
Référente	Iryna ANDRIYANOVA	Professeure à CY Cergy Paris Université

ABSTRACT

The many challenges posed by the evolution of communication systems often require an interdisciplinary approach. In particular, discrete mathematics, algebra and number theory play an important role in the design of codes for wireless communications and security. This thesis focuses on the applications of lattice coding (and to a lesser extent, of error-correction codes) to multi-antenna wireless communications, physical layer security and post-quantum cryptography.

First, we consider lattice space-time codes for multiple-antenna systems. We propose new techniques to reduce their decoding complexity, which is one of the main obstacles to their practical implementation. We also study the trade-off between rate and reliability in the high signal-to-noise-ratio regime. Finally, we consider a question of a more fundamental nature, namely the problem of approaching the capacity of multi-antenna channels when the number of blocks tends to infinity.

The second part of this work focuses on the applications of lattices to information-theoretic security. Borrowing tools from lattice-based cryptography, we design codes such that the output distributions induced by different confidential messages are indistinguishable, which achieve semantic security over Gaussian and wireless wiretap channels. This technique can also be applied to the problem of extracting secret keys from correlated Gaussian sources. We also propose reconciliation methods for two cryptographic key generation protocols based on lattices.

In the final part of the manuscript, we revisit the problem of approximation of output statistics from a different angle. We consider a two-node network comprised of an information source and a noisy channel, and we require the coordination of the signals at the input and at the output of the channel with the source and the reconstruction. Our objective is to characterize the set of achievable joint behaviors. Furthermore, we develop explicit polar coding schemes for coordination.

RÉSUMÉ

Les nombreux défis posés par l'évolution des systèmes de communication demandent souvent une approche interdisciplinaire. En particulier, les mathématiques discrètes, l'algèbre et la théorie de nombres jouent un rôle primordial dans la conception de codes pour les communications sans fil et la sécurité. Cette thèse porte sur les applications des réseaux euclidiens (et, en moindre mesure, des codes correcteurs) aux communications sans fil multi-antennes, à la sécurité au niveau de la couche physique et à la cryptographie post-quantique.

En premier lieu, nous considérons les schémas de codage espace-temps pour les systèmes multi-antennes. Nous proposons de nouvelles techniques pour réduire leur complexité de décodage. Nous étudions également le compromis fondamental entre fiabilité et débit dans le régime de fort rapport signal à bruit. Enfin, nous considérons une question de nature plus fondamentale, c'est-à-dire le problème de s'approcher de la capacité des canaux multi-antennes quand le nombre de blocs tend vers l'infini.

La deuxième partie de ce mémoire porte sur les applications des réseaux euclidiens à la sécurité en théorie de l'information. Nous exploitons des outils issus de la cryptographie basée sur les réseaux pour concevoir des codes tels que les distributions en sortie induites par des messages différents soient indistinguables, qui atteignent la sécurité sémantique des canaux wiretap gaussiens et sans fil. Cette technique s'applique également au problème de la génération de clés secrètes à partir de sources gaussiennes corrélées. Nous proposons aussi des méthodes de réconciliation pour deux protocoles de génération de clés cryptographiques basées sur les réseaux euclidiens.

Dans la partie finale du mémoire, nous revisitons le problème de l'approximation des statistiques en sortie d'un canal sous un angle différent. Nous considérons un modèle point-à-point composé d'une source d'information, d'un encodeur, d'un canal bruité, d'un décodeur, d'une information commune et nous cherchons à coordonner les signaux en entrée et en sortie du canal avec la source et sa reconstruction. Notre objectif est la caractérisation de l'ensemble des distributions de probabilité conjointes réalisables. De plus, nous développons de nouveaux codes polaires pour la coordination.

ACKNOWLEDGEMENTS

"Of course, the most rewarding part is the 'Aha' moment, the excitement of discovery and enjoyment of understanding something new – the feeling of being on top of a hill and having a clear view. But most of the time, doing mathematics for me is like being on a long hike with no trail and no end in sight."

–MARYAM MIRZAKHANI

My first thanks go to all my collaborators, without whom this work would not have been possible - in particular to Cong Ling for sharing his ideas and intuition, to Roope Vehkalahti for his enthusiasm and sense of humour, and to Matthieu Bloch for his amazing insight and vision. I am also indebted to my students Giulia Cervia, Charbel Saliba and Cécile Bouette: it has been a pleasure to see them grow more self-assured and independent during their PhD. Special thanks go to Maël Le Treust and Ligong Wang for sharing the adventure of co-supervising - we have quite different styles which hopefully complement each other - and to Inbar Fijalkow for being a mentor to my students.

I am very grateful to Iryna Andriyanova, Inbar Fijalkow, Caroline Fontaine, Vincent H. Poor, Jean-Pierre Tillich, Michèle Wigger and Ram Zamir for kindly accepting to participate in my HDR committee, and for their thought-provoking questions during the defense. I would especially like to thank the three referees for their attentive reading of the manuscript and in-depth reviews, and Inbar for guiding me through the whole process.

The ETIS lab has been a great place to work during these years thanks to many people, including those who left and whom I really miss - my colleagues and friends from the Information, Communication and Imaging team (Iryna Andriyanova, Veronica Belmega, Sara Berri, Kévin Carrier, Luan Chen, Arsenia Chorti, Inbar Fijalkow, Maël Le Treust, Ligong Wang, Claudio Weidmann) and from other teams (Myriam Ariaudo, Emmanuelle Bourdel, Aymeric Histace, Lilyana Petrova, David Picard, Camille Simon-Chane, Son Vu...) as well as the teaching colleagues at ENSEA (Christophe Barès, Philippe Bouafia, Matthieu Guerquin-Kern, Nicolas Simond, Antoine Tauvel...) - the list goes on.

Finally, thanks to Rob for being at my side during all these years and for his unconditional support through thick and thin (as well as occasional help with some proof!).

CONTENTS

Laura Luzzi

Date of birth: 26/3/1980
Nationality: Italian and French
Office phone: 01 30 73 62 96
Address: 6, avenue du Ponceau, 95014 Cergy-Pontoise
E-mail: laura.luzzi@ensea.fr
Webpage: https://perso.etis-lab.fr/luzzi/

Research interests

Wireless communications, MIMO systems, information theory, physical layer security, post-quantum cryptography

Professional experience

Sept. 2012 - present Assistant professor (Maître de conférences) at ENSEA
ETIS, UMR 8051, CY Cergy Paris Université, ENSEA, CNRS

Oct. 2011 - Aug. 2012 Marie Curie Research Fellow, Department of Electrical and Electronic Engineering, Imperial College London
Collaboration: Cong Ling

Oct. 2010 - Sept. 2011 Postdoctoral fellow at Supélec, Alcatel Lucent Chair in Flexible Radio
Co-funded by Georgia-Tech Lorraine
Collaboration: Matthieu Bloch

Oct. 2007 - May 2010 Postdoctoral researcher at Télécom-ParisTech, Comélec Department
Funding: ANR project ORIANA
Collaborations: Jean-Claude Belfiore, Ghaya Rekaya - Ben Othman

Feb. 2007 - Dec. 2008 Teaching assistant at the University Paris VI, Paris
Department of Mathematics

Education

2004 - 2007 **Ph.D.** in Applied Mathematics, Scuola Normale Superiore, Pisa, Italy
"*Continued fractions, coding and wireless channels*", 70/70 *cum laude*
Advisors: Stefano Marmi, Emanuele Viterbo

1998 - 2003 **Degree in Mathematics**, University of Pisa, Italy
"*Continued fraction expansions and geodesic flows: ergodic properties and symbolic dynamics* ", *110/110 cum laude*
Advisor: Stefano Marmi

Awards and distinctions

2017 "Women and Science" prize of the Paris Seine University
2011 Marie Curie Intra-European Fellowship at Imperial College London, U.K.
2006 - 2007 Exchange scholarship at the Ecole Normale Supérieure, Paris, France
2004 PhD Fellowship in Applied Mathematics, Scuola Normale Superiore, Pisa, Italy

Advising

PhD students 2 defended (50%, 90%), 1 ongoing (45%)
Postdocs 2

Publications

13 (10*) international journal papers
28 (27*) international conference papers (3 invited)
2 (2*) book chapters
2 (2*) patents
* Post-PhD

Scientific Responsibilities

2023 Jury member, PhD Prize in Signal, Image and Computer Vision (GDR ISIS/Club EEA)

Local Responsibilities

2021 - present Head of the ICI team of ETIS
2020 - 2021 Member of the Scientific Board of ETIS
2016 - 2019 Member of the Scientific Board of ENSEA

Teaching Responsibilities

2014 - 2023 Responsible for the 3rd year Option in Networks and Telecommunications at ENSEA

Workshop Organization

2014 GDR ISIS workshop “Physical layer security in wireless networks”, Paris

1 | ACTIVITY REVIEW

This section provides a detailed overview of my teaching and research activities since my PhD, as well as administrative and scientific responsibilities.

1.1 Teaching activities

Most of my teaching takes place in the Information Processing Department (DTI)[1] of ENSEA. In particular, since 2014 I have been **responsible for the 3rd year Option in Networks and Telecommunications (RT)**[2] (approximately 20 students).

A **reform of the 2nd year** was set up at ENSEA in 2017-18, with a new teaching organization into majors and minors. In this context, I have contributed to setting up a new major in Information Theory (with S. Vu). In 2020, a 2nd year **international group** was created, with full-immersion teaching in English, to which I participate. I also contributed to translating the course materials and problem sets of Information Theory as well as the labs of Random Signal Modelling.

In relation to my new research interest in cryptography, in 2021 I've taken the responsibility of the Network Security module in 3rd year RT. I have also contributed to creating a new course in Physical Layer Security in the Master I&ISC, option "Signal, Information and Telecommunications".

My typical teaching charge **per year** is **approximately 230 hours ETD**. The years 2018-19 (maternity leave) and 2019-20 (CRCT) correspond to a half teaching service.

Responsibility of teaching modules

In particular, I have been responsible for the following modules:

1. **Advanced Signal Processing** (2012 - 2016) - with I. Fijalkow, 3rd year ENSEA, option SyM[3]
2. **Communication Systems / Wireless Communications** (2012 - present) - with I. Fijalkow, 3rd year ENSEA, option RT(S) and SyM
3. **Telecommunications** (2017 - present) - with V. Belmega, 3rd year ENSEA, option RT(S)
4. **Information Theory and Multimedia Compression** (2017 - present) - with S. Vu, 2nd year ENSEA
5. **Physical Layer Security** (2021 - present), M2R I&ISC, CY University / ENSEA

I proposed and supervised **3rd year projects** on the following topics: fountain codes, low complexity detection for MIMO systems, polar codes.

[1] Formerly Signal and Telecommunications Department (DST).
[2] Now called Networks, Telecommunications and Security (RTS).
[3] Now SIA.

Administrative responsibilities

I have been the **coordinator for the 3rd year option "Networks and Telecommunications" (RT)** since 2014[4]. This responsibility is associated to a significant workload and corresponds to a mission of 48 hours ETD. It includes the creation of the timetables for the whole first semester, finding and helping guest teachers (printing lecture notes, exam surveillance), the organization of projects, validating the final year internship topics and the curriculums for our outgoing foreign exchange and dual degree students. Since 2019 we've set up a partnership with Nokia with conferences and résumé workshops, as well as a visit to their labs in Nozay.
In 2018 I coordinated a re-organization of teaching modules to improve readability and coherence of modules, which in the past were fragmented and often taught by guest teachers. In 2022 we've proposed a new reform of the option into Networks - Telecommunication - Security (RTS) by reinforcing the security modules (cryptography, network security, cybersecurity and software security).

1.2 Research Activities

1.2.1 Summary and main research interests

My research themes are at the interface between information and coding theory and discrete mathematics. In particular, they focus on lattice codes and error-correcting codes for wireless communications and security, and more recently on lattice-based cryptography. In several works, I have considered the application of algebraic tools such as the theory of Euclidean lattices and their theta series, of division algebras over algebraic number fields, and the ergodic theory of Lie groups and their arithmetic subgroups, in order to solve problems in coding theory.

After my degree in mathematics at the University of Pisa, Italy, I chose to pursue my PhD in applied mathematics at the Scuola Normale Superiore, Pisa. My PhD thesis is composed of two distinct sections. The first section, under the supervision of Stefano Marmi, focuses on the ergodic properties of a family of continued fraction expansions which generalize the Gauss algorithm [J1]. A follow-up work in collaboration with Keio University in Japan focuses on the same topic [J3].
The second part of the thesis was motivated by the applications of continued fractions to block coding for MIMO systems, and proposes new coded modulation schemes for slow fading channels [J2]. This work was in collaboration with Emanuele Viterbo during a visit to the Politecnico di Torino.

After this transition from pure mathematics to applications, I decided to focus on research in digital communications. A postdoctoral fellowship in Télécom-ParisTech with Jean-Claude Belfiore and Ghaya Rekaya-Ben Othman was the opportunity to acquire a solid background in this area, and to develop a new expertise in the topic of low-complexity decoding for MIMO systems [J4],[J5], which also led to two patent applications. In particular, thanks to my background in discrete mathematics, I proposed new techniques to exploit the algebraic properties of codes based on division algebras in order to simplify decoding.
I then joined the Alcatel-Lucent Chair in Flexible Radio at Supélec for a second postdoctoral fellowship on the topic of physical layer security (with Mérouane Debbah, in collaboration with Matthieu Bloch from Georgia-Tech Lorraine [C5]). At the same time I kept working on low-complexity MIMO decoding in collaboration with Frédérique Oggier of Nanyang University in Singapore [C6], with Cong Ling (Imperial College) and Damien Stehlé (ENS Lyon) [C7] and with Télécom-Paris [C4].
My collaboration with Cong Ling led to a successful application for a Marie Curie IEF Fellowship at Imperial College, where I worked on lattice coding for physical layer security [J8] and for the interference channel [C10]. I also started a new collaboration with Roope Vehkalahti at the University of Turku to study the diversity-multiplexing gain trade-off of division algebra codes [J7].
Finally, in 2012 I joined the Information, Communication and Imaging (ICI) team of the ETIS laboratory (UMR 8051 Cergy Paris University, ENSEA, CNRS) as an assistant professor. From February 2020 to July 2020, I benefited from a 6-month sabbatical (*Congé de Recherche et Conversion Thématique*) to focus on the new topic of post-quantum cryptography. Since November 2021, I am the head of the ICI team.
At ETIS, I have co-supervised three PhD students: Giulia Cervia on the topic of coordination of autonomous

[4]Shared with V. Belmega in 2014-15 and 2018-20, and with A. Chorti in 2022-23.

agents, Charbel Saliba on lattice-based cryptography, and Cécile Bouette (still ongoing) on covert communications.

In summary, my research post-PhD has focused on the following topics:

- **Low-complexity decoding techniques for MIMO systems**

  Publications: 3 journals: [J4],[J5],[J6], 5 conferences: [C2], [C3], [C4], [C6], [C7]
  Collaborators: Jean-Claude Belfiore (Télécom-Paris), Ghaya Rekaya Ben-Othman (Télécom-Paris), Frédérique Oggier (Nanyang University, Singapore), Cong Ling (Imperial College London), Damien Stehlé (ENS Lyon)
  Students: Asma Mejri (M1)
  Supported by: ANR project ORIANA

- **Diversity-multiplexing gain trade-off of space-time codes for MIMO systems**

  Publications: 2 journals: [J7], [J12], 5 conferences: [C13], [C14], [C15], [C18], [C22]
  Collaborators: Roope Vehalahti (University of Turku, Aalto University, and University of Jyväskylä, Finland), Hsiao-Feng Lu (National Chiao Tung University, Taiwan), Jean-Claude Belfiore (Télécom-Paris), Alexander Gorodnik (University of Bristol)
  Supported by: ENSEA, Academy of Finland, Finnish Cultural Foundation

- **Almost universal codes for MIMO channels with constant gap to capacity based on class field towers**

  Publications: 1 journal [J9], 2 conferences [C16], [C17], one book chapter [B2]
  Collaborators: Roope Vehalahti (University of Turku, Finland)
  Supported by: ENSEA, Academy of Finland, Finnish Cultural Foundation

- **Semantically secure lattice codes for wiretap channels**

  Publications: 2 journals [J8], [J10], 2 conferences [C8], [C19]
  Collaborators: Cong Ling (Imperial College London), Jean-Claude Belfiore (Télécom-Paris), Damien Stehlé (ENS Lyon), Roope Vehkalahti (Aalto University, Finland)
  Supported by: Marie Curie IEF Fellowship LACONIC, ENSEA

- **Secret key generation from correlated Gaussian sources**

  Publications: 1 conference [C12], 1 journal [J13]
  Collaborators: Cong Ling (Imperial College London), Matthieu Bloch (Georgia Tech)
  Supported by: Marie Curie IEF Fellowship LACONIC, INEX Ambition PHEBE

- **Strong coordination of signals and actions over noisy channels**

  Publications: 1 journal [J11], 4 conferences [C11], [C20], [C21], [C23], 1 national conference [NC1]
  Collaborators: Matthieu Bloch (Georgia Tech), Maël Le Treust (ETIS), Jörg Kliewer (New Mexico State University)
  Students: Giulia Cervia (PhD)
  Supported by: ENSEA, INS2I CNRS

- **Error correction and reconciliation for lattice-based post-quantum cryptography**

  Publications: 2 conferences [C24], [C27]
  Collaborators: Cong Ling (Imperial College London)
  Students: Charbel Saliba (PhD)
  Funded by: INEX project Lattice Hashing

- **Covert communications over non-Gaussian noise channels**

  Publications: 1 conference [C28]
  Students: Cécile Bouette (ongoing PhD)
  Collaborators: Ligong Wang (formerly at ETIS, now at ETH Zurich)
  Funded by: INEX Ambition PHEBE

Other current research interests include short blocklength wiretap code constructions based on polar and Reed–Muller codes [C26], and low-complexity decoders for non-binary polar codes [C25].

1.2.2 Advising

Defended PhD theses

1. **Giulia CERVIA**, "Coordination of autonomous devices over noisy channels: capacity results and coding techniques"

   Period: October 1st, 2015 - November 30th, 2018
   Advising: 50% (with Maël Le Treust (40%) and Inbar Fijalkow (10%, official PhD advisor)
   Funding: Ministry fellowship, "Science and Engineering" Doctoral School, University of Cergy-Pontoise
   Career path: Assistant professor at IMT Lille Douai since September 2020, after a postdoctoral fellowship at KTH Royal Institute of Technology Stockholm in 2019-20.
   Publications: 1 journal [J11], 3 conferences [C20], [C21], [C23], 1 national conference [NC1]

2. **Charbel SALIBA**, "Error correction and reconciliation techniques for lattice-based key generation protocols"

   Period: October 1st, 2017 - May 24th, 2022
   Advising: 90% (with Inbar Fijalkow (10%), official PhD advisor)
   Funding: INEX Paris-Seine project Lattice Hashing
   Publications: 2 conferences [C24], [C27]

Ongoing PhD theses

3. **Cécile BOUETTE**, "Information and coding-theoretic study of covert communication"

   Start date: November 1st, 2021
   Advising: 45% (with Ligong Wang (45%) and Inbar Fijalkow (10%), official PhD advisor)
   Funding: INEX Ambition project PHEBE "Physical-Layer Security for Beyond 5G"
   Publications: 1 conference [C28]

Unofficial advising

During my postdoctoral fellowship at Imperial College, I have worked closely with a PhD student, including unofficial advising:

- **Maria Constanza ESTELA ZAMORA**, "Interference management for interference channels: performance improvement and lattice techniques"
  PhD: Imperial College London
  Period: October 2011 - July 2012
  Collaboration with Cong Ling (Imperial College London)
  Publications: 1 conference [C10]

M2R students

- **Qi XUAN**, "Practical implementation of polar codes"
  Master: M2R SIC, University of Cergy-Pontoise
  Period: April - September 2015
  Advising: 100%

M2R Research Initiation Projects

- **Ivonne CHEBIB**, "Short packet transmission in 5G MIMO systems"
  M2R SIT option, November 2020 - March 2021

- **Qi XUAN**, "Analysis of Polar Codes for channel coding"
  M2R SIC, University of Cergy-Pontoise, November 2014 - March 2015

- **Xin YE**, "Wireless Network Coding"
  Master project at Imperial College, January - June 2012

M1 students

- **Kaiyu MU**, "Fountain Codes"
  Final year project at Imperial College London.
  Period: October 2011 - June 2012

- **Asma MEJRI**, "Diversity gain of MIMO decoders"
  Final year project at Télécom-ParisTech.
  Period: March - June 2010
  Publications: 1 conference [C4]

Postdocs

- **Franklin COCHACHIN**
  Topic: Low-complexity decoders for non-binary polar codes
  Period: July 2020 - June 2021
  Collaboration with Fakhreddine Ghaffari (ETIS)
  Funding: ANR QCSP
  Publications: 1 conference [C25]

- **Mahdi SHAKIBA-HERFEH**
  Topic: Secrecy analysis of short blocklength linear codes for the wiretap channel
  Period: December 2020 - November 2021
  Collaboration with Arsenia Chorti (ETIS)
  Publications: 1 conference [C26]

1.2.3 Research projects

- **Marie Curie Intra-European Fellowship** at Imperial College London
  Projet FP7 LACONIC *"Lattice Codes for Multiuser Wireless Communications"*, budget 192K€ (2011-2012)

- **Principal investigator** of the **INEX** Paris Seine project AAP 2017 (CY Initiative), "**Lattice hash functions** for secret key generation", budget 112.5k€ . Funding of Charbel Saliba's PhD thesis (2017-2021)
- **INEX Ambition PHEBE** "Physical-Layer Security for Beyond 5G" (CY Initiative), budget 391k€ (with L. Wang, A. Chorti, M. Le Treust, M. Chafii). Funding of Cécile Bouette's PhD thesis (2020-2024)
- **ANR QCSP** "Quasi-Cyclic Short Packet" with F. Ghaffari. Participants: Université de Bretagne Sud, IMT Atlantique, ETIS, IPB/ENSEIRB-MATMECA, Orange Labs, Sequans, CEA-LETI. Budget for ETIS: 80k€. Funding of the postdoc of F. Cochachin (2019-2023)
- **ANR JCJC DECODE** "Generic decoding for various metrics - A toolbox for post-quantum cryptography". Principal Investigator: K. Carrier, with N. Sendrier of Inria Paris. Budget: 183k€. Funding of Valérian Hatey's PhD thesis (2022-2025)
- **PEPR 5G** Project PC8 **E2ESec** (End-to-end Security for 5G), with A. Chorti. Budget: 71k€. (2023-2027)
- **EU HORIZON project** JU-SNS-2023 **ROBUST-6G** "Smart, Automated, and Reliable Security Service Platform for 6G", leader: ERICSSON Turkey. ETIS participants: A. Chorti (PI), L. Chen, S. Berri. Total budget for ETIS: 394k€ (provisional) (2023-2026)

I have also led several **local projects funded by ENSEA and UCP/CYU** for an overall amount of 7.7k€:

- **BQR-ENSEA-2013** "Lattice codes for strong secrecy in fading wiretap channels", 2k€. Funding of a research visit to Imperial College London.
- **BQR-ENSEA-2014** "Algebraic space-time codes for MIMO wireless systems", 1.5k€. Funding of a research visit to Turku University, Finland.
- **UCP-Invited-Professor-2016**, 1.2k€. *Invited professor:* Roope Vehkalahti, Turku University, Finland.
- **SRV-ENSEA-Invited-Professor-2017**, 2k€. *Invited professor:* Matthieu Bloch, Georgia Tech, U.S.A.
- **SRV-ENSEA-Invited-Professor-2018**, 1k€. *Invited professor:* Roope Vehkalahti, Aalto University, Finland.

1.2.4 Scientific responsibilities

Local responsibilities

- **Head of the ICI team** of the ETIS laboratory (Information, Communication and Imaging) since November 2021. The team currently comprises 4 Full Professors, 7 Assistant Professors, 10 PhD students, 1 postdoc.
  Duties: managing the team budget for missions and investments; participating to the monthly meetings of the Steering Group and to the Scientific Board of ETIS; writing the yearly activity report; organizing team meetings; coordinating outreach activities.
- Elected member of the Scientific Board of ETIS (October 2020 - November 2021)
- Elected member of the Scientific Board of ENSEA (April 2016 - January 2019)

National responsibilities

- Member of the jury of the Best PhD Prize in Signal, Image and Computer Vision awarded jointly by the French Club EEA, the GdR ISIS and GRETSI, June 2023

Workshop organization

- GdR ISIS workshop "Physical layer security in wireless networks", Télécom-Paris, May 2014
- Workshop "The arithmetics of wireless communications", Centro di Ricerca Matematica Ennio de Giorgi, Pisa, Italy, November 2008

Selection committees

- Member of the selection committee for the Assistant Professor (MCF) position n. 4046, Section 27/61 at ENSEA, April-May 2017.
- Member of the selection committee for the Assistant Professor (MCF) position n. 4069, Section 27/61 at ENSEA, April-May 2022.
- Member of the selection committee for the Assistant Professor (MCF) position n. 4079, Section 61 at ENSEA, May 2023
- Member of the selection committee for the Tenure-track Assistant Professor (MCF CDD) position in Systems, Networks and Security, Sections 27/61 at CY Cergy Paris Université, June 2023

Examiner for PhD and mid-term committees

- **PhD committee** of Hamed Mirghasemi at Télécom-Paris: "Lattice Codes for the Continuous Wiretap Channels", October 2014
- **Mid-term committee** of Sarah Kamel at Télécom-ParisTech: "Secure coding for Cloud-assisted Wireless Networks", July 2015
- **Mid-term committee** of Aymen Askri at Télécom-Paris: "Space-Time Codes for 5G", August 2019

Technical Program Committees

- IEEE Information Theory Workshop 2022
- IEEE Globecom Workshop on Enabling Security, Trust, and Privacy in 6G Wireless Systems, 2023

Session chair

- "Security, Privacy and Sharing", *IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013
- "Reed–Muller codes", *IEEE Information Theory Workshop* (online), October 2021
- "Security", *IEEE Information Theory Workshop* , Saint-Malo, France, May 2023

Reviewing

- **Journals:** regular reviewer for *IEEE Transactions on Information Theory* and *IEEE Transactions on Communications* and occasionally for other journals: *IEEE Communication Letters*, *IEEE Journal on Selected Areas in Information Theory*, *EURASIP Journal on Wireless Communications and Networking*, *Entropy*, *Transactions on Wireless Communications*, *Advances in Mathematics of Communications*.
- **International conferences:** regular reviewer for *IEEE International Symposium on Information Theory*, *IEEE Information Theory Workshop* and occasionally for others: *GLOBECOM*, *SPAWC*, *ICC*, *PIMRC*, *ISTC*.
- **National conferences:** *GRETSI*.
- **Research projects:** Reviewer for the French ANR Evaluation committee CE-48 "Foundations of digital technology - information technology, automation, signal processing" in 2019 and 2023

1.2.5 Collaborations

International Collaborations

- **Matthieu Bloch**, GeorgiaTech, U.S.A.
  Publications: 2 journals, 6 conferences
- **Cong Ling**, Imperial College London, U.K.
  Publications: 4 journals, 7 conferences
  Students: Maria Constanza Estela, Charbel Saliba
- **Roope Vehkalahti**, University of Jyväskylä (previously Turku and Aalto Universities), Finland
  Publications: 4 journals, 9 conferences, 1 book chapter
- **Alexander Gorodnik**, University of Bristol, U.K.
  Publications: 1 conference
- **Frédérique Oggier**, Nanyang University, Singapore
  Publications: 1 conference

National Collaborations

- **Jean-Claude Belfiore**, Télécom-Paris
  Publications: 4 journals, 6 conferences
- **Ghaya Rekaya-Ben Othman**, Télécom-Paris
  Publications: 3 journals, 4 conferences
- **Damien Stehlé**, ENS Lyon (LIP)
  Publications: 2 journals, 1 conference
- **Mérouane Debbah**, Centrale-Supélec
  Publications: 1 book chapter

Local Collaborations

- **Maël Le Treust**
  Publications: 1 journal, 3 conferences
  Students: Giulia Cervia
- **Arsenia Chorti**
  Publications: 1 conference
- **Ligong Wang**
  Publications: 1 conference
  Students: Cécile Bouette
- **Fakhreddine Ghaffari**
  Publications: 1 conference

1.2.6 Scientific visits

- **Nanyang University, Singapore**
  September 2010 (one month)
  Collaboration: Frédérique Oggier
  Related publication: [C6]

- **Finnish universities**
  - Aalto University, Helsinki, December 2011 (two weeks, invited by Camilla Hollanti).
  - Turku University, March 2013 (one week), June 2014 (one week), April 2015 (one month).
  - University of Jyväskylä, July 2022 (one week).
  Collaboration: Roope Vehkalahti
  Related publications: [J7], [J9], [C9], [C13], [C14], [C15], [C16], [C17], [C19]

- **Imperial College London, U.K.**
  March 2014 (one week), April-May 2016 (one month), February 2020 (one week)
  Collaboration: Cong Ling
  Related publications: [C19], [J10], [C24]

1.2.7 Invited talks

- "Algebraic reduction for low-complexity lattice decoding", workshop "Lattice Coding & Crypto Meeting", Imperial College London, U.K., September 2018

- "DMT classification of MIMO codes and ergodic theory of Lie groups", workshop "Interactions between number theory and wireless communication", University of York, U.K., July 2016

- "An introduction to algebraic coding for wireless channels", workshop "Interactions between algebra, coding theory and cryptography", University of Durham, U.K., January 2016

- "Secret key generation for Gaussian sources using lattices", workshop "Mathematical Tools of Information-Theoretic Security", Huawei Technologies, Paris, September 2015

- "Semantically secure lattice codes for the Gaussian wiretap channel", Special Session on Fundamentals and PHY, *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, London, U.K., September 2013

1.3 Awards and distinctions

Post PhD

- **Women and Science Prize** of the Paris Seine University (CY Alliance), March 2017

- **Marie Curie Intra-European Fellowship** at Imperial College London, U.K., 2011

Before and during the PhD

- **Exchange scholarship** at the Ecole Normale Supérieure, Paris, France, 2006

- **PhD Fellowship** in Applied Mathematics, Scuola Normale Superiore, Pisa, Italy (classed 2nd after written and oral examination)

1.4 Publications

The underlined author names refer to the PhD and post-doc students that I have co-advised officially and unofficially.

International journal papers (post-PhD):

[J13] **L. Luzzi**, C. Ling, M. Bloch, "Optimal rate-limited secret key generation from Gaussian sources using lattices", *IEEE Transactions on Information Theory*, vol. 69, n. 8, pp. 4944-4960, August 2023

[J12] R. Vehkalahti, **L. Luzzi**, "The DMT of Real and Quaternionic Lattice Codes and DMT Classification of Division Algebra Codes", *IEEE Transactions on Information Theory*, vol 68, n. 5, pp. 2999-3013, May 2022

[J11] G. Cervia, **L. Luzzi**, M. Le Treust, M. Bloch, "Strong coordination of signals and actions over noisy channels with two-sided state information", *IEEE Transactions on Information Theory* vol. 66, no 8, pp. 4681–4708, Aug. 2020

[J10] **L. Luzzi**, R. Vehkalahti, C. Ling, "Almost universal codes for MIMO wiretap channels", *IEEE Transactions on Information Theory*, vol. 64 n. 11, pp. 7218 – 7241, November 2018

[J9] **L. Luzzi**, R. Vehkalahti, "Almost universal codes achieving ergodic MIMO capacity within a constant gap", *IEEE Transactions on Information Theory*, vol. 63, n. 5, pp. 3224–3241, May 2017

[J8] C. Ling, **L. Luzzi**, J.-C. Belfiore, D. Stehlé, "Semantically Secure Lattice Codes for the Gaussian Wiretap Channel", *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6399-6416, 2014

[J7] R. Vehkalahti, H-F. Lu, **L. Luzzi**, "Inverse Determinant Sums and Connections Between Fading Channel Information Theory and Algebra", *IEEE Transactions on Information Theory*, vol 59, n. 9, pp. 6060–6082, 2013

[J6] **L. Luzzi**, D. Stehlé, C. Ling, "Decoding by embedding: correct decoding radius and DMT-optimality", *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2960–2973, 2013

[J5] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, "Algebraic reduction for the Golden Code", *Advances in Mathematics of Communications*, vol 6 n.1, pp. 1–26, 2012

[J4] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, "Augmented lattice reduction for MIMO decoding", *IEEE Transactions on Wireless Communications*, vol. 9, n.9, pp. 2853–2859, 2010

International journal papers (PhD):

[J3] **L. Luzzi**, S. Marmi, H. Nakada, R. Natsui, "Generalized Brjuno functions associated to α-continued fractions", *Journal of Approximation Theory*, vol. 162, n.1, pp. 24–41, 2010

[J2] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, E. Viterbo, "Golden Space-Time Block Coded Modulation", *IEEE Transactions on Information Theory* vol 55 n.2, pp. 584–597, 2009

[J1] **L. Luzzi**, S. Marmi, "On the entropy of Japanese continued fractions", *Discrete and Continuous Dynamical Systems Series A*, vol 20, n. 3, pp. 673–711, 2008

International conferences (post-PhD):

[C28] C. Bouette, **L. Luzzi**, L. Wang, "Covert Communication over two types of additive noise channels", *IEEE Information Theory Workshop*, Saint-Malo, France, April 2023

[C27] C. Saliba, **L. Luzzi**, C. Ling, "Error Correction for FrodoKEM Using the Gosset Lattice", *International Zurich Seminar on Information and Communication*, March 2022

[C26] M. Shakiba-Herfeh, **L. Luzzi**, A. Chorti, "Finite Blocklength Secrecy Analysis of Polar and Reed-Muller Codes in BEC Semi-Deterministic Wiretap Channels", *IEEE Information Theory Workshop*, Kanazawa, Japan, October 2021

[C25] F. Cochachin, **L. Luzzi**, F. Ghaffari, "Reduced Complexity of a Successive Cancellation Based Decoder for NB-Polar Codes", *International Symposium on Topics in Coding (ISTC)*, Montreal, Canada, August-September 2021

[C24] C. Saliba, **L. Luzzi**, C. Ling, "A reconciliation approach to key generation based on Module-LWE", *IEEE International Symposium on Information Theory*, Melbourne, Australia, July 2021

[C23] G. Cervia, **L. Luzzi**, M. Le Treust, M. Bloch, "Strong coordination over noisy channels with strictly causal encoding", *56th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, October 2018

[C22] **L. Luzzi**, R. Vehkalahti, "The DMT classification of real and quaternionic lattice codes", *IEEE International Symposium on Information Theory*, Vail, Colorado, June 2018

[C21] G. Cervia, **L. Luzzi**, M. Le Treust, M. Bloch, "Strong coordination of signals and actions over noisy channels", *IEEE International Symposium on Information Theory*, Aachen, Germany, June 2017

[C20] G. Cervia, **L. Luzzi**, M. Bloch, M. Le Treust, "Polar coding for empirical coordination of signals and actions over noisy channels", *IEEE Information Theory Workshop*, Cambridge (UK), September 2016

[C19] **L. Luzzi**, C. Ling, R. Vehkalahti, "Almost universal codes for fading wiretap channels", *IEEE International Symposium on Information Theory*, Barcelona, Spain, July 2016

[C18] **L. Luzzi**, R. Vehkalahti, A. Gorodnik, "Towards a complete DMT classification of division algebra codes", *IEEE International Symposium on Information Theory*, Barcelona, Spain, July 2016

[C17] **L. Luzzi**, R. Vehkalahti, "Division algebra codes achieve MIMO block fading channel capacity within a constant gap", *IEEE International Symposium on Information Theory*, Hong Kong, China, June 2015

[C16] R. Vehkalahti, **L. Luzzi**, "Number field lattices achieve Gaussian and Rayleigh channel capacity within a constant gap", *IEEE International Symposium on Information Theory*, Hong Kong, China, June 2015

[C15] R. Vehkalahti, **L. Luzzi**, J.-C. Belfiore, "Shifted inverse determinant sums and new bounds for the DMT of space-time lattice codes", *IEEE International Symposium on Information Theory*, Honolulu, HI, July 2014.

[C14] R. Vehkalahti, **L. Luzzi**, "Measuring the growth of inverse determinants sums of a family of quasi-orthogonal codes", *Proc. International Zurich Seminar on communications*, **invited paper**, February 2014

[C13] **L. Luzzi**, R. Vehkalahti, "A new design criterion for spherically-shaped division algebra-based space-time codes", *IEEE Information Theory Workshop*, Seville, Spain, September 2013

[C12] C. Ling, **L. Luzzi**, M. Bloch, "Secret key generation from Gaussian sources using lattice hashing", *IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013

[C11] M. Bloch, **L. Luzzi**, J. Kliewer, "Strong coordination with Polar Codes", *50th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, October 2012

[C10] M. C. Estela, **L. Luzzi**, C. Ling, J.-C. Belfiore, "Analysis of lattice codes for the many-to-one interference channel", IEEE Information Theory Workshop (ITW 2012), Lausanne, Switzerland, September 2012

[C9] R. Vehkalahti, **L. Luzzi**, "Connecting DMT of Division Algebra Space-Time Codes and Point Counting in Lie Groups", *IEEE International Symposium on Information Theory*, Cambridge (MA), July 2012

[C8] C. Ling, **L. Luzzi**, J.-C. Belfiore, "Lattice codes achieving strong secrecy over the mod-Λ Gaussian Channel", *IEEE International Symposium on Information Theory*, Cambridge (MA), July 2012

[C7] C. Ling, S. Liu, **L. Luzzi**, D. Stehlé, "Decoding by embedding: correct decoding radius and DMT-optimality", *IEEE International Symposium on Information Theory*, St. Petersburg, Russia, July 2011

[C6] **L. Luzzi**, F. Oggier, "A family of fast-decodable MIDO codes from crossed-product algebras over $\mathbb{Q}$", *IEEE International Symposium on Information Theory*, St. Petersburg, Russia, July 2011

[C5] **L. Luzzi**, M. Bloch, "Capacity-based random codes cannot achieve strong secrecy over symmetric wiretap channels", *1st International ICST Workshop on Secure Wireless Networks* (*Securenets 2011*), Cachan, France, **invited paper**, May 2011

[C4] A. Mejri, **L. Luzzi**, G. Rekaya-Ben Othman, "On the diversity of the Naive Lattice Decoder", *International Workshop on Systems, Signal Processing and their Applications*, Tipaza, Algeria, **invited paper**, May 2011

[C3] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, "Augmented lattice reduction for low-complexity MIMO decoding", *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, Istanbul, Turkey, September 2010

[C2] G. Rekaya-Ben Othman, **L. Luzzi**, J.-C. Belfiore, "Algebraic reduction for the Golden Code", *IEEE International Conference on Communications 2009*, Dresden, Germany, June 2009

International conference (PhD):

[C1] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, E. Viterbo, "Golden Space-Time Block Coded Modulation", *IEEE Information Theory Workshop 2008*, Porto, Portugal, May 2008

National conference (peer-reviewed, with proceedings):

[NC1] G. Cervia, **L. Luzzi**, M. Le Treust, M. Bloch, "Polar codes for empirical coordination over noisy channels with strictly causal encoding", in *Colloque GRETSI*, Juan-Les-Pins, September 2017

Book chapters:

[B2] R. Vehkalahti, **L. Luzzi**, "Algebraic Lattice Codes for Linear Fading Channels", in "Number Theory Meets Wireless Communications", Mathematical Engineering, Springer, 2020

[B1] A. Reznik, Y. Shah, **L. Luzzi**, M. Debbah, "Information-Theoretic Security in Wireless Systems", in "Security Technologies for an Ambient Lifestyle - Security, Privacy and Trust in the Wireless World", Wiley, 2013

Patents:

[P2] G. Rekaya-Ben Othman, J.-C. Belfiore, **L. Luzzi**, "Procédé de décodage d'un signal ayant subi un codage espace-temps avant émission, dans un système multi-antennaire, produit programme d'ordinateur et dispositif de décodage correspondant", (*Decoding procedure for a space-time encoded signal in a multi-antenna system, and the corresponding software and decoding device*), French patent application filed at the INPI (National Institute of Industrial Property), September 2, 2008

[P1] **L. Luzzi**, G. Rekaya-Ben Othman, J.-C. Belfiore, "Méthode de décodage par réseau de points augmenté pour système multi-source", (*Augmented lattice decoding method for multi-source systems*), French patent application filed at the INPI (National Institute of Industrial Property), December 30, 2009

2 | INTRODUCTION: LATTICES AND ERROR-CORRECTING CODES FOR COMMUNICATIONS AND SECURITY

The many challenges posed by the evolution of communication systems often require an interdisciplinary approach. In particular, discrete mathematics, algebra and number theory play an important role in the design of codes for wireless communications and security. This thesis focuses on the applications of lattice coding (and to a lesser extent, of error-correction codes) to multi-antenna wireless communications, physical layer security and post-quantum cryptography. In this introduction, I will put the accent on fundamental tools in geometry of numbers and information theory that are relevant to my work.

2.1 Lattice codes for communications over Gaussian and fading channels

A *lattice* is a discrete subgroup of $\mathbb{R}^n$. The properties of lattices have been studied in mathematics at least since the 18th century; since the beginning of the 20th century, the field of *geometry of numbers* [103] was advanced by H. Minkowski, C. L. Siegel and E. Hlawka, particularly in connection with number theory and Lie algebras. Despite a century of work on the topic, many questions are still open in this field, such as the problem of finding the densest sphere packings in any dimension, which has its roots in Hilbert's 18th problem; the optimality of lattice packings in dimension 8 and 24 was only settled recently thanks to a breakthrough by M. Viazowska.

In information theory and communications, lattice signal constellations are the natural counterpart of linear codes for continuous channels, where the Hamming metric is replaced by Euclidean distance. The constructions of good lattice packings from error-correcting codes have been studied at great length [49]. A series of works by Poltyrev [188], Loeliger [157] and Erez and Zamir [83] led to the proof that random lattices obtained from linear codes over finite fields achieve the capacity of the Gaussian channel. The existence of these good families of lattices can be essentially shown using the Minkowski-Hlawka-Siegel theorem [206], which states that the expected number of lattice points in a bounded measurable set S over the ensemble of random lattices of unit volume (with respect to the Haar measure) is equal to the volume of S.
Besides channel coding, lattices are a versatile tool to solve many other information-theoretic problems [236], such as source coding [82], side information problems [235, 84] and multiterminal settings (distributed source coding, broadcast, interference alignment [173]).

While the additive white Gaussian noise channel is a good model for deep-space links, modern wireless communications require more general channel models including time or frequency varying fading and multiple transmit and receive antennas. Good lattice "space-time" codes for fading and multiple input multiple-output (MIMO) channels must harness the diversity and multiplexing gain provided by wireless channels [211]. Due to fading, the multiplicative structure of the code also plays a role alongside its additive structure. In particular, lattice constructions from number fields and division algebras allow to design rotated lattice constellations that are optimal in terms of diversity and multiplexing gain [179, 178]. These algebraic lattice codes have been included in the DVB-T2 video broadcasting transmission scheme and in the WiMAX standard for wireless communications.

2.2 Lattice-based cryptography

Lattices also have important applications in computer science and cryptography, which have become especially relevant in the current search for next-generation cryptosystems. Lattice problems such as the *shortest vector problem* (SVP), the *shortest independent vector problem* (SIVP) the *closest vector problem* (CVP) and the corresponding approximate versions are believed to be computationally hard even for quantum computers except for very large approximation factors, and can be used to instantiate cryptographic primitives. Moreover, lattice-based protocols led to the first example of fully homomorphic encryption, which could revolutionize cloud computing [95].
Lattice problems are easier to solve if a "good" basis of the lattice is available, that is, a nearly orthogonal basis with relatively short vectors. *Lattice reduction algorithms* such as LLL [148] or BKZ [203] aim to find an improved basis when given an arbitrary basis as input, and are widely used for the cryptanalysis of lattice-based protocols, together with other techniques such as *sieving* [6] or *enumeration* [126].
A significant advantage of lattice-based cryptography compared to other cryptographic techniques is that the security of most lattice primitives is based on the *worst-case* hardness of lattice problems. In cryptography a problem is considered to be hard only if it is hard in the average-case, i.e. it is hard for all but a negligible fraction of instances. In 1996, a breakthrough paper by Ajtai [5] showed that solving SVP on a random lattice on average involves a solution for the approximate SVP for *any* lattice within a polynomial approximation factor. This connection illustrates the *worst-case to average-case reduction*: if the latter problem is hard in some (worst) cases, then the former is also hard on average.

Many cryptographic primitives were constructed based on this concept; one of the most versatile and widely used is the Learning With Errors (LWE) problem introduced by Regev [193]. Informally speaking, the (decision) LWE problem can be stated as follows. Given a secret vector $\mathbf{s} \in \mathbb{Z}_q^n$ and a matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$, let $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$, where $\mathbf{e}$ is an error term drawn from a rounded Gaussian or discrete Gaussian distribution. The problem is to distinguish between a finite number of samples $(\mathbf{A}, \mathbf{b})$ of the previous form and uniform samples.
Regev's work builds on the *transference theorems* developed by Banaszczyk [15] using harmonic analysis techniques, which relate the lengths of the successive minima of a lattice and of its dual. He introduces the *smoothing parameter* of a lattice which corresponds to the smallest standard deviation such that a discrete Gaussian distribution over the lattice behaves like a continuous Gaussian. Equivalently, the distribution of a Gaussian noise modulo the lattice is close to uniform provided that its variance is larger than the smoothing parameter. Regev showed that there exists a polynomial-time quantum reduction from approximate SIVP with approximation factor γ to LWE, provided that γ is small compared to the smoothing parameter of the lattice. The proof is based on an iterative reduction to the the problem of sampling from a discrete Gaussian over the lattice (*Discrete Gaussian Sampling* or DGS). Its connection to SIVP comes from the fact that if the variance of the discrete Gaussian is not too large with respect to the smoothing parameter, one can obtain short lattice vectors with high probability by sampling.

Later works introduced structured variants of LWE such as RLWE [161] and MLWE [135] which involve ideal lattices and module lattices respectively. Their cryptographic applications are more efficient compared to LWE. In particular, the suites of post-quantum key encapsulation mechanisms and digital signatures CRYSTALS-KYBER [13] and DILITHIUM, selected by NIST for standardization in 2022, are both based on MLWE. However, it is still an open question whether the additional algebraic structure might make these variants more vulnerable to attacks.

2.3 Fundamental information-theoretic metrics for secrecy and coordination

In my work I have mostly considered the security applications of lattices at the physical layer. While cryptography operates at the upper layers assuming that the communication links are error-free and employs pseudorandom noise, physical layer security aims to exploit the noise inherent in physical channels to guarantee the confidentiality of communications [149, 30]. Although physical layer security has remained largely theoretical up to now, mostly due to the difficulty of verifying the required hypotheses about the channel or source models required to obtain a physical advantage over attackers, the sixth generation of wireless networks will provide some key enablers for its practical implementation [170, 44]. Unlike cryptography, physical layer security guarantees information-theoretic

secrecy, which is measured in terms of statistical independence between the confidential message (or secret key) and the eavesdropper's observations. Thus, it is secure even against computationally unbounded adversaries, and consequently, also quantum-secure.

In my work, I have mostly focused on the channel resolvability approach [29] to obtain information-theoretic security. The *resolvability of a channel* is defined as the minimum coding rate such that the output of the code through the channel is asymptotically statistically similar to the output of a given source (in terms of variational distance or normalized Kullback-Leibler divergence). It was first studied by Han and Verdù [108], who showed that for a large class of models, the resolvability of a channel is equal to its capacity.
In wiretap coding, the average variational distance between output distributions corresponding to different confidential messages provides a bound for the mutual information between the message and the signal observed by an eavesdropper, following an approach that can be traced back to Csiszàr [53]. Thus, one technique to design wiretap codes is to employ a binning structure where each bin is a resolvability code for the eavesdropper's channel [114, 29].
Besides secrecy applications, the notion of resolvability and approximation of output statistics is also the basis for a new framework in information theory which considers other purposes for communication beyond the transfer of information, namely the problem of coordinating the actions of autonomous agents [56]. In particular, the degree of coordination towards a specific goal can be measured by the variational distance of the distribution of the sequence of joint actions of the network nodes to a target distribution. This problem falls within the scope of goal-oriented communications.

2.4 Main research contributions and organization of this thesis

My main research contributions after the PhD can be grouped into three axes.

Lattice coding and decoding for MIMO systems

The first axis concerns lattice space-time coding and decoding for multiple antenna systems.

In my postdoctoral research work, I considered the problem of **reducing the complexity of decoding of space-time block codes**, which is one of the main obstacles to their adoption in communication standards. Note that the decoding problem essentially corresponds to the closest vector problem (CVP) in the "faded" lattice constellation. We propose two low-complexity decoding techniques; the first is specialized to lattice codes based on division algebras and exploits their particular algebraic structure, and in particular the structure of their unit group, in order to simplify the decoding [J5]. The second technique is more general and consists in embedding the faded lattice into a higher-dimensional lattice exhibiting a large gap between the first two successive minima, so that lattice reduction algorithms are guaranteed to solve CVP [J6].

Another interesting question for the design of space-time codes is the characterization of the trade-off between rate and reliability [238] depending on their algebraic structure. In the high signal-to-noise ratio (SNR) regime, the question had already been settled in the symmetric setting (i.e. when the number of transmit and receive antennas are equal) under the non-vanishing determinant condition [80], but remained open for the asymmetric setting. In collaboration with Roope Vehkalahti, we proposed a general framework to compare the **trade-off between diversity and multiplexing gain (DMT) for asymmetric space-time codes based on division algebras**. In particular, we uncovered a surprising connection with the growth of the size of sets of elements of bounded norm in arithmetic subgroups of Lie groups [J7],[C18]. More precisely, we showed that the union bound for the pairwise error probability is essentially determined by the behavior of an inverse determinant sum over the unit group of the code. Recent results in the ergodic theory of arithmetic subgroups of Lie groups [100], which allow to approximate a sum over the arithmetic group with an integral over the corresponding Lie group, provide the bridge between discrete and continuous settings and allow to recover the diversity-multiplexing trade-off. Finally, we were able to propose a complete classification of the DMT of asymmetric space-time codes [J12].

Another line of research concerns more fundamental questions. In fact, even though the characterization of the ergodic capacity of multi-antenna systems was well-known [213], there were no known families of explicit

codes achieving this capacity. We made a first step towards answering this question by analyzing the asymptotic performance of multi-block algebraic space-time codes when the number of blocks tends to infinity [J9].

For the single antenna Gaussian channel, it was known using sphere packing arguments that the Hermite invariant of a family of lattice codes determines its gap to capacity. We showed that the normalized minimum determinant plays a similar role for space-time codes over MIMO fading channels. Thanks to this design criterion, we show that there exist **families of multi-block space-time codes from division algebras** which are *approximately universal* [C17], in the sense that they **achieve a constant gap to capacity for a general class of fading models**. The universality property guarantees robustness with respect to imperfect estimation of channel statistics in high-mobility scenarios, or in broadcast mode. This construction is unfortunately not explicit since it requires the computation of Hilbert class field towers [163] which is still a difficult problem in computational algebra.

These works are presented in **Chapter 3**.

Lattice coding for physical layer security

Borrowing tools from lattice-based cryptography such as the smoothing parameter of a lattice [166], we propose a fundamental parameter to design lattice codes that are good for secrecy, the *flatness factor*, which can be computed from the theta series of the lattice. We obtain secrecy through channel resolvability [29], by designing a code such that the output distributions induced by each confidential message are indistinguishable. The flatness factor (based on the L^∞ metric) provides an upper bound for the variational distance of output distributions and, consequently, for the information leakage. This new figure of merit for lattices allows to define the notion of *secrecy-goodness*: a family of lattices is secrecy-good if its flatness factor vanishes asymptotically. Using a Minkowski-Hlawka type bound on the average behaviour of the theta series, we show the existence of secrecy-good lattices under suitable volume conditions. We propose a new wiretap coding scheme where each confidential message is encoded according to a *discrete Gaussian* distribution over a coset of a secrecy-good lattice. This leads to a construction of **lattice codes which achieve strong secrecy over Gaussian wiretap channels** for rates up to $1/2$ nat from the secrecy capacity [J8]. These codes are also **semantically secure** [22] since no assumption on the distribution of the confidental messages is required.

The previous approach can be **generalized to fading and multiple-antenna channels** [J10]. As a consequence of Banaszczyk's transference theorems [15], we show that the flatness factor of the faded lattice vanishes if its dual lattice has good minimum distance. This leads to a simple code design criterion: the product between the minimum determinants of the lattice and of its dual should be maximized. Moreover, we propose a wiretap code construction based on ideal lattices from class field towers, which achieves strong secrecy and semantic security up to a constant gap to secrecy capacity for general fading models under the hypothesis of partial statistical channel state information at the transmitter as well as some compound channel models. Universality is an important property in practice, since the statistics of the eavesdropper's channel are typically unknown.

In the recent work [J13], we propose a **lattice-based scheme for secret key generation from Gaussian sources in the presence of an eavesdropper**. This type of protocol could allow the distribution of secret keys in decentralized networks. Typically, secret key generation consists of two distinct procedures: information reconciliation, in which Alice and Bob exchange public messages to agree on a common sequence, and privacy amplification to extract a secret key from this shared sequence. The main novelty is the use of the *modulo lattice operation* for privacy amplification, which allows to extract the intrinsic randomness of the channel [28]. Our information reconciliation step follows the outline of lattice Wyner-Ziv coding as in [235], but we introduce a *randomized lattice quantization technique*. Furthermore, we introduce two new notions of flatness factors based on L^1 distance and KL divergence, which improve upon the volume-to-noise ratio threshold of the L^∞ flatness factor and may be of independent interest. Consequently, our scheme achieves the strong secret key capacity of degraded source models, as well as the optimal secret key rate / public communication rate trade-off.

Lattice-based reconciliation techniques can also be applied to secret key generation in lattice cryptography. As part of Charbel Saliba's thesis, we considered **reconciliation and error-correction techniques for two cryptographic key generation protocols** based on LWE, KYBERKEM and FRODOKEM, with the aim to improve their reliability, secrecy and bandwidth requirements [C24][C27].

These contributions are presented in **Chapter 4**.

Strong coordination in point-to-point networks

As part of Giulia Cervia's PhD thesis, we revisited the problem of approximation of output statistics from a different angle, namely the coordination of actions of autonomous agents, where the goal is to induce a prescribed output distribution at the nodes with the least amount of communication. We consider a two-node network comprised of an information source and a noisy channel, and we require the **coordination of the signals at the input and at the output of the channel with the source and the reconstruction**. We assume that the encoder and decoder share a common source of randomness and we introduce a state capturing the effect of the environment. Our objective is to **characterize the strong coordination region**, i.e. the set of achievable joint behaviors and the required minimal rates of common randomness. We prove general inner and outer bounds for this region, and characterize the exact coordination region in three particular cases: when the channel is perfect, when the decoder is lossless and when the random variables of the channel are separated from the random variables of the source. The study of the latter case allows us to show that the joint source-channel separation principle does not hold for strong coordination. We also prove that strong coordination offers "free" security guarantees at the physical layer. Furthermore, we develop **explicit polar coding schemes** for coordination by exploiting the technique of source polarization.

These results are presented in **Chapter 5**.

Finally, **Chapter 6** illustrates some open problems and **research perspectives** in the fields of physical layer security and post-quantum cryptography.

2.5 Notation and definitions

We list here some notations that are used in the remainder of the manuscript.
We use column notation for vectors. We define the integer interval $[\![a, b]\!]$ as the set of integers between a and b. We use the notation $M_{m,n}(\mathbb{F})$ for the set of $m \times n$ matrices with elements in the field $\mathbb{F}$. Given a complex matrix $X \in M_{m,n}(\mathbb{C})$, $\|X\| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |x_{i,j}|^2}$ denotes its Frobenius norm. The notation I_n stands for the identity matrix of size n.
Given a matrix A, its transpose is denoted by A^t, and its Hermitian transpose by $A^\dagger$. Given Hermitian matrices A and B, the notation $A \succcurlyeq 0$ indicates that A is positive semidefinite; the notation $A \succcurlyeq B$ means that $A - B \succcurlyeq 0$. The notation $\mathrm{diag}(A_1, \ldots, A_n)$ will stand for the block diagonal matrix with diagonal blocks $A_1, \ldots, A_n$.

Given a finite set $\mathcal{A}$, we denote the uniform distribution on $\mathcal{A}$ by $\mathcal{U}_\mathcal{A}$. The *variational distance* or *statistical distance* between two distributions p and q taking values in $\mathcal{X}$ is defined as $\mathbb{V}(p, q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|$ in the discrete case, and $\mathbb{V}(p, q) = \int_\mathcal{X} |p(x) - q(x)|\, dx$ in the continuous case. Their *Kullback-Leibler divergence* is $\mathbb{D}(p||q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$ in the discrete case and $\mathbb{D}(p||q) = \int_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)} dx$ in the continuous case.

The special linear group $\mathrm{SL}_n(\mathbb{F})$ of degree n over a field $\mathbb{F}$ is the set of $n \times n$ matrices with determinant 1.
Given an number field F, its ring of integers will be denoted $\mathcal{O}_F$ and its discriminant will be denoted d_F.
We denote by $\mathcal{N}_\mathbb{C}(\mu, \sigma^2)$ a circularly symmetric complex Gaussian random variable with mean μ and variance σ^2 per complex dimension (equivalently, variance $\sigma^2/2$ per real dimension).

Lattices. The lattice generated by a basis matrix B is denoted by $\mathcal{L}(B) = \{B\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^n\}$. For a vector $\mathbf{x}$, the nearest-neighbor quantizer associated with Λ is $Q_\Lambda(\mathbf{x}) = \arg\min_{\boldsymbol{\lambda} \in \Lambda} \|\boldsymbol{\lambda} - \mathbf{x}\|$. We define the usual modulo lattice operation by $\mathbf{x} \mod \Lambda \triangleq \mathbf{x} - Q_\Lambda(\mathbf{x})$. A measurable set $\mathcal{R}(\Lambda) \subset \mathbb{R}^n$ is a fundamental region of the lattice Λ if $\cup_{\boldsymbol{\lambda} \in \Lambda} (\mathcal{R}(\Lambda) + \boldsymbol{\lambda}) = \mathbb{R}^n$ and if $(\mathcal{R}(\Lambda) + \boldsymbol{\lambda}) \cap (\mathcal{R}(\Lambda) + \boldsymbol{\lambda}')$ has measure 0 for any $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}'$ in Λ. The Voronoi cell of Λ, defined by $\mathcal{V}(\Lambda) = \{\mathbf{x} : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}$, specifies the nearest-neighbor decoding region. The Voronoi cell is one example of the fundamental region of a lattice. Given a fundamental region $\mathcal{R}(\Lambda)$, the mod $\mathcal{R}(\Lambda)$ operation is defined by $\mathbf{x} \mapsto \check{\mathbf{x}}$ where $\check{\mathbf{x}}$ is the unique element of $\mathcal{R}(\Lambda)$ such that $\check{\mathbf{x}} - \mathbf{x} \in \Lambda$. Obviously, the usual mod-Λ operation corresponds to the case where $\mathcal{R}(\Lambda) = \mathcal{V}(\Lambda)$.

For a (full-rank) sublattice $\Lambda' \subset \Lambda$, the finite group Λ/Λ' is defined as the group of distinct cosets $\boldsymbol{\lambda} + \Lambda'$ for $\boldsymbol{\lambda} \in \Lambda$. Denote by $[\Lambda/\Lambda']$ a set of coset representatives. The lattices Λ' and Λ are often said to form a pair of

nested lattices, in which Λ is referred to as the fine lattice while Λ' the coarse lattice. The order of the quotient group Λ/Λ' is equal to $V(\Lambda')/V(\Lambda)$. The norm of any shortest vector of Λ, often referred to as the *minimum distance*, is denoted by $\lambda_1(\Lambda)$ or $\lambda_1(B)$ when a basis B of Λ is given. The distance of a vector $\mathbf{y} \in \mathbb{R}^n$ to a lattice $\Lambda \subset \mathbb{R}^n$ is $\mathrm{dist}(\mathbf{y}, \Lambda) = \min_{\lambda \in \Lambda} \|\mathbf{y} - \lambda\|$.

Definition 2.1 (Hermite constant) *The Hermite constant is defined as*

$$\gamma_n \triangleq \sup_{\Lambda} \frac{\lambda_1^2(\Lambda)}{\det(\Lambda)^{2/n}}, \tag{2.1}$$

where the supremum is taken over all lattices Λ of dimension n.

Lattice problems. We now give precise definitions of the lattice problems that are relevant for this work. In all these problems, the input lattice Λ is described by an arbitrary basis B, and $\gamma = \gamma(n)$ is the approximation factor.

- *Closest Vector Problem (*CVP*):*
  Given a lattice Λ and a vector $\mathbf{y} \in \mathbb{R}^m$, find a vector $B\hat{\mathbf{x}} \in \Lambda$ such that $\|\mathbf{y} - B\hat{\mathbf{x}}\|$ is minimal.
- *γ-Approximate CVP (*CVP_γ*), with $\gamma \geq 1$:*
  Given a lattice Λ and a vector $\mathbf{y} \in \mathbb{R}^m$, find a vector $B\hat{\mathbf{x}} \in \Lambda$ such that $\|\mathbf{y} - B\hat{\mathbf{x}}\| \leq \gamma \mathrm{dist}(\mathbf{y}, B)$.
- *η-Bounded Distance Decoding (*BDD_η*) with $\eta \leq 1/2$:*
  Given a lattice Λ and a vector $\mathbf{y}$ such that $\mathrm{dist}(\mathbf{y}, B) < \eta\lambda_1$, find the lattice vector $B\hat{\mathbf{x}} \in \mathcal{L}(B)$ closest to $\mathbf{y}$.
- *Shortest Vector Problem (*SVP*):*
  Given a lattice Λ, find a vector $\mathbf{v} \in \Lambda$ of norm λ_1.
- *γ-Approximate SVP (*SVP_γ*), with $\gamma \geq 1$:*
  Given a lattice Λ, find a vector $\mathbf{v} \in \Lambda$ such that $0 < \|\mathbf{v}\| \leq \gamma\lambda_1$.
- *γ-Approximate Shortest Independent Vector Problem (*SIVP_γ*), with $\gamma \geq 1$:*
  Given an n-dimensional lattice Λ, output a set of n linearly independent vectors in Λ of length at most $\gamma \cdot \lambda_n(\Lambda)$.
- *γ-unique SVP (*uSVP_γ*), with $\gamma \geq 1$:*
  Given a lattice Λ such that $\lambda_2(\Lambda) > \gamma\lambda_1(\Lambda)$, find a vector $\mathbf{v} \in \Lambda$ of norm $\lambda_1(\Lambda)$.

3 | ALGEBRAIC SPACE-TIME CODES FOR MIMO SYSTEMS: PERFORMANCE AND DECODING COMPLEXITY

Algebraic number theory is an effective tool to design lattice "space-time" codes which exploit the diversity and multiplexing gain of multiple input multiple output (MIMO) wireless systems, which are now adopted in modern standards such as Digital Video Broadcasting and WiMAX.

This chapter presents my research contributions related to algebraic space-time coding for MIMO systems, focusing on the following problems:

1) Reducing the complexity of decoding, which is one of the main obstacles to the practical implementation of algebraic space-time codes;
2) Characterizing their diversity-multiplexing gain trade-off in the high signal-to-noise ratio regime;
3) Studying their asymptotic performance and gap to capacity when the number of space-time blocks grows to infinity.

Before presenting these contributions, we review the main design criteria and algebraic constructions for space time codes in the next section.

3.1 MIMO systems and design criteria for lattice space-time codes

In a MIMO system with n transmit antennas and m receive antennas, the transmitted signal can be represented in the form of a matrix or space-time block $X = (x_{i,j}) \in M_{n,T}(\mathbb{C})$, where $x_{i,j}$ represents the signal emitted by the antenna $i \in \{1, \dots, n\}$ at time $j \in \{1, \dots, T\}$, and T is the duration of a frame. The received signal is given by

$$Y = \sqrt{\rho} H X + W, \tag{3.1}$$

where $H \in M_{m,n}(\mathbb{C})$ denotes the channel which acts multiplicatively, $W \in M_{m,T}(\mathbb{C})$ is the additive Gaussian noise, and ρ represents the signal-to-noise ratio (SNR).

In this chapter, we consider a simple i.i.d. Rayleigh flat fading model which ignores channel correlations across time and antennas, so that H and W have i.i.d. complex Gaussian entries $h_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0,1)$, $w_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0,1)$[1]. We consider an "open-loop" mode where the transmitter has no channel state information (CSI), with the simplified assumption of perfect CSI at the receiver[2]. A *space-time block code* $\mathcal{C} \subset M_{n,T}(\mathbb{C})$ is a set of matrices satisfying the average power constraint[3]

$$\frac{1}{|\mathcal{C}|} \frac{1}{nT} \sum_{X \in \mathcal{C}} \|X\|^2 \leq 1. \tag{3.2}$$

The *rate* of the code $\mathcal{C}$ is $R = \frac{1}{T} \log |\mathcal{C}|$.

[1] Admittedly, this model is highly simplified, but it already captures some of the trade-offs of space-time coding. The hypothesis of independent fading across antennas is reasonable if there is sufficient spacing among them. While the flat fading assumption generally doesn't hold for wireless channels, in systems employing orthogonal frequency division multiplexing (OFDM), the wideband channel is divided into narrowband channels that are approximately flat.

[2] In practice, the receiver must estimate the channel using training symbols, and the quality of the estimation will depend on the variability of the channel with time.

[3] Practical systems are actually subject to peak power limitations, but characterizing the capacity is more difficult under peak constraints.

Rank and determinant criterion. Maximum likelihood (ML) decoding for this system is given by

$$\hat{X}_{\text{ML}} = \underset{X' \in \mathcal{C}}{\operatorname{argmin}} \|Y - \sqrt{\rho} H X'\|^2 .$$

The probability of error under ML decoding can be estimated using the union bound

$$P_e = \mathbb{P}\{\hat{X} \neq X\} \leq \sum_{X' \neq X} \mathbb{P}\{X \to X'\},$$

where $\mathbb{P}\{X \to X'\} = \mathbb{P}\left\{\left\|Y - \sqrt{\rho} H X'\right\| \leq \left\|Y - \sqrt{\rho} H X\right\|\right\}$ is the pairwise error probability (PEP).
For a fixed channel realization H, the PEP is bounded by the Chernoff bound on the Q-function:

$$P_e(H) = \mathbb{P}\{\hat{X} \neq X \mid H\} \leq \sum_{X' \in \mathcal{C}, X' \neq X} e^{-\rho \left\| H(X - X') \right\|^2}. \tag{3.3}$$

By averaging the PEP over the MIMO Rayleigh channel, the following bound was derived by Tarokh *et al.* [211]:

$$P_e = \int_{M_{m,n}(\mathbb{C})} P_e(H) p(H) dH \leq \sum_{X' \in \mathcal{C}, X' \neq X} \frac{1}{(\det(I_n + \rho (X - X')(X - X')^\dagger))^m}. \tag{3.4}$$

At high SNR, a good approximation for this bound is given by an *inverse determinant sum* over nonzero codewords:

$$P_e \leq \sum_{X' \in \mathcal{C}, X' \neq X} \frac{1}{\rho^{nm} (\det((X - X')(X - X')^\dagger))^m}. \tag{3.5}$$

For finite constellations carved from linear codes, this bound leads to the following criteria:

- *rank criterion*: each nonzero codeword $X \in \mathcal{C}$ should be full-rank in order to achieve the maximum diversity order mn.
- *determinant criterion*: the minimum determinant over all nonzero codewords X should be maximized.

Observe that for the rank criterion to hold, the frame length needs to satisfy the *minimum delay* condition $T \geq n$. Note that even if an (infinite) lattice code satisfies the rank condition, when choosing constellations $\mathcal{C}$ of increasing size in order to transmit more data, the minimum determinant over non-zero codewords might decrease, and might eventually vanish for the infinite code. It is thus important to choose lattice codes with the so called *non-vanishing determinant* (NVD) property .

Diversity-multiplexing gain trade-off. In the high SNR regime with fixed blocklength, the rate-reliability performance of space-time codes is measured by their diversity-multiplexing gain trade-off [238]. The maximum diversity of the MIMO system (3.1) is equal to mn, the maximum number of independent transmit-receive paths. In this regime, it is well known [213] that the ergodic channel capacity $C = \mathbb{E}_H \left[\log \det \left(I_n + \rho H H^\dagger\right)\right]$ scales like $\min(m, n) \log \rho$; that is, the maximum multiplexing gain is $\min(m, n)$. The approach of [238] is to consider the regime in which the rate $R(\rho)$ of the code is a fraction of the capacity.

Definition 3.1 (Diversity-multiplexing gain trade-off) *We will say that a family of codes* $\{C(\rho)\}$ *achieves the* diversity-multiplexing gain trade-off (DMT) *of* spatial multiplexing gain r *and* diversity gain $d(r)$ *if the rate satisfies*

$$\lim_{\rho \to \infty} \frac{R(\rho)}{\log \rho} = r, \tag{3.6}$$

and the average error probability is such that

$$P_e(\rho) \doteq \rho^{-d(r)},$$

where the dotted equality $f(M) \doteq g(M)$ stands for

$$\lim_{M\to\infty} \frac{\log(f(M))}{\log(M)} = \lim_{M\to\infty} \frac{\log(g(M))}{\log(M)}. \tag{3.7}$$

It was shown in [238, 80] that $\forall T \geq n$, the optimal DMT $d_{\max}(r)$ of the MIMO channel model (3.1) is a piecewise linear curve joining the points

$$(r, [(n-r)(m-r)]^+), \quad r \in \mathbb{Z}. \tag{3.8}$$

The intuitive interpretation is that r transmit antennas and r receive antennas are used for multiplexing, leaving $n-r$ transmit antennas and $m-r$ receive antennas available for diversity. Moreover, Elia *et al.* [80] proved that in a MIMO system with n transmit and m receive antennas and minimal delay $T = n$, the non-vanishing determinant property is a sufficient condition for a $2n^2$-dimensional lattice code in $M_n(\mathbb{C})$ to achieve the optimal DMT hen received with an arbitrary number of antennas m.

Lattice space-time codes from division algebras The rank and determinant criteria show that in fading channels, the *multiplicative structure* of the code plays a role in addition to the additive structure. Due to the matrix form of space-time codewords, the theory of non-commutative algebras is a useful tool to build high performance codes with the non-vanishing determinant property [178].

We now review how to build such lattice codes from division algebras. We refer the reader to Reiner's book [194] for the relevant algebraic concepts.

We consider a number field K and a cyclic field extension E/K of degree n with Galois group $\mathrm{Gal}(E/K) = \langle\sigma\rangle$. Consider the cyclic algebra

$$\mathcal{D} = (E/K, \sigma, \gamma) = E \oplus uE \oplus u^2E \oplus \cdots \oplus u^{n-1}E, \tag{3.9}$$

where $u \in \mathcal{D}$ is an auxiliary generating element subject to the relations $xu = u\sigma(x)$ for all $x \in E$ and $u^n = \gamma \in \mathbb{Q}^*$. Considering $\mathcal{D}$ as a right vector space over E, every element $x = x_0 + ux_1 + \cdots + u^{n-1}x_{n-1}$ admits the following *left regular representation* as a matrix $\psi(x) \in M_n(E)$:

$$\begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \cdots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & & \gamma\sigma^{n-1}(x_2) \\ x_2 & \sigma(x_1) & \sigma^2(x_0) & & \gamma\sigma^{n-1}(x_3) \\ \vdots & & & & \vdots \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \cdots & \sigma^{n-1}(x_0) \end{pmatrix}. \tag{3.10}$$

The mapping ψ is an injective K-algebra homomorphism that allows us to identify $\mathcal{D}$ with its image in $M_n(\mathbb{C})$. We assume that $\mathcal{D}$ is a *division algebra*, namely that every nonzero element in $\mathcal{D}$ is invertible. Consequently, every non-zero matrix in the set $\psi(\mathcal{D}) \subset M_n(\mathbb{C})$ is invertible, but $\psi(\mathcal{D})$ is dense and therefore not directly suitable for space-time coding. In order to obtain a discrete subset of codewords, one can choose an $\mathcal{O}_E$*-order* Γ in $\mathcal{D}$, namely a subring of $\mathcal{D}$, having the same identity element as $\mathcal{D}$, such that Γ is a finitely generated module over the ring of integers $\mathcal{O}_E$ of E and generates $\mathcal{D}$ as a linear space over E.

The determinant of the regular representation $X = \psi(x)$ of an element $x \in \Gamma$ is its reduced norm:

$$\det(\psi(x)) = N_{\mathcal{D}/K}(x) \in \mathcal{O}_K.$$

In particular, if $K = \mathbb{Q}$ or $K = \mathbb{Q}(\sqrt{-d})$, the ring of integers $\mathcal{O}_K$ is discrete and $|\det(\psi(x))| \geq 1$ for all $x \in \Gamma$. More generally, we can consider codes which of the form $\psi(\Gamma\alpha)$, where $\Gamma\alpha$ is a principal ideal of the order Γ. Any finite signal constellation carved from these codes will have the *non-vanishing determinant* property.

Examples of non-vanishing determinant codes built using the previous construction include the Alamouti Code [7], the Golden Code [19] and the Perfect Codes [179]. Elia *et al.* [80] showed that such full-dimensional division algebra based codes are also DMT optimal, and gave a general construction for DMT-achieving $2n^2$-dimensional

lattice codes in $M_n(\mathbb{C})$.

3.2 Low-complexity decoding of algebraic space-time codes

While full-dimensional algebraic space-time codes are optimal in terms of DMT under maximum likelihood decoding, their decoding complexity is a significant challenge to their practical implementation.

Decoding MIMO space-time block codes amounts to solving the closest vector problem (CVP) in a finite subset of the lattice generated by the channel matrix, whose dimension grows with the number of antennas. Maximum likelihood decoding of these codes can be performed using the sphere decoding algorithm [220], although its complexity is in general exponential in the lattice dimension [122, 123]. Low-complexity decoding techniques such as zero-forcing (ZF) and minimum mean square error (MMSE) decoding do not preserve the diversity of the system. Their performance can be improved by a pre-processing step using lattice reduction algorithms. For instance, pre-processing using the LLL algorithm [148], which has polynomial complexity, allows to achieve the optimal receive diversity order [209]. It was also shown that lattice reduction (LR) aided regularized lattice decoding preserves DMT-optimality [121].

3.2.1 Algebraic reduction for space-time codes based on division algebras

As we have seen in Section 3.1, algebraic space-time codes are endowed with an additional multiplicative structure through the left regular representation of division algebras. It is natural to ask whether this extra structure can be exploited to improve closest point search or lattice reduction.

In a joint work with J.-C. Belfiore and G. Rekaya Ben-Othman [J5], we proposed *algebraic reduction*, a pre-processing method that exploits the multiplicative structure of the code. The main idea is to absorb part of the channel into the code, by approximating the channel matrix by a unit in the corresponding division algebra. This extends the algebraic reduction technique in [195] for lattice constructions based on number fields for single-antenna fading channels, which was shown to achieve the optimal diversity order together with zero-forcing (ZF) detection, and to outperform LLL reduction followed by ZF detection in high dimension.

We consider space-time block codes which can be represented in the form $\psi(\Gamma\alpha)$, where Γ is a maximal order of a division algebra $\mathcal{D}$ of index n over $K = \mathbb{Q}(i)$, and ψ is the left regular representation (3.10).

Example 3.2 (The Golden code) The Golden code [19] is a full rate, full diversity space-time code for 2×2 MIMO systems which was selected as an optional profile in the WiMAX standard. It is based on the left regular representation (3.10) of the quaternion division algebra $\mathcal{D} = (\mathbb{Q}(i,\theta)/\mathbb{Q}(i), \sigma, i)$, where θ is the golden number, and $\sigma : \mathbb{Q}(i,\theta) \to \mathbb{Q}(i,\theta)$ is such that $\sigma(\theta) = 1 - \theta$ and leaves the elements of $\mathbb{Q}(i)$ fixed.
More precisely, the infinite code is given by $\psi(\Gamma\alpha)$, where Γ is a maximal order of $\mathcal{D}$, and $\alpha = 1 + i\sigma(\theta)$. Its codewords have the form

$$X = \frac{1}{\sqrt{5}} \begin{pmatrix} \alpha(s_1 + s_2\theta) & \alpha(s_3 + s_4\theta) \\ \sigma(\alpha)i(s_3 + s_4\sigma(\theta)) & \sigma(\alpha)(s_1 + s_2\sigma(\theta)) \end{pmatrix}, \tag{3.11}$$

where $s_1, s_2, s_3, s_4 \in \mathbb{Z}[i]$ are QAM symbols.

The lattice representation of the code can be obtained as follows.

Remark 3.3 (Vectorization of matrices) Let ϕ be the function $M_n(\mathbb{C}) \to \mathbb{C}^{n^2}$ that vectorizes matrices column by column. The left multiplication function $M_n(\mathbb{C}) \to M_n(\mathbb{C})$ that maps B to AB induces a linear mapping represented by the block diagonal matrix $A_l = I_n \otimes A \in M_{n^2}(\mathbb{C})$.

Notation 3.4 (Lattice point representation) Let $\{w_1, w_2, \ldots, w_{n^2}\}$ be a basis of $\psi(\Gamma\alpha)$ as a $\mathbb{Z}[i]$-module. Every codeword X can be written as

$$X = \sum_{i=1}^{n^2} s_i w_i, \quad \mathbf{s} = (s_1, s_2, \ldots, s_{n^2})^t \in (\mathbb{Z}[i])^{n^2}.$$

Let Φ be the matrix whose columns are $\phi(w_1), \phi(w_2), \ldots, \phi(w_{n^2})$. Then the lattice point corresponding to X is

$$\mathbf{x} = \phi(X) = \sum_{i=1}^{n^2} s_i \phi(w_i) = \Phi \mathbf{s} \tag{3.12}$$

We denote by Λ the $\mathbb{Z}[i]$-lattice with generator matrix Φ.

Definition 3.5 (Unit group) *Let Γ be an order in a division algebra $\mathcal{D}$. The* unit group Γ^* *of Γ consists of nonzero elements $x \in \Gamma$ such that their multiplicative inverse $x^{-1} \in \Gamma$.*
If the center K of the division algebra $\mathcal{D}$ is $\mathbb{Q}$ or an imaginary quadratic field, then

$$\Gamma^* = \{x \in \mathcal{D} \ : \ |\det(\psi(x))| = 1\}.$$

We will consider the subgroup of units of norm 1

$$\Gamma^1 = \{x \in \mathcal{D} \ : \ \det(\psi(x)) = 1\}.$$

Remark 3.6 (Units and unimodular transformations) Consider the generator matrix Φ for the representation of $\psi(\Gamma\alpha)$ as a $\mathbb{Z}[i]$-lattice in (3.12).
If $U \in \psi(\Gamma^*)$, then

$$U_l \Phi = \Phi T_U$$

where $T_U \in M_{n^2}(\mathbb{C})$ is unimodular (with elements in $\mathbb{Z}[i]$).

Algebraic reduction. Suppose the received signal is of the form

$$Y = HX + W,$$

where H has i.i.d. entries $h_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ and W has i.i.d. entries $w_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$.

Assuming perfect CSI at the receiver, the latter can perform the normalization

$$Y' = \frac{Y}{\sqrt{\det(H)}} = H_1 X + W', \tag{3.13}$$

where $\det(H_1) = 1$. We will approximate H_1 with a unit $U \in \psi(\Gamma^1)$. More precisely, we want to obtain a decomposition of the form

$$H_1 = EU, \tag{3.14}$$

where $U \in \psi(\Gamma^1)$ and $E \in \mathrm{SL}_n(\mathbb{C})$ is a suitable approximation error.

Applying the matrix vectorization mapping ϕ to both sides of equation (3.13), we obtain

$$\mathbf{y}' = E_l U_l \Phi \mathbf{s} + \mathbf{w}',$$

where $E_l = I_n \otimes E$, $U_l = I_n \otimes U$, Φ is the generator matrix of the code lattice defined in (3.12), and $\mathbf{s} \in \mathbb{Z}[i]^{n^2}$ is the vector of QAM information symbols. By Remark 3.6,

$$\mathbf{y}' = E_l \Phi T_U \mathbf{s} + \mathbf{w}' = E_l \Phi \mathbf{s}' + \mathbf{w}', \qquad \mathbf{s}' \in \mathbb{Z}[i]^{n^2}.$$

In order to decode, we can apply ZF detection to estimate $\mathbf{s}'$:

$$\hat{\mathbf{s}}' = \left\lfloor \Phi^{-1} E^{-1} \mathbf{y}' \right\rceil = \left\lfloor \mathbf{s}' + \Phi^{-1} E_l^{-1} \mathbf{w}' \right\rceil = \lfloor \mathbf{s}' + \mathbf{n} \rceil .$$

Finally, we can recover the estimate of the initial signal $\hat{\mathbf{s}} = T_U^{-1} \hat{\mathbf{s}}'$. Note that the variance σ_i^2 of the i-th component

of the noise $\mathbf{n}$ is bounded by

$$\sigma_i^2 \leq \frac{n^2\sigma^2}{|\det(H)|^{\frac{2}{n}}} \left\|\Phi^{-1}\right\|_F^2 \left\|E^{-1}\right\|^2 \quad \forall i = 1, \ldots, n^2.$$

Consequently, in order to maximize performance, we find the following criterion to choose U: the norm $\left\|E^{-1}\right\|_F = \left\|UH_1^{-1}\right\|_F$ should be minimized. If this norm is bounded by a constant, then this method achieves the optimal receive diversity order:

Proposition 3.7 *Suppose that there exists a reduction algorithm which, given $H_1 \in \mathrm{SL}_n(\mathbb{C})$, outputs a decomposition of the form (3.14) such that*

$$\left\|E^{-1}\right\|_F^2 \leq C_\Gamma. \tag{3.15}$$

where the constant C_Γ depends only on Γ.
Then the algebraic reduction followed by ZF detection achieves the full receive diversity n.

Approximation with a unit in the quaternion case In the quaternion case (i.e. when $n = 2$, such as in the case of the Golden Code), we propose an explicit algorithm to find U, by exploiting the fact that Γ^1 is a cocompact discrete subgroup of $\mathrm{SL}_2(\mathbb{C})$. We consider the action of $\mathrm{SL}_2(\mathbb{C})$ on the hyperbolic 3-space

$$\mathbb{H}^3 = \{(z, r) \mid z \in \mathbb{C},\ r \in \mathbb{R}^+\}$$

with the hyperbolic distance ρ such that $\cosh\rho(P, P') = 1 + \frac{d(P,P')}{2rr'}$ for $P, P' \in \mathbb{H}^3$, where $d(P, P')^2 = |z - z'|^2 + (r - r')^2$ is the squared Euclidean distance. It will be enough to consider the action on the point $J = (0, 0, 1)$ given by

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \quad A(J) = \left(\frac{\Re(b\bar{d} + a\bar{c})}{|c|^2 + |d|^2}, \frac{\Im(b\bar{d} + a\bar{c})}{|c|^2 + |d|^2}, \frac{1}{|c|^2 + |d|^2} \right).$$

Note that given $A \in \mathrm{SL}_2(\mathbb{C})$, its Frobenius norm is equal to

$$\|A\|_F^2 = 2\cosh\rho(J, A(J)).$$

Therefore, $\left\|H_1U^{-1}\right\|_F$ is small if and only if $U^{-1}(J)$ is close to $H_1^{-1}(J)$ in hyperbolic distance.

By Poincaré's Polyhedron Theorem [81], a fundamental region for the action of Γ^1 on $\mathbb{H}^3$ is given by the Dirichlet fundamental polyhedron of Γ^1 (with center J), which is defined as the intersection of all the bisectors corresponding to non-trivial elements:

$$\mathcal{P} = \bigcap_{\substack{g \in \Gamma^1, \\ g \neq \mathbb{1}}} D_g(J) \tag{3.16}$$

where

$$D_g(J) = \{P \in \mathbb{H}^3 \mid \rho(J, P) \leq \rho(g(J), P)\}. \tag{3.17}$$

The hyperbolic space $\mathbb{H}^3$ is tiled by the copies $g(\mathcal{P})$, $g \in \Gamma^1$. The hyperbolic volume of $\mathcal{P}$ is given by the Tamagawa Volume Formula

$$\mathrm{Vol}(\mathcal{P}_{\Gamma^1}) = \frac{1}{4\pi^2}\zeta_F(2)\,|d_F|^{\frac{3}{2}} \prod_{p \mid d(\Gamma/\mathcal{O}_F)} (N_p - 1). \tag{3.18}$$

In the previous formula, ζ_K denotes the *Dedekind zeta function*[4] relative to the field K, d_K is the discriminant of K, $d(\Gamma/\mathcal{O}_K)$ is the $\mathcal{O}_K$-discriminant of Γ, p varies among the primes of $\mathcal{O}_K$, and $N_p = [\mathcal{O}_K : p\mathcal{O}_K]$, where $\mathcal{O}_K$ is the ring of integers of K.

[4]The Dedekind zeta function is defined as $\zeta_K(s) = \sum_I ([\mathcal{O}_K : I])^{-s}$, where I varies among the proper ideals of $\mathcal{O}_K$.

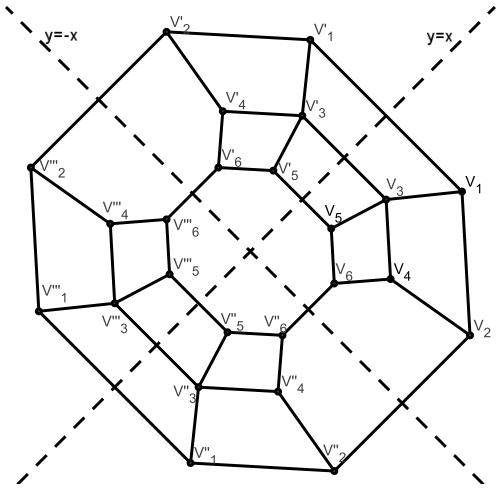


Figure 3.1: The projection of the polyhedron $\mathcal{P}$ on the plane $\{r = 0\}$.

Algorithm 1: The unit search algorithm

input: $h_1 \in \mathrm{SL}_2(\mathbb{C})$.

$h = h_1$, $\bar{u} = \mathbb{1}$, $i_0 = 0$
repeat
  Compute $h^{-1}(J) = (x, y, r)$
  $d_0 = 2\cosh\rho(h^{-1}(J), J)$
  for $i = 1, \ldots, 2r$ **do**
    $d_i = 1 + \frac{(x-x_i)^2+(y-y_i)^2+(r-r_i)^2}{2rr_i}$
  end
  $i_0 = \underset{i\in\{0,1,\ldots,2r\}}{\operatorname{argmin}} d_i$.
  $\bar{u} \leftarrow \bar{u}u_{i_0}$, $h \leftarrow hu_{i_0}$
until $i_0 = 0$
output: $\hat{u} = \bar{u}^{-1}$ is the chosen unit.

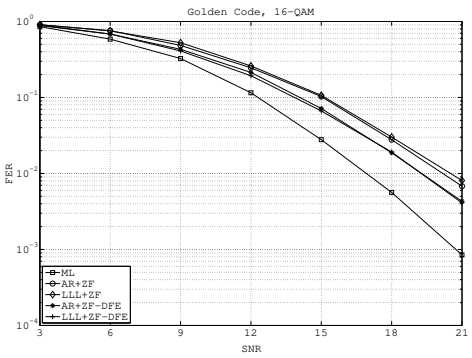


Figure 3.2: Comparison of algebraic reduction and LLL reduction using MMSE-GDFE preprocessing combined with ZF or ZF-DFE decoding with 16-QAM constellations.

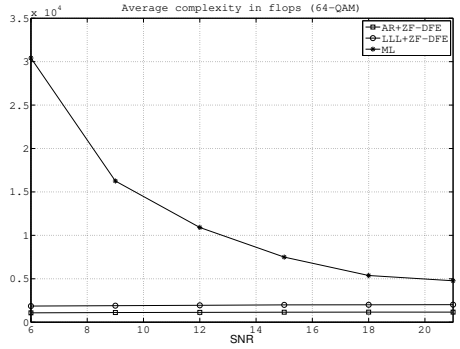


Figure 3.3: Comparison of the average complexity in floating point operations of Sphere Decoding, complex LLL-ZF-DFE decoding and simplified AR-ZF-DFE decoding (both using MMSE-GDFE preprocessing) using 64-QAM constellations.

In the case of the Golden code, we are able to obtain a complete characterization of the unit group Γ^1: we compute a set of 8 generators of the group and a complete set of relations between the generators, we characterize the polyhedron $\mathcal{P}$ (see Figure 3.1) and compute its volume explicitly.

Given a normalized channel matrix $H_1 \in \mathrm{SL}_2(\mathbb{C})$, we describe an algorithm to find a unit $\hat{U}$ such that $H_1^{-1}(J) \in \hat{U}(\mathcal{P})$ (Algorithm 1). Let $U_1, \ldots, U_r$ be the generators of Γ^1 and $U_{r+1} = U_1^{-1}, \ldots, U_{2r} = U_r^{-1}$ their inverses. The neighboring polyhedra of $\mathcal{P}$ are all of the form $U_i(\mathcal{P})$, $i = 1, \ldots, 2r$.
The idea is to begin the search from $\mathcal{P}$ and the neighboring polyhedra, corresponding to the generators of the group and their inverses, and choose U_i such that $U_i(J)$ is the closest to $H_1^{-1}(J)$. Since U_i is an isometry of $\mathbb{H}^3$, at the next step we can apply U_i^{-1} and start again the search of the $U_{i'}$ that gives the closest point to ${U_i}^{-1}H_1^{-1}(J)$. With this strategy we only need to perform $2r$ comparisons at each step of the search.

Figure 3.2 shows the performance of algebraic reduction (AR) followed by ZF and ZF-DFE decoding for the Golden Code using 16-QAM constellations. With MMSE-GDFE preprocessing [171], algebraic reduction-aided decoding is within 2.3 dB of the ML using ZF-DFE detection, at the FER of 10^{-2}. A comparison of algebraic reduction and LLL reduction followed by ZF-DFE detection evidences that the two methods have nearly identical performance.

Figure 3.3 shows the average complexity in floating point operations of the AR-ZF-DFE decoder using 64-QAM constellations in an i.i.d. Rayleigh fading channel. As shown in the figure, the sphere decoding complexity

is very high in the low-to-moderate SNR regime, while the complexity of LLL-ZF-DFE and AR-ZF-DFE doesn't depend on the SNR.
Numerical simulations evidence that AR-ZF-DFE provides a complexity saving of about 70% with respect to LLL-ZF-DFE[5]. We expect the complexity savings to be higher for slow fading channels:

Remark 3.8 (Advantage of the algebraic reduction in the case of slow fading channels) If the channel varies slowly from one time block to the next, it is reasonable to expect that the polyhedron $\hat{U}(\mathcal{P})$ containing $H_1^{-1}(J)$ at the time t will be the same, or will be adjacent, to the polyhedron chosen at the time $t-1$. Thus, algebraic reduction requires only a slight adjustment of the search at each step. On the contrary, classical lattice reduction-aided decoding techniques require a full lattice reduction at each time block.

Follow-up work and open problems

Generalization to other space-time codes based on quaternion algebras. Both the performance and the complexity of algebraic reduction depend in general on the structure of the unit group, and so it might not be advantageous in every case. Indeed, the *quality* of approximation by a unit is related to the diameter $R_{\max}$ of the fundamental polyhedron, while the *speed* of the algorithm depends on the cardinality r of a minimal set of generators for the group. Unfortunately, the structure of the unit group can be very complex in general [181]. Finding good space-time codes from quaternion algebras such that r and $R_{\max}$ are small is an open problem; [10] considered the problem of finding suitable quaternion algebras so that the volume (3.18) is minimized.

Generalization to higher-dimensional space-time codes. The principle of algebraic reduction as well as the proof that it achieves the maximal receive diversity order hold in general for space-time codes based on maximal orders of division algebras of index n over $\mathbb{Q}(i)$ provided that the multiplicative approximation error can be bounded as in (3.15). However, finding the generators of the unit group in an order for general division algebras is a difficult problem in computational algebra [130].

3.2.2 Decoding by embedding

The algebraic reduction technique presented in the previous section applies only to space-time codes with a special algebraic structure. Moreover, it does not provide a performance gain over LLL reduction, although it offers a complexity gain. In this section, we present a more general decoding technique for both coded and uncoded MIMO systems based on lattice reduction: the *embedding technique* [J4],[J6]. The core idea is to embed the basis of the decoding lattice and the received vector into an $(n+1)$-dimensional lattice, in order to convert an n-dimensional instance of the closest vector problem (CVP) is converted into an $(n+1)$-dimensional instance of the shortest (nonzero) vector problem (SVP). The LLL algorithm can recover the transmitted vector when the norm of the noise vector is small compared to the minimum distance λ_1 of the lattice. This condition corresponds to a variant of the CVP known as *Bounded Distance Decoding* (BDD). More precisely, BDD_η (with $\eta \leq 1/2$) is a special instance of CVP where the norm of the noise vector (or, equivalently, the distance from the target vector to the lattice) is less than $R = \eta \cdot \lambda_1$. The radius R is referred to as the (correct) *decoding radius* of the algorithm.

We consider the $n \times m$ flat fading MIMO system model

$$Y = HX + W,$$

where the entries of the channel gain matrix H are normalized to unit variance, and the entries of W are i.i.d. complex Gaussian with variance σ^2. The codewords X satisfy the average power constraint $E[\|X\|_{\mathrm{F}}^2/T] = 1$. Hence, the SNR at each receive antenna is $1/\sigma^2$.

When a lattice space-time block code is employed, the QAM information vector $\mathbf{s}$ is multiplied by the generator matrix Φ of the encoding lattice. The $n \times T$ codeword matrix X is defined by column-wise stacking of consecutive

[5] Our simulations refer to the complex version of LLL-ZF-DFE presented in [92], which already obtains a complexity saving of about 50% with respect to real LLL-ZF-DFE.

n-tuples of the vector $\Phi\mathbf{s} \in \mathbb{C}^{nT}$. By column-by-column vectorization, the received signal can be expressed as

$$\mathbf{y} = (I_T \otimes H)\,\Phi\mathbf{s} + \mathbf{w} \tag{3.19}$$

When $T = 1$ and $\Phi = I_n$, equation (3.19) reduces to the model for uncoded MIMO communication $\mathbf{y} = H\mathbf{s}+\mathbf{w}$. After separating real and imaginary parts, we obtain the equivalent $N \times M$ real-valued MIMO system model

$$\mathbf{y} = B\mathbf{x} + \mathbf{n}, \tag{3.20}$$

where $N = 2nT$, $M = 2mT$, and where $B \in \mathbb{R}^{M\times N}$ can be interpreted as the basis matrix of the decoding lattice.

The principle of Kannan's embedding technique [126] is to embed the basis B and the received vector $\mathbf{y}$ into a higher dimensional lattice. More precisely, we consider the following $(M + 1) \times (N + 1)$ basis matrix:

$$\widetilde{B} = \begin{bmatrix} B & -\mathbf{y} \\ \mathbf{0}_{1\times N} & t \end{bmatrix} \tag{3.21}$$

where $t > 0$ is a parameter to be determined, which we refer to as the embedding parameter.
The strategy is to reduce CVP to SVP in the following way. For a suitable choice of t and for sufficiently small noise norm, the vectors $\mathbf{v} = \pm[(B\mathbf{x} - \mathbf{y})^T \quad t]^T$ are the shortest vectors in the lattice $\mathcal{L}(\widetilde{B})$ generated by B. Thus an SVP algorithm will find $\mathbf{v}$, and the message $\mathbf{x}$ can be recovered from the coordinates of this vector in the basis $\widetilde{B}$:

$$\text{if } \mathbf{v} = \widetilde{B}\begin{pmatrix}\mathbf{x}'\\ 1\end{pmatrix} = \begin{pmatrix}B\mathbf{x}' - \mathbf{y}\\ t\end{pmatrix}, \text{ then } \hat{\mathbf{x}} = \mathbf{x}'. \tag{3.22}$$

Decoding radius of the embedding technique. In [J4], we showed that the LLL algorithm with parameter α can be used to find the shortest vector in the lattice $\mathcal{L}(\widetilde{B})$, and that the correct decoding radius is lower bounded by

$$\frac{1}{2\sqrt{2}\alpha^{n-\frac{1}{2}}}\lambda_1(B). \tag{3.23}$$

In [160], it is proven that by choosing $t = \mathrm{dist}(\mathbf{y}, \mathcal{L}(B))$, the embedding technique allows one to reduce $\mathsf{BDD}_{1/(2\gamma)}$ to uSVP_γ. We show that one can achieve the same correct decoding radius by setting $t = \frac{1}{2\gamma}\lambda_1(B)$, thus bypassing the assumption from [160] that $\mathrm{dist}(\mathbf{y}, \mathcal{L}(B))$ is known.

Theorem 3.9 (Decoding Radius of Embedding) *Applying* uSVP_γ *($\gamma \geq 1$) to the extended lattice (3.21) with parameter t ($0 < t < \lambda_1(B)/\gamma$) guarantees a correct decoding radius*

$$R_{\textit{uSVP-Emb}} \geq \sqrt{\frac{t}{\gamma}\lambda_1(B) - t^2}. \tag{3.24}$$

Setting $t = \frac{1}{2\gamma}\lambda_1(B)$ maximizes this lower bound. This gives:

$$R_{\textit{uSVP-Emb}} \geq \frac{1}{2\gamma}\lambda_1(B). \tag{3.25}$$

As the LLL algorithm can solve uSVP_γ with $\gamma = \alpha^{\frac{n}{2}}$ for the basis (3.21) of dimension $N + 1$, the correct decoding radius using LLL satisfies

$$R_{\text{uSVP-Emb}} \geq \frac{1}{2\alpha^{\frac{n}{2}}}\lambda_1(B) \tag{3.26}$$

by choosing $t = \frac{1}{2\alpha^{\frac{n}{2}}}\lambda_1(B)$. This decoding radius improves the bound (3.23) from [J4]. However, it can still be improved. The reason is that the estimate $\gamma = \alpha^{\frac{n}{2}}$ is pessimistic for uSVP_γ. In fact, the quantity $\alpha^{\frac{n}{2}}$ is just the approximation factor for the approximate SVP achieved by LLL. Any algorithm solving SVP_γ necessarily solves uSVP_γ, but the latter might be an easier problem.

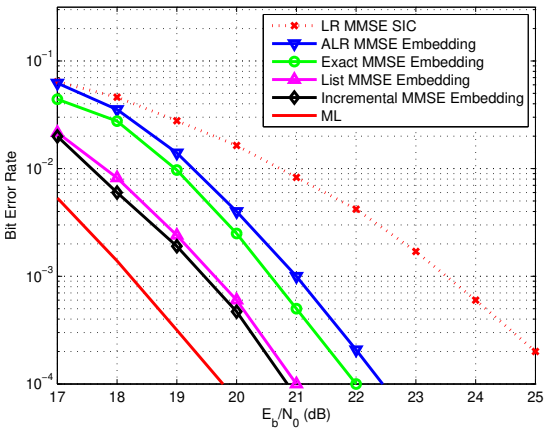


Figure 3.4: Bit error rate vs. average SNR per bit for the uncoded 10×10 system using 64-QAM.

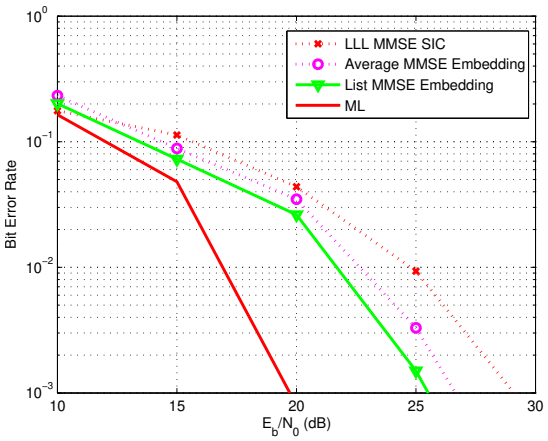


Figure 3.5: Bit error rate vs. average SNR per bit for the 4×4 perfect code using 64-QAM.

Lemma 3.10 (Improved bound to solve uSVP_γ with LLL) *In an n-dimensional lattice, the LLL algorithm can solve uSVP_γ for $\gamma = \sqrt{\bar{\gamma}_{n-1}}\alpha^{\frac{n}{4}}$, where $\bar{\gamma}_n = \max_{1 \le i \le n} \gamma_i$. Here γ_i denotes the Hermite constant (2.1) in dimension i.*

DMT-optimality of the embedding technique. On the MIMO communications front, we prove that bounded distance decoding of the regularized lattice is DMT-optimal over Rayleigh fading channels.

We suppose for the sake of simplicity that $M = N$. We consider the equivalent normalized channel model where the noise variance is equal to 1:

$$\mathbf{y}' = B'\mathbf{x} + \mathbf{n}',$$

where $B' = \sqrt{\rho}B$, $n_i' = \sqrt{\rho}n_i \sim \mathcal{N}_{\mathbb{R}}(0,1)$, $\forall i = 1, \ldots, n$. Here $\rho = \frac{1}{\sigma^2}$ denotes the SNR. Moreover, we consider the equivalent *regularized system* [121]

$$\mathbf{y}_1 = R\mathbf{x} + \mathbf{n_1}, \tag{3.27}$$

where

$$\begin{pmatrix} B' \\ I_n \end{pmatrix} = QR, \quad \mathbf{y}_1 = Q^\dagger \begin{pmatrix} \mathbf{y}' \\ 0_{n\times 1} \end{pmatrix}.$$

From the point of view of receiver architecture, this amounts to performing left preprocessing before decoding, by using a maximum mean square error generalized decision-feedback equalizer (MMSE-GDFE) [79].

Theorem 3.11 *For any constant $\eta > 0$, any decoding technique which always provides a solution for the regularized BDD_η is DMT-optimal.*

Remark 3.12 This represents a nontrivial extension of the analysis in [121] for γ-approximation algorithms of CVP. Indeed, γ-approximate algorithms are a special case of BDD: it is easy to see that any decoding technique which provides a CVP_γ solution $\hat{\mathbf{x}}$ to (3.27) is also able to solve $\mathsf{BDD}_{\frac{1}{2\gamma}}$. In fact, suppose that $\mathbf{y}_1$ is such that $\mathrm{dist}(\mathbf{y}_1, R) < \frac{1}{2\gamma}\lambda_1(R)$. Then the CVP_γ solution $\hat{\mathbf{x}}$ satisfies

$$\|\mathbf{y}_1 - R\hat{\mathbf{x}}\| < \gamma \min_{\mathbf{x}\in\mathbb{Z}^n} \|\mathbf{y}_1 - R\mathbf{x}\| = \gamma \mathrm{dist}(\mathbf{y}_1, \mathcal{L}(R)) < \frac{\lambda_1(R)}{2},$$

so that $\hat{\mathbf{x}}$ is the optimal solution of (3.27). However, the converse is apparently not true, that is, BDD does not necessarily provide CVP_γ solutions for all $\mathbf{y}_1$.

Performance of the embedding technique. Fig. 3.4 shows the bit error rate for an uncoded MIMO system with $n_T = n_R = 10$, 64-QAM. We found that the list versions of the embedding technique achieves near-optimum performance in this setting; the SNR loss is about 1 dB.

Fig. 3.5 shows the achieved performance of embedding decoding for the 4×4 Perfect code [179] using 64-QAM. The decoding lattices have dimension 32. The list version of embedding enjoys 3.5 dB gain over LLL reduction followed by successive interference cancellation.

3.3 Diversity-multiplexing trade-off of asymmetric space-time codes

As we have seen in Section 3.1, for full $2n^2$ dimensional lattice space-time codes in $M_n(\mathbb{C})$, the NVD condition is a sufficient condition for DMT-optimality. This criterion was generalized by Tavildar and Viswanath [212] who introduced the notion of *approximately universal* codes, such that the product of the smallest m singular values of any non-zero matrix in a $2nm$-dimensional lattice $L \subset M_n(\mathbb{C})$ stays above some fixed constant. They showed that such codes achieve the optimal DMT curve in the $n \times m$ MIMO channel. In the case where $n = m$, this criterion coincides with the NVD condition.

When receiving a full $2n^2$-dimensional space-time lattice code with minimum delay ($T = n$) with $m < n$ antennas, the dimension of the receiver space is only $2mT = 2nm$ and so the image of the infinite lattice is no longer a lattice, but a dense set of points. Thus the standard sphere decoding algorithm [220] cannot be employed, although special techniques such as generalized sphere decoding have been proposed [61]. Therefore, when the number of receive antennas m is smaller than the number of transmit antennas n, it is desirable to use *asymmetric space-time codes* that are $2nm$-dimensional, which represent the "best fit" for the $n \times m$ MIMO channel. One natural question is whether these codes can be DMT-optimal assuming some conditions, such as NVD. The main interest in non-full dimensional codes, such as the codes arising from division algebras with center $K = \mathbb{Q}$, is that such codes are fast-decodable[6] [218].

Before our work, the only available general criterion for DMT-optimality was the approximate universality criterion given in [212], and for $m < n$ no asymmetric codes satisfying this condition were known except in the case $m = 1$. It was also known that there are space-time codes that are DMT optimal despite not satisfying the approximate universality criterion [217]. This motivated our search for a more general and easily applicable DMT criterion. This work is in collaboration with R. Vehkalahti.

Approach based on point counting in Lie groups [J7],[C18]. The minimum determinant criterion (see Section 3.1) focuses on the worst-case pairwise error probability in the sum (3.5), and does not consider the global distribution of codewords. On the other hand, the DMT takes into account the overall error probability, but is too coarse for practical code design. For instance, while all full-rate division-algebra codes are DMT-optimal, their actual performances can be very different.

Our first approach was to study directly the asymptotic behavior of inverse determinant sums of the form (3.5), which can be seen as an intermediate concept between the former two [J7]. We consider lattice codes arising from the left regular representation of an order Γ in a division algebra $\mathcal{D} = (E/K, \sigma, \gamma)$ and its left regular representation ψ (see Section 3.1). In order to have multiplexing gain r, we take spherical codebooks of the form $\mathcal{C}(M) = \frac{1}{M}(\psi(\Gamma) \cap \mathcal{B}_M)$, where $\mathcal{B}_M$ denotes the ball of radius M in $M_n(\mathbb{C})$ with respect to the Frobenius norm, and where the radius M grows asymptotically with the SNR as

$$M = \rho^{\frac{rn}{k}}, \quad k = \dim_{\mathbb{Z}}\Gamma.$$

In the case of minimum delay codes with $n = T$, the PEP bound (3.5) takes the form

$$P_e \leq \sum_{\substack{x \in \Gamma \setminus \{0\} \\ \|\psi(x)\|_F \leq 2M}} \frac{1}{|\det(\psi(x))|^{2m}}$$

[6] For example, over a 2×1 MIMO channel, both the Alamouti code [7] and the Golden Code [19] are DMT-optimal. However the Alamouti code is fast-decodable, while the Golden Code is hard to decode with one receive antenna, since the image of the code is an 8-dimensional lattice projected onto a 4-dimensional space.

We show that the behavior of this inverse determinant sum is essentially determined by the behavior of the sum over the unit group Γ^* (see Definition 3.5.) From the ideal theory of orders we have that if $x\Gamma = y\Gamma$, then x and y must differ by a unit, i.e. $y = xu$ for some $u \in \Gamma^*$. Therefore we can write

$$\sum_{\substack{x\in\Gamma\setminus\{0\}\\ \|\psi(x)\|_F\leq M}} \frac{1}{|\det(\psi(x))|^{2m}} = \sum_{x\in\mathcal{I}(M)} \frac{|\psi(x\Gamma^*)\cap\mathcal{B}_M|}{|\det(\psi(x))|^{2m}},$$

where $\mathcal{I}(M)$ is a collection of non-zero elements $x \in \Lambda$, $\|\psi(x)\|_F \leq M$, each generating a separate (right) ideal. In order to make our bound more treatable, we consider the subgroup Γ^1 of *units of norm* 1 (see again Definition 3.5) which is known to have finite index in Γ^* , i.e. $[\Gamma^* : \Gamma^1] < \infty$ [129]. Thanks to this property, one can show that

$$|\psi(x\Gamma^*)\cap\mathcal{B}_M| \leq K\left|\psi(\Gamma^1)\cap\mathcal{B}_M\right|.$$

for some constant K, independent of x and M. Finally, we obtain the bound

$$\sum_{\substack{x\in\Gamma\setminus\{0\}\\ \|\psi(x)\|_F\leq M}} \frac{1}{|\det(\psi(x))|^{2m}} \leq K\left|\psi(\Gamma^1)\cap\mathcal{B}_M\right| \sum_{x\in\mathcal{I}(M)} \frac{1}{|\det(\psi(x))|^{2m}} \tag{3.28}$$

We show that the sum over ideals in (3.28) has logarithmic growth (and is thus negliglible from the point of view of DMT): there exist constants K_1, K_2 independent of M such that

$$\sum_{x\in\mathcal{I}(M)} \frac{1}{|\det(\psi(x))|^{2m}} \leq \sum_{\substack{x\in\mathcal{I}(M)\\ [\Gamma:x\Gamma]<M^k}} \frac{1}{[\Gamma : x\Gamma]^s} \leq \zeta_\Gamma(s, M^k) \leq K_1(\log M)^{K_2}.$$

Here $\zeta_\Gamma(s, M^k)$ denotes the truncated *Hey zeta function* of Γ

$$\zeta_\Gamma(s) = \sum_{I\in\mathbf{I}_\Gamma} \frac{1}{[\Gamma : I]^s},$$

where $\Re(s) > 1$ and $\mathbf{I}_\Gamma$ is the set of right ideals of Γ [34], and the exponent $s = \frac{m}{n}$ if $K = \mathbb{Q}(\sqrt{-d})$, $s = \frac{m}{2n}$ if $K = \mathbb{Q}$.

The final step in order to characterize the behaviour of the inverse determinant sum (3.28) is to study the term $\left|\psi(\Gamma^1)\cap\mathcal{B}_M\right|$, i.e. to count the units inside the ball $\mathcal{B}_M$. Fortunately, this is a well-known problem in the ergodic theory of Lie groups. In fact, the group of units Γ^1 can be seen as a subgroup of a suitable Lie group.
Before stating this result, we need a technical definition to distinguish two types of $\mathbb{Q}$-central division algebras.

Definition 3.13 (Ramification of $\mathbb{Q}$-central division algebras) *Let $\mathcal{D}$ be a $\mathbb{Q}$-central division algebra of index n. We say that $\mathbb{D}$ is* ramified at the infinite place *if $\mathbb{D}\otimes_{\mathbb{Q}}\mathbb{R} \simeq M_{n/2}(\mathbb{H})$. If it is not, then $\mathbb{D}\otimes_{\mathbb{Q}}\mathbb{R} \simeq M_n(\mathbb{R})$.*

Then, $\psi(\Gamma^1)$ is a discrete cocompact subgroup of the Lie group G, where

$$G = \begin{cases} \mathrm{SL}_n(\mathbb{C}) & \text{if } K = \mathbb{Q}(\sqrt{-d}) \\ \mathrm{SL}_n(\mathbb{R}) & \text{if } K = \mathbb{Q}, \quad \mathcal{D} \text{ not ramified at } \infty \\ \mathrm{SL}_{n/2}(\mathbb{H}) & \text{if } K = \mathbb{Q}, \quad \mathcal{D} \text{ ramified at } \infty. \end{cases} \tag{3.29}$$

In this case, it follows by a general result by Gorodnik and Nevo [100, Corollary 1.11 and Remark 1.12] that the number of units inside the ball is close to the Haar volume of the ball:

$$\lim_{M\to\infty} \frac{\left|\psi(\Gamma^1)\cap\mathcal{B}_M\right|}{\mathrm{Vol}_G(\mathcal{B}_M)} = C_L$$

where C_L is some nonzero constant independent of M. The previous theorem transforms the point counting

problem into an integration problem. It has been shown by Gorodnik and Weiss that [99, Theorem 7.4] the Haar volume of the ball scales like

$$\mathrm{Vol}_G(\mathcal{B}_M) \doteq M^T,$$

where the constant T can be computed by determining some invariants of the Lie group G. The value of T is well known in the case $G = \mathrm{SL}_n(\mathbb{R})$ [71]. The corresponding results for $\mathrm{SL}_n(\mathbb{H})$ and $\mathrm{SL}_n(\mathbb{C})$, although probably well-known to specialists, were not readily available in the literature, but can be computed by determining some invariants of Lie algebras, related to the Lie groups under consideration (see the Appendix of [J7]):

$$T = \begin{cases} 2n^2 - 2n & \text{for} \quad \mathrm{SL}_n(\mathbb{C}) \\ n^2 - n & \text{for} \quad \mathrm{SL}_n(\mathbb{R}) \\ n^2 - 2n & \text{for} \quad \mathrm{SL}_{n/2}(\mathbb{H}). \end{cases}$$

With this technique we can bound the DMT for multiplexing gain $r \in [0, 1]$ for all division algebra codes with $K = \mathbb{Q}$ or $\mathbb{Q}(\sqrt{-d})$. However, it turns out to be suboptimal for higher multiplexing gains. This is due to the fact that the bound (3.5) is too loose to capture the DMT.

We refine this approach in [C18] by going back to the channel-dependent pairwise error probability (3.3), before averaging:

$$P_e(H) \leq \sum_{X \in \mathcal{C} \setminus \{0\}} e^{-\rho \|HX\|^2} = \sum_{X \in \psi(\Gamma) \cap \mathcal{B}_M \setminus \{0\}} e^{-c\|HX\|^2}.$$

where $c = \rho^{1 - \frac{2rn}{k}}$.

Similarly to the previous approach, by considering a finite set $\{a_1, \ldots, a_j\}$ of coset leaders of Γ^1 in Γ^*, we can obtain a bound of the form

$$P_e(H) \leq \sum_{x \in \mathcal{I}(M)} \sum_{i=1}^{j} \sum_{u \in \Gamma^1} e^{-c\|H\psi(xa_iu)\|^2}$$

in terms of a sum over the unit group Γ^1 and over ideals $\mathcal{I}(M)$. Using a simplified argument inspired by the Strong Wavefront Lemma in [101], we show that the previous sum can be bounded by an integral over the corresponding ball in G:

$$\sum_{\substack{u \in \Gamma^1, \\ \|\psi(u)\|_F \leq M}} e^{-c\|H\psi(au)\|^2} \leq \frac{1}{\mathrm{Vol}_G(\mathcal{F})} \int_{\mathcal{B}_{MR_{\mathcal{F}}}} e^{-\frac{c}{R_{\mathcal{F}}^2}\|H\psi(a)g\|^2} d\mu(g)$$

where $R_{\mathcal{F}}$ is such that the fundamental domain $\mathcal{F}$ of Γ^1 in G is contained in $\mathcal{B}_{R_{\mathcal{F}}}$, and μ is the Haar measure of G. As before, the sum over $\mathcal{I}(M)$ only contributes a term $\sim O(\log M)$ and does not affect the DMT.

After averaging over the channel H and recalling that [211]

$$\int_{M_{m,n}(\mathbb{C})} e^{-c\|HX\|^2} p(H) dH = \frac{1}{(\det(I + cXX^\dagger))^m},$$

the problem is reduced to computing integrals of the form

$$\int_{\mathcal{B}_{MR_{\mathcal{F}}}} \frac{1}{\det\left(I + cgg^*\right)^m} d\mu(g) \tag{3.30}$$

over the Lie group G in (3.29). This can be done using the Cartan decomposition of G (see [C18] for details).

In [C18], we were able to bound this integral explicitly and recover a lower bound for the DMT of general families of lattice codes derived from division algebras with center $\mathbb{Q}$, for a larger range of multiplexing gains (see Table 3.1).

Approach based on equivalent channels. More recently [J12], we were able to confirm that the lower bounds in Table 3.1 indeed give the DMT curves of different types of asymmetric space-time codes for every multiplexing

Lie group	piecewise linear function	condition	
$\mathrm{SL}_n(\mathbb{C})$	$(r, [(n-r)(m-r)]^+),\ r \in \mathbb{Z}$	$m \geq 2\lceil r \rceil - 1$	known
$\mathrm{SL}_n(\mathbb{R})$	$(r, [(n-2r)(m-r)]^+),\ 2r \in \mathbb{Z}$	$m \geq \lceil 2r \rceil - \frac{1}{2}$	new
$\mathrm{SL}_{n/2}(\mathbb{H})$	$(r, (n-2r)(m-r)]^+),\ r \in \mathbb{Z}$	$m \geq 2\lceil r \rceil - 1$	new

Table 3.1: Lower bounds for the DMT of division algebra-based space-time codes using the Strong Wavefront Lemma approach.

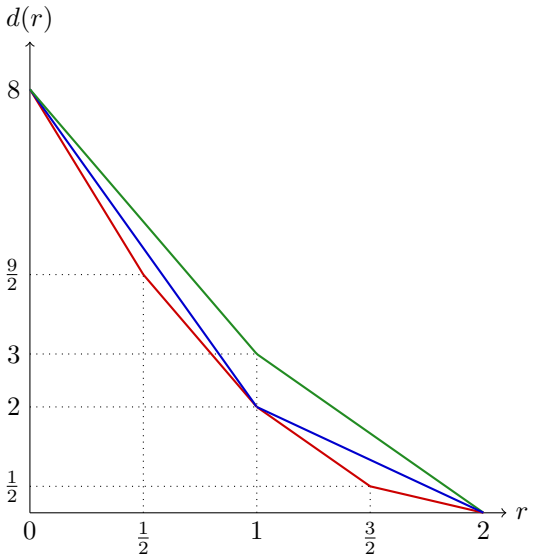


Figure 3.6: DMT upper bounds for $n = 4$, $m = 2$ for codewords restricted to $M_n(\mathbb{C})$ (green), $M_{n/2}(\mathbb{H})$ (blue), and $M_n(\mathbb{R})$ (red).

gain. The approach used in this work is quite different and more general, and applies to a large class of asymmetric codes. In fact, we are not just considering division algebra codes, but all space-time codes whose codewords are restricted to the real and quaternionic matrices $M_n(\mathbb{R})$ or $M_{n/2}(\mathbb{H})$ respectively.

We prove that if the codewords of the space-time scheme belong to these restricted sets of matrices, its DMT is automatically upper bounded by a limit that is tighter than the general DMT bound. The proof is based on re-deriving the bounds by Zheng and Tse [238] for an equivalent real or quaternionic channel model.

Diversity-multiplexing gain trade-off of real and quaternionic lattice codes. Using this approach we establish a general upper bound for the DMT of real and quaternionic codes, as illustrated in Figure 3.6.

Theorem 3.14

1) Suppose that $\forall \rho$, $\mathcal{C}(\rho) \subset M_n(\mathbb{R})$. Then the DMT of the code $\mathcal{C}$ is upper bounded by the piecewise linear function $d_1(r)$ connecting the points $(r, [(m-r)(n-2r)]^+)$ where $2r \in \mathbb{Z}$.

2) Suppose that $\forall \rho$, $\mathcal{C}(\rho) \subset M_{n/2}(\mathbb{H})$. Then the DMT of the code $\mathcal{C}$ is upper bounded by the piecewise linear function $d_2(r)$ connecting the points $(r, [(m-r)(n-2r)]^+)$ for $r \in \mathbb{Z}$.

Then, we show that real and quaternionic spherically shaped lattice codes with the NVD property achieve the DMT upper bounds of Theorem 3.14. This result extends Propositions 4.2 and 4.3 in [C18] (obtained using the Lie group approach presented in the previous subsection) to all multiplexing gains.

Theorem 3.15

1) Let $\mathcal{L}$ be an n^2-dimensional lattice in $M_n(\mathbb{R})$, and consider the code $\mathcal{C}(\rho) = \rho^{-\frac{r}{n}} \mathcal{L}(\rho^{\frac{r}{n}})$. If $\mathcal{L}$ has the NVD property, then the DMT of the code $\mathcal{C}(\rho)$ under ML decoding is the function $d_1(r)$.

2) Let $\mathcal{L}$ be an n^2-dimensional lattice in $M_{n/2}(\mathbb{H})$ with the NVD property. Then the DMT of the code $\mathcal{C}(\rho) = \rho^{-\frac{r}{n}}\mathcal{L}(\rho^{\frac{r}{n}})$ under ML decoding is $d_2(r)$.

These results can be extended to obtain the DMT of multi-block space-time codes; see [J12] for details.

Remark 3.16 We note that the proofs of Theorem 3.14 and Theorem 3.15 also implicitly rely on the theory of Lie groups. In fact, to establish our bounds for equivalent real and quaternionic MIMO channels, we use the joint eigenvalue distributions of real and quaternionic Wishart matrices which were derived in [76] using Lie group methods. More details can be found in [J12].

Theorem 3.15 states that n^2-dimensional NVD lattices in $M_n(\mathbb{R})$ and $M_{n/2}(\mathbb{H})$ do achieve the respective DMT upper bounds of Theorem 3.14. The following Lemma, proven in [218], shows that for every n there exists an n^2-dimensional NVD lattice $\mathcal{L} \subset M_n(\mathbb{R})$, and for every even n there exists an n^2-dimensional NVD lattice $\mathcal{L} \subset M_{n/2}(\mathbb{H})$.

Lemma 3.17 *Let Λ be an* order *in an index n $\mathbb{Q}$-central division algebra $\mathcal{D}$. Then the following statements hold:*

1) If the infinite prime is ramified in the algebra $\mathbb{D}$, then there exists an embedding $\psi_{\text{abs}} : \mathcal{D} \to M_{n/2}(\mathbb{H})$ such that $\psi_{\text{abs}}(\Lambda)$ is an n^2-dimensional NVD lattice.

2) If $\mathbb{D}$ is not ramified at the infinite place, then there exists an embedding $\psi_{\text{abs}} : \mathcal{D} \to M_n(\mathbb{R})$ such that $\psi_{\text{abs}}(\Lambda)$ is an n^2-dimensional NVD lattice.

3) For every n there exists an index n $\mathbb{Q}$-central division algebra that is ramified at the infinite place and one which is not.

As a Corollary, we obtain a complete DMT characterization of $\mathbb{Q}$-central division algebra codes.

Corollary 3.18 *Let Γ be an order in an index n $\mathbb{Q}$-central division algebra $\mathcal{D}$.*

1) If $\mathcal{D}$ is not ramified at the infinite place, then the DMT of the code $\psi_{\text{abs}}(\Lambda) \subset M_n(\mathbb{R})$ achieves the first upper bound of Theorem 3.14.

2) If $\mathcal{D}$ is ramified at the infinite place, then the code $\psi_{\text{abs}}(\Gamma) \subset M_{n/2}(\mathbb{H})$ achieves the second upper bound of Theorem 3.14.

The previous result refers to lattice codes constructed using the abstract embedding of Lemma 3.17. In contrast, explicit codes are typically built using regular representations as in (3.10). Note that if Γ is a $\mathbb{Z}$-order in an index n $\mathbb{Q}$-central division algebra $\mathcal{D}$, then $\psi_{reg}(\Gamma)$ is an n^2-dimensional NVD lattice in $M_n(\mathbb{C})$ [194].

Unfortunately, in general, while these lattices have the correct dimension and the NVD property, there is no guarantee that they are always contained in $M_n(\mathbb{R})$ or in $M_{n/2}(\mathbb{H})$ and we can not directly apply Theorem 3.15. However, the following result from [J7] shows that all the lattices produced by regular representations are conjugated versions of lattices whose DMT we know:

Lemma 3.19 *Let $\mathbb{D}$ be an index n $\mathbb{Q}$-central division algebra and $\Gamma \subset \mathcal{D}$ an order. If the infinite prime is ramified in the algebra $\mathcal{D}$, then there exists an invertible matrix $A \in M_n(\mathbb{C})$ such that*

$$A\psi_{reg}(\Gamma)A^{-1} = \psi_{abs}(\Gamma) \subset M_{n/2}(\mathbb{H}).$$

If $\mathcal{D}$ is not ramified at the infinite place, then there exists an invertible matrix $B \in M_n(\mathbb{C})$ such that

$$B\psi_{reg}(\Gamma)B^{-1} = \psi_{abs}(\Gamma) \subset M_n(\mathbb{R}).$$

The following conjecture then seems to be plausible, but its proof has eluded us.

Conjecture 3.20 *Let $\mathcal{D}$ be an index n $\mathbb{Q}$-central division algebra and $\Gamma \subset \mathcal{D}$ an order. If $\mathcal{D}$ is not ramified at the infinite prime, then the DMT of $\psi_{reg}(\Gamma)$ under ML decoding is equal to the first upper bound of Theorem 3.14. If $\mathcal{D}$ is ramified at the infinite prime, then the DMT of $\psi_{reg}(\Gamma)$ under ML decoding is equal to the second upper*

bound of Theorem 3.14.

Open problems and perspectives Based on the finite-blocklength bounds by Polyanskiy, Poor and Verdù [189], Durisi *et al.* studied finite blocklength, finite SNR rate bounds for multiple-antenna communications [72], which can be seen as generalizing ergodic capacity and the diversity-multiplexing gain trade-off at the same time. Their work reveals a new trade-off between the rate gain achieved by exploiting the available space-time diversity resources and the rate loss caused by the channel estimation overhead, which becomes significant for short packets. In particular, they derive tight upper and lower rate bounds for the Alamouti code. An interesting question is whether this analysis can be extended to more general space-time codes.

3.4 Approaching capacity with multi-block space-time codes

In the previous section, we have studied the performance of algebraic lattice space-time codes over MIMO wireless channels at fixed block length in the asymptotic regime where the SNR tends to infinity. In this section, we consider another asymptotic scenario, where the SNR is fixed, and the blocklength tends to infinity, and study the gap to capacity of multi-block algebraic space-time codes [J9].

This work, in collaboration with R. Vehkalahti, was motivated by the fact that, although the capacity of multi-antenna fading channels was well-known [213], at that time there were no known families of explicit codes achieving this capacity. However, it was known that with simple modulation and strong outer codes such as turbo or LDPC codes, one can operate at rates close to capacity with small error probability [119, 198]. Moreover, in the slow-fading scenario, it was shown that linear precoding and perfect space-time codes achieve a constant gap to capacity [180].

In contrast, for the classical complex Gaussian single antenna channel, it was known that several lattice code constructions achieve $\log \mathrm{SNR} - C$ rates for some constant gap C. These constructions are based on sphere packing arguments showing that the performance of a lattice code in the classical Gaussian channel can be roughly estimated by the size of a geometrical invariant of the lattice, the *Hermite invariant*. This connection has led to the outstanding work of Conway and Sloane connecting algebra, geometry and information theory [49].

In the case of fading channels, it was well-known that the minimum determinant criterion allows to improve the worst-case pairwise error probability in the high-SNR regime, when coding over a single fading block (see Section 3.1). However, no design criterion had been suggested to approach the MIMO capacity with explicit lattice codes. In [J9] we addressed this problem and showed that when we are allowed to encode and decode over a growing number of fading blocks, the *normalized minimum determinant* plays a similar role to the Hermite constant in Gaussian channels. In particular it can be used to measure how close to capacity a given family of lattice codes can get.

Multi-block space-time codes. We consider a MIMO system with n transmit and m receive antennas, where transmission takes place over k quasi-static fading blocks of delay $T = n$. A multi-block codeword $X \in M_{n\times nk}(\mathbb{C})$ has the form $[X_1, X_2, \ldots, X_k]$, where the submatrix $X_i \in M_n(\mathbb{C})$ is sent during the i-th block. The received signals are given by

$$Y_i = H_i X_i + W_i, \qquad i \in \{1, \ldots, k\} \tag{3.31}$$

where $H_i \in M_{m\times n}(\mathbb{C})$ and $W_i \in M_{m\times n}(\mathbb{C})$ are the channel and noise matrices. The coefficients of W_i are modeled as circular symmetric complex Gaussian with zero mean and unit variance per complex dimension. Perfect channel state information is available at the receiver but not at the transmitter, and decoding is performed after all k blocks have been received. We will call such a channel an (n, m, k)-*multi-block channel*.
For the sake of simplicity, we will suppose that $m \geq n$ unless explicitly stated. We also assume that for all $i \geq 1$, H_i is full-rank with probability 1, and that the random variable $\sum_{i=1}^{k} \frac{1}{k} \log\det(H_i^\dagger H_i)$ converges in probability to some constant when the number of blocks k tends to infinity. This model covers several standard MIMO channels such as the Rayleigh block fading channel and the Gaussian MIMO channel.

A *multi-block code* $\mathcal{C}$ in a (n, m, k)-channel is a set of matrices in $M_{n\times nk}(\mathbb{C})$. In particular we will concentrate on finite codes that are drawn from lattices. Let R denote the code rate in bits per complex channel use;

equivalently, $|\mathcal{C}| = 2^{Rkn}$. We assume that every matrix X in a finite code $\mathcal{C} \subset M_{n\times nk}(\mathbb{C})$ satisfies the average power constraint

$$\frac{1}{nk}\mathbb{E}[\|X\|^2] \leq P. \tag{3.32}$$

We denote by $B(r)$ the set of matrices in $M_{n\times nk}(\mathbb{C})$ with Frobenius norm smaller or equal to r. Given a family of lattices $L_{n,k} \subseteq M_{n\times nk}(\mathbb{C})$ and $R > 0$, we want to design spherically shaped multi-block codes $\mathcal{C}$ of the form

$$\mathcal{C}_L = B(\sqrt{Pkn}) \cap (X_R + \alpha L_{n,k}), \tag{3.33}$$

having rate greater or equal to R, and satisfying the average power constraint (3.32). In the above formula, α is a suitable energy normalization constant, and X_R is a suitable shift. One can show [103] that we can choose $X_R \in M_{n\times nk}(\mathbb{C})$ such that

$$2^{Rnk} = |\mathcal{C}_L| \geq \frac{\mathrm{Vol}(B(\sqrt{Pkn}))}{\mathrm{Vol}(\alpha L_{n,k})} = \frac{C_{n,k}P^{n^2k}}{\alpha^{2n^2k}\,\mathrm{Vol}(L_{n,k})},$$

where $C_{n,k} = \frac{(\pi nk)^{n^2k}}{(n^2k)!}$. We then find the following condition for the scaling constant:

$$\alpha^2 = \frac{C_{n,k}^{\frac{1}{n^2k}}P}{2^{\frac{R}{n}}\,\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}}} \tag{3.34}$$

Given a matrix $X = [X_1, \ldots, X_k] \in M_{n\times nk}(\mathbb{C})$, we define its *product determinant*

$$\mathrm{pdet}(X) = \prod_{i=1}^{k} \det(X_i). \tag{3.35}$$

The *minimum determinant* of the lattice $L \subseteq M_{n\times nk}(\mathbb{C})$ is defined as

$$\det_{\min}(L) := \inf_{X\in L\setminus\{\mathbf{0}\}} |\mathrm{pdet}(X)|.$$

We can now define the *normalized minimum determinant* $\delta(L)$, which is obtained by first scaling the lattice L to have a unit size fundamental parallelotope and then taking the minimum determinant of the resulting scaled lattice. A simple computation shows that

$$\delta(L) = \frac{\det_{\min}(L)}{(\mathrm{Vol}(L))^{1/2n}}. \tag{3.36}$$

Reduced Hermite invariant and "incompressible" lattices. We define the following notation for component-wise multiplication of multi-block matrices: given $X = [X_1, \ldots, X_k]$ and $H = [H_1, \ldots, H_k] \in M_{n\times nk}(\mathbb{C})$,

$$H * X \doteq [H_1X_1, \ldots, H_kX_k]. \tag{3.37}$$

With this notation, the channel output $Y = [Y_1, \ldots, Y_k]$ can be written as

$$Y = H * X + W, \tag{3.38}$$

where $W = [W_1, \ldots, W_k]$ is the multi-block noise. From the receiver's point of view, this is equivalent to an additive white Gaussian noise channel where the lattice code is

$$H * \mathcal{C}_L = \{H * X \mid X \in \mathcal{C}_L\}.$$

Even if the lattice $L_{n,k}$ (and therefore the code $\mathcal{C}_L$) has good minimum distance, there is no guarantee that the same can be said about the lattice $H * \mathcal{C}_L$ after transmission over the channel. This leads us to consider the following

design criterion: the lattice should be "incompressible", in the sense that it should still have good minimum distance after fading.

First, we recall the classical definition of the Hermite constant, which characterizes the density of a lattice packing:

Definition 3.21 *The Hermite constant of a d-dimensional lattice $L \subset M_{n\times nk}(\mathbb{C})$ can be defined as*

$$\mathrm{h}(L) = \frac{\inf\{\, ||X||^2 \mid X \in L, X \neq 0\}}{\mathrm{Vol}(L)^{2/d}}.$$

For convenience we introduce the group of matrices

$$\mathcal{H} = \left\{ H \in M_{n\times nk}(\mathbb{C}) \;\middle|\; \prod_{i=1}^{k} \det(H_i) = 1 \right\}. \tag{3.39}$$

Definition 3.22 *The* reduced Hermite constant *of an m-dimensional lattice $L \subset M_{n\times nk}(\mathbb{C})$ with respect to the group $\mathcal{H}$ is defined as*

$$\mathrm{rh}_{\mathcal{H}}(L) = \inf_{H\in\mathcal{H}} \{\mathrm{h}(H * L)\}.$$

For any lattice L, the Hermite constant $\mathrm{h}(L) > 0$. The same is not true for the reduced Hermite invariant. Let us now describe the set of lattices L for which $\mathrm{rh}_{\mathcal{H}}(L) > 0$.

The normalized minimum determinant (3.36) provides an alternative characterization of the reduced Hermite invariant:

Proposition 3.23 *If $L \subset M_{n\times nk}(\mathbb{C})$ is a $2n^2k$-dimensional lattice, then*

$$\mathrm{rh}_{\mathcal{H}}(L) = nk\,(\delta(L))^{2/nk}.$$

Consequently, in order to maximize the reduced Hermite constant, we should maximize the normalized minimum determinant.

Algebraic constructions of multi-block space-time codes. We have seen in Section 3.1 how cyclic division algebras can be used to design single block space-time codes through the left regular representation ψ in equation (3.10). A generalization of the embedding ψ to the multi-block case was proposed in [229, 158] for division algebras whose center K contains an imaginary quadratic field. In this work we consider a more general multi-block construction developed in [153], which applies to any totally complex center K.

Assume that E/K is a cyclic Galois extension of degree n with Galois group $\mathrm{Gal}(E/K) = \langle\sigma\rangle$, and that $\mathcal{D} = (E/K, \sigma, \gamma)$ is a cyclic division algebra.

Recall that a totally complex field K has $2k$ distinct $\mathbb{Q}$-embeddings $\beta_i : K \hookrightarrow \mathbb{C}$ in complex conjugate pairs. We will denote by $\overline{\beta_i}$ the embedding given by $x \mapsto \overline{\beta_i(x)}$. We can extend each β_i to an embedding $\alpha_i : E \hookrightarrow \mathbb{C}$ such that $\alpha_i|_K = \beta_i$ and $\overline{\alpha_i}|_K = \overline{\beta_i}$. We order the embeddings $\{\alpha_1, \ldots, \alpha_{2k}\}$ so that $\alpha_i = \overline{\alpha_{i+k}}$, for $1 \leq i \leq k$. Let a be an element of $\mathcal{D}$ and $A = \phi(a)$. Consider the mapping $\varphi : \mathcal{D} \mapsto M_{n\times nk}(\mathbb{C})$ given by

$$a \mapsto (\alpha_1(A), \ldots, \alpha_k(A)), \tag{3.40}$$

where each α_i is extended to an embedding $\alpha_i : M_n(E) \hookrightarrow M_n(\mathbb{C})$.

If Γ is an $\mathcal{O}_K$-order in $\mathcal{D}$, then $L_{n,k} = \varphi(\Gamma)$ is a $2kn^2$-dimensional multi-block lattice in $M_{n\times nk}(\mathbb{C})$ with the non-vanishing (multiblock) determinant property: for $X \in \varphi(\Gamma)$, $X \neq 0$,

$$\prod_{i=1}^{k} |\det(X_i)|^2 = \prod_{i=1}^{2k} \det(\alpha_i(A)) = \prod_{i=1}^{2k} \alpha_i(\det(A)) = N_{K/\mathbb{Q}}(N_{\mathcal{D}/K}(a)) = N_{\mathcal{D}/\mathbb{Q}}(a) \geq 1, \tag{3.41}$$

and therefore $\det_{\min}(L_{n,k}) = \inf_{X\neq 0}\prod_{i=1}^{k}|\det(X_i)| = 1$. Moreover, we have the volume formula

$$\mathrm{Vol}(\varphi(\Gamma)) = \frac{1}{2^{kn^2}}\sqrt{N_{K/\mathbb{Q}}(d(\Gamma/\mathcal{O}_K))}\sqrt{|d_K|^{n^2}}, \tag{3.42}$$

where $d(\Gamma/\mathcal{O}_K)$ is the $\mathcal{O}_K$-discriminant of the order Γ, and d_K is the discriminant of the field K. We refer the reader to [194] for the relevant definitions. In order to maximize the normalized minimum determinant, this volume should be minimized.

In the case of fixed center K, [216] addressed the problem of finding the division algebras with the smallest $\mathcal{O}_K$-discriminant, yielding the densest MIMO lattices. The main construction is based on the following result (Theorem 6.14 in [216]):

Theorem 3.24 *Let K be a number field of degree $2k$ and P_1 and P_2 be two prime ideals of K. Then there exists a degree n division algebra $\mathcal{D}$ having an $\mathcal{O}_K$-order Γ with discriminant*

$$d(\Gamma/\mathcal{O}_K) = (P_1P_2)^{n(n-1)} \tag{3.43}$$

The volume formula (3.42) and Theorem 3.24 suggest that in order to build families of (n,n,k) multi-block codes with the largest normalized minimum determinant, we should proceed in two steps:

a) choose a sequence of center fields K of degree $2k$ such that their discriminants d_K grow as slowly as possible;
b) given the center K, choose an algebra $\mathcal{D}$ and an order Γ satisfying (3.43), where P_1 and P_2 are the prime ideals in K with the smallest norms[7].

We now discuss the choice of a suitable sequence of center fields. The following theorem by Martinet [163] proves the existence of infinite sequences of totally complex number fields K with small discriminants d_K.

Theorem 3.25 (Martinet) *There exists an infinite tower of totally complex number fields $\{K_k\}$ of degree $2k$, where $2k = 5\cdot 2^{2+t}$, such that*

$$|d_{K_k}|^{\frac{1}{2k}} = G, \tag{3.44}$$

for $G \approx 92.368$.

In particular for the fields K_k we can choose primes P_1 and P_2 such that

$$N_{K/\mathbb{Q}}(P_1) \leq 23^{k/10} \text{ and } N_{K/\mathbb{Q}}(P_2) \leq 23^{k/10}. \tag{3.45}$$

Taking into account the volume formula (3.42), Theorem 3.25 and the bound (3.45), we have shown the existence of a family of multiblock lattices $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ such that

$$\mathrm{Vol}(L_{n,k}) \leq 23^{\frac{kn(n-1)}{10}}\left(\frac{G}{2}\right)^{n^2k}. \tag{3.46}$$

Unfortunately, Martinet's theorem is not constructive, being based on Golod and Shafarevich's result [98] on the existence of infinite Hilbert class field towers. For fixed degree, the required number fields can be found using computational algebra software, but this process is computationally taxing [78].

Remark 3.26 The number field towers in Theorem 3.25 are not the best known possible. It was shown in [107] that one can construct a family of totally complex fields such that $G < 82.2$, but this choice would add some notational complications. The optimal value of G is still not known.

Algebraic codes achieving constant gap to capacity Suppose that an infinite family of lattices $L_k \in \mathbb{C}^k$ has Hermite invariants satisfying $\frac{\mathrm{h}(L_k)}{k} \geq c$, for some positive constant c. Then a classical result in information theory

[7] However, we note [153] that *a priori* there may be a trade-off between these two choices, so that minimizing the two terms in (3.42) separately may be suboptimal.

states that with this family of lattices, all rates satisfying

$$R < \log P - \log\left(\frac{4}{\pi e}\right) + \log c,$$

are achievable in the additive complex Gaussian channel [49, Chapter 3].

We want to establish an analogue of this result for fading channels. To this end, we will study the performance of the spherical codes $\mathcal{C}_L$ of the form (3.33) built using the algebraic multiblock construction in the previous section over a general class of fading channels.
We consider both ML decoding and "naive lattice decoding" [208], namely the closest point search in the infinite lattice $L_{n,k}$.

Theorem 3.27 *Suppose that $m \geq n$, and let $\{H_i\}_{i\in\mathbb{Z}}$ be a fading process such that $H_i \in M_{m\times n}$ is full-rank with probability 1, and such that the weak law of large numbers holds for the random variables $\{\log\det(H_i^\dagger H_i)\}$, i.e. $\exists \mu > 0$ such that $\forall \epsilon > 0$,*

$$\lim_{k\to\infty} \mathbb{P}\left\{\left|\frac{1}{k}\sum_{i=1}^{k}\log\det(H_i^\dagger H_i) - \mu\right| > \epsilon\right\} = 0. \tag{3.47}$$

Let $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that

$$\det_{\min}(L_{n,k}) = 1, \quad \textit{and} \quad \mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L \tag{3.48}$$

for some constant $C_L > 0$. Then, any rate

$$R < \mu + n\left(\log P - \log\frac{4n^2}{\pi e} - \log C_L\right) \tag{3.49}$$

is achievable using the codes $L_{n,k}$ both with ML decoding and naive lattice decoding.

The proof of Theorem 3.27 is based on the sphere bound for the error probability, which holds for both ML decoding and naive lattice decoding:

$$P_e \leq \mathbb{P}\left\{\|W\|^2 \geq \left(\frac{d_H}{2}\right)^2\right\},$$

where d_H is the minimum distance in the received constellation (see Figure 3.7). A lower bound for d_H is given by

$$d_H^2 = \min_{X\neq\overline{X}}\sum_{i=1}^{k}\left\|H_i(X_i - \overline{X}_i)\right\|^2 \overset{(a)}{\geq} \alpha^2 nk \min_{X\neq 0}\prod_{i=1}^{k}|\det(H_iX_i)|^{\frac{2}{nk}} \overset{(b)}{\geq} \alpha^2 nk\prod_{i=1}^{k}|\det(H_i)|^{\frac{2}{nk}}.$$

where (a) follows from the arithmetic - geometric mean inequality and (b) follows from the NVD property (3.41). We then find sufficient conditions in order to have $P_e \to 0$ by using the weak law of large numbers (3.47). The rate conditions follow from the volume bound (3.46) and the choice of a suitable normalization as in (3.34). See [J9] for details.

Remark 3.28 For the ML decoder we can prove an analogue of Theorem 3.27 also in the case $m < n$, although the bound on achievable rates is more involved [J9].

Remark 3.29 We note that existence of a family of lattices with $C_L \leq 23^{\frac{(n-1)}{10n}}G/2$, was shown in the previous section.

We now consider a deterministic model, where $H_i = H$ is constant. If the channel is known at the receiver but not at the transmitter, the transmitter cannot use optimal power allocation and waterfilling, and can only achieve

the *white-input capacity* corresponding to uniform power allocation $C_{\text{WI}} = \log\det\left(I_{n_r} + P/nHH^\dagger\right)$. This is for example the case for an open-loop broadcast channel where the transmitter cannot perform rate adaptation for all the users. The following corollary then shows that a constant gap to white-input capacity is achievable:

Corollary 3.30 (Deterministic channel) *Consider a deterministic channel with $m \geq n$ such that $H_i = H$ for all $i \geq 1$, and let $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\text{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, this coding scheme can achieve any rate*

$$R < \log\det\frac{P}{n}H^\dagger H - n\log C_L - n\log\frac{4n}{\pi e}.$$

Another class of channels satisfying the hypotheses of Theorem 3.27 is the set of ergodic and stationary fading processes $\{H_i\}$. If we suppose that the channel is *isotropically invariant*, i.e. the distribution of H is invariant under right multiplication by unitary matrices, then under the assumption of no CSI at the transmitter, the optimal input allocation is uniform [213] and the capacity is $C = \mathbb{E}_H\left[\log\det\left(I_{n_r} + P/nHH^\dagger\right)\right]$.

Corollary 3.31 (Stationary ergodic channels) *Suppose that $n_r \geq n$ and that the fading process $\{H_i\}$ is ergodic, stationary and isotropically invariant. Moreover, suppose that $\mathbb{E}\left[\left|\log\det H^\dagger H\right|\right] < \infty$. Let $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\text{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, any rate*

$$R < \mathbb{E}_H\left[\log\det\frac{P}{n}H^\dagger H\right] - n\log C_L - n\log\frac{4n}{\pi e}$$

is achievable using the codes $L_{n,k}$ both with ML decoding and naive lattice decoding.

Remark 3.32 We can actually show that the achievable rate is within a constant gap from capacity, although this constant will depend on the channel statistics.

In the case of an i.i.d Rayleigh fading MIMO channel, the achievable rate can be computed explicitly (see Figures 3.9 and 3.10), and the error probability vanishes exponentially fast. See [J9] for details.

Open questions and follow-up work In this work we proved the existence of lattice codes achieving constant gap to capacity in ergodic fading channels. Unlike existence results based on random coding, our codes are always built from the same family of lattices, irrespective of the SNR and even of the fading statistics. Hence, using the minimum determinant as a design principle leads to *approximately universal codes* which guarantee robustness with respect to imperfect channel estimation. This property can be useful in open loop mode for high-mobility users, or for broadcast channels when the transmitter cannot adapt their rate to all the receivers.

However, our codes still have a considerable gap to capacity and further research is needed. We note that this gap depends on several factors:

- First of all, the normalized minimum determinant affects the value of the gap. One could try to find families of lattices with larger normalized minimum determinant, for instance by replacing the centers in our constructions with families of number fields having smaller discriminants. This is a hard problem in number theory, since the optimal value of the constant G for number fields is not known. For finite degrees, Odlyzko's bounds [176] may give much better discriminants. One can also consider more general examples of lattices than those arising from orders in division algebras, such as ideals of orders.

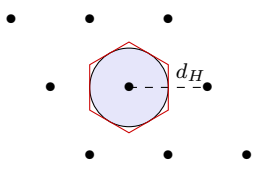


Figure 3.7: An illustration of the sphere upper bound.

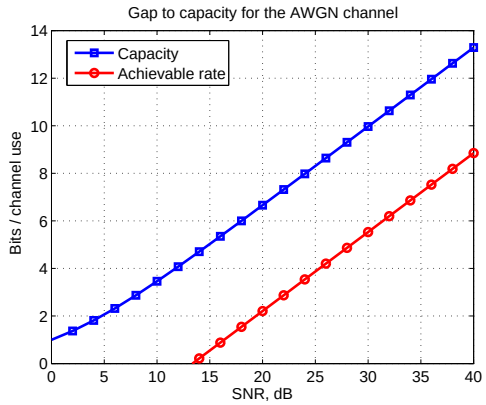


Figure 3.8: Achievable rate on a single-antenna AWGN channel.

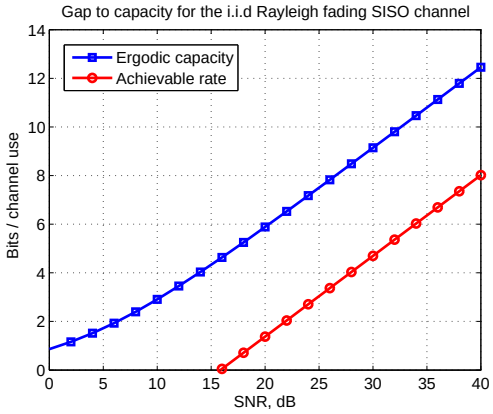


Figure 3.9: Achievable rate and channel capacity for the single antenna i.i.d. Rayleigh fading channel.

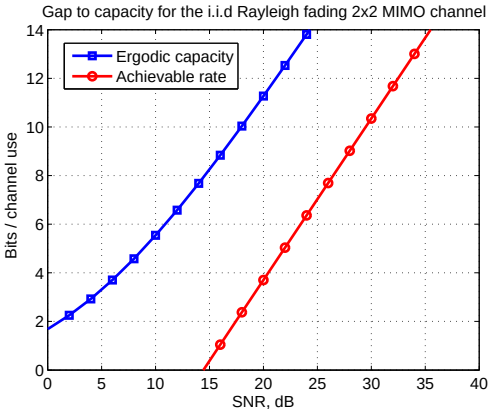


Figure 3.10: Achievable rate and channel capacity for the 2×2 MIMO i.i.d. Rayleigh fading channel.

- Second, our bound for the error probability is based on sphere packing and might be suboptimal, since it bounds the performance of ML decoding with bounded distance decoding.
- Finally, in this work we have not considered the issue of shaping. On one hand, improving the shaping properties of our lattices might lead to a better error probability bound; one interesting question is under which hypotheses the canonical embedding of a number field yields a lattice whose fundamental region is close to a sphere. Moreover, we have assumed that the codewords are chosen uniformly inside a spherical lattice code; replacing the uniform distribution with a discrete Gaussian distribution would compensate the loss of "+1" in the capacity expression.

Finally, we note that while the family of codes in question is well-defined and deterministic, the best known algorithms to compute Hilbert class fields of arbitrary number fields have high computational complexity [78], and thus our construction cannot be made explicit at present.

Follow-up work Since the submission of this work, there have been several advances on the topic. In [221, Section 4.5], S. Vituri gave a proof of existence of lattice codes achieving a constant gap to capacity for ergodic SISO channels. It appears that with minor modifications this proof implies the existence of capacity-achieving lattices. In [156] the authors prove that polar lattices achieve capacity in i.i.d fading channels. This is not only an existence result, but provides an explicit low-complexity code construction as well. In [36] the authors prove the existence of lattice codes achieving capacity in the compound SISO channel, where the fading is random during the first s time units, but then gets repeated in blocks of length s.

4 | LATTICE CODES FOR PHYSICAL LAYER SECURITY AND CRYPTOGRAPHY

Physical layer security aims to exploit the random nature of noisy channels in order to offer an additional level of protection. The notion of cryptographic security, which is based on the computational complexity of mathematical primitives which are hard to invert, is replaced by information-theoretic security, which is measured in terms of statistical independence between the confidential message M and the channel output Z^n, and implies that even a computationally unlimited adversary cannot extract any useful information from the signal. This notion was first introduced by Shannon [205], who required *perfect secrecy*, i.e. $\mathbb{I}(\mathsf{M};\mathsf{Z}^n) = 0$. However, perfect secrecy is impractical, since in Shannon's noise-free setting, it entails the use of a one-time-pad.

In the context of noisy wiretap channels, Wyner [228] proved that both error correction and data confidentiality can be achieved by channel coding without any secret key. He replaced Shannon's perfect secrecy with the asymptotic weak secrecy condition $\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Z}^n) = 0$ as the number of channel uses tends to infinity. However, this notion is too weak for many practical applications, since the total amount of leaked data can still tend to infinity, and now it is widely accepted that a physical-layer security scheme should be secure in the sense of Csiszár's *strong secrecy* $\lim_{n\to\infty} \mathbb{I}(\mathsf{M};\mathsf{Z}^n) = 0$ [53]. In information theory, plaintexts are often assumed to be uniformly distributed. This assumption is deemed problematic from a cryptographic perspective, since real-life messages are often not uniformly random, and might have low entropy. This issue can be resolved by using the standard notion of *semantic security* [97] which requires that the probability that the eavesdropper can guess any function of the message given the ciphertext should not be significantly higher than the probability of guessing it using a simulator that does not have access to the ciphertext. The relation between strong secrecy and semantic security was revealed in [22] for discrete wiretap channels, namely, achieving strong secrecy for all distributions of the plaintext messages is equivalent to achieving semantic security.

It can be shown that stochastic encoding is necessary to ensure information-theoretic security over wiretap channels [30, Section 3.4.1]. Each confidential message m is then associated to a subcodebook or "bin" rather than to a single codeword.

Fundamental mechanisms for secrecy and relation to channel resolvability Our approach to achieve strong secrecy is based on the notion of *channel resolvability*, which was introduced by Han and Verdù [108]. We review here its simplest formulation in the case of discrete memoryless channels.

Definition 4.1 (Resolvability code and resolution rate) *Consider a discrete memoryless channel $W : \mathcal{X} \to \mathcal{Z}$ with transition probability $p_{\mathsf{Z}|\mathsf{X}}$, whose input is an i.i.d. source distributed according to q_{X}; the output of the channel is then an i.i.d. process distributed according to q_{Z}. The aim is to construct a sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$ of rate R and increasing block length n, such that the output distribution p_{Z^n} induced by a uniform choice of the codewords in $\{\mathcal{C}_n\}$ approaches the distribution $q_{\mathsf{Z}^n} \sim \prod_{i=1}^n q_{\mathsf{Z}}$ in variational distance, i.e.*

$$\lim_{n\to\infty} \mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n}) = 0. \tag{4.1}$$

In this case, the sequence $\{\mathcal{C}_n\}_{n\geq 1}$ is called a sequence of resolvability codes *achieving* resolution rate R *for the channel W. The* channel resolvability *of W is then defined as the minimum resolution rate such that resolvability codes exist for any input source.*

This definition can be extended to a very general class of channels using information spectrum methods; it can be shown that the resolvability is equal to the channel capacity as long as the channel satisfies the strong converse [108].

The relation between resolvability and secrecy was investigated in [53, 114, 29], and in my own work in collaboration with M. Bloch [C5]. The key idea is that each bin corresponding to a confidential message should be a resolvability code for the eavesdropper's channel, i.e. the output distributions $p_{\mathsf{Z}^n|\mathsf{M}=m}$ corresponding to different bins should be indistinguishable in variational distance. In the case of random wiretap codes for discrete memoryless channels, the following Lemma holds [57]:

Lemma 4.2 (Cloud mixing) *For a memoryless channel $p_{\mathsf{Z}|\mathsf{X}}$, an input distribution q_{X} and the corresponding output q_{Z}, if the random codebook sequence $\{\mathcal{C}_n\}$ has rate $R' > \mathbb{I}(q_{\mathsf{X}}; q_{\mathsf{Z}})$, then $\exists \beta > 0$ such that*

$$\forall n, \quad \mathbb{E}_{\mathcal{C}_n}\left[\mathbb{V}(p_{\mathsf{Z}^n|\mathcal{C}_n}, q_{\mathsf{Z}^n})\right] \leq e^{-\beta n},$$

where the expectation is computed over the random code ensemble.

Applying this result to the wiretap channel, we find that in order to have strong secrecy, the bin rate R' should be greater than the capacity C_e of the eavesdropper's channel.

4.1 Semantically secure lattice codes for Gaussian wiretap channels

Wireless systems are particularly vulnerable to eavesdropping since every transmission can potentially be overheard by all the neighboring nodes in the network. Moreover, the wireless medium is inherently a source of randomness, which can be harnessed to provide security. Therefore, the design of wiretap codes over continuous channels such as Gaussian and fading channels is of particular interest for the practical integration of physical layer security into future communication systems.

Previous works For continuous wiretap channels such as the Gaussian channel, the earliest code designs focused on the maximization of the eavesdropper's error probability [131]. In particular, [21, 177] considered nested lattice codes for the Gaussian wiretap channel, introduced the notion of *secrecy gain* and showed the existence of families of even unimodular lattices such that the eavesdropper's error probability tends to 1 when the lattice dimension tends to ∞. These lattices were also investigated in [85]. In [43] it was shown that there exist families of lattice codes achieving the weak secret key capacity. Lattice codes were also used to provide weak / strong secrecy in the settings of cooperative jamming and interference channels [116–118].

In this section, I review our main contributions on the design of strongly and semantically secure lattice codes for the Gaussian wiretap channel [J8].

Strong secrecy and semantic security in continuous channels Extending the results of [22, Section 3], we showed that semantic security and strong secrecy for all message distributions are equivalent for continuous channels. Details can be found in Section II.B of [J8].

The following continuous channel adaptation of Csiszár's Lemma [53, Lemma 1] shows how to bound the leakage for arbitrary message distributions p_{M} using the variational distance of output distributions. The lower bound is a consequence of Pinsker's inequality [54]. The proof of the upper bound is similar to the discrete case.

Lemma 4.3 *Let Z^n be a random variable defined on $\mathbb{R}^n$ and M be a random variable taking values in the finite set $\mathcal{M}_n$ such that $|\mathcal{M}_n| \geq 4$. Then*

$$\frac{1}{2} d_{\mathrm{av}}^2 \leq \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \leq d_{\mathrm{av}} \log \frac{|\mathcal{M}_n|}{d_{\mathrm{av}}},$$

where

$$d_{\mathrm{av}} = \sum_{m \in \mathcal{M}_n} p_{\mathsf{M}}(m)\, \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n})$$

is the average variational distance of the conditional output distributions from the global output distribution.

Note that for any distribution q_{Z^n} on $\mathbb{R}^n$, we have [53]

$$d_{\mathrm{av}} \le 2 \sum_{m \in \mathcal{M}_n} p_{\mathsf{M}}(m)\, \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, q_{\mathsf{Z}^n}). \tag{4.2}$$

Together with Lemma 4.3, this leads to an upper bound on the mutual information, in case we can approximate $p_{\mathsf{Z}^n|\mathsf{M}=m}$ by a density that is independent of m.

Lemma 4.4 *Suppose that $\forall n$ there exists some density q_{Z^n} in $\mathbb{R}^n$ such that $\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, q_{\mathsf{Z}^n}) \le \varepsilon_n$, for all $m \in \mathcal{M}_n$. Then we have $d_{\mathrm{av}} \le 2\varepsilon_n$ and so*

$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \le 2\varepsilon_n nR - 2\varepsilon_n \log(2\varepsilon_n). \tag{4.3}$$

Discrete Gaussian distributions and the flatness factor. For $\sigma > 0$ and $\mathbf{c} \in \mathbb{R}^n$, we define the Gaussian distribution of variance σ^2 centered at $\mathbf{c} \in \mathbb{R}^n$ as

$$f_{\sigma,\mathbf{c}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|\mathbf{x}-\mathbf{c}\|^2}{2\sigma^2}},$$

for all $\mathbf{x} \in \mathbb{R}^n$. For convenience, we write $f_\sigma(\mathbf{x}) = f_{\sigma,\mathbf{0}}(\mathbf{x})$.

Let Λ be an n-dimensional lattice in $\mathbb{R}^n$. To study the effect of Gaussian noise on lattice signalling, we consider the Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \sum_{\lambda \in \Lambda} f_{\sigma,\lambda}(\mathbf{x}) \tag{4.4}$$

for all $\mathbf{x} \in \mathbb{R}^n$. We denote by $f_{\sigma,\mathcal{R}(\Lambda)} = f_{\sigma,\Lambda|\mathcal{R}(\Lambda)}$ its restriction to the fundamental region $\mathcal{R}(\Lambda)$. Note that $f_{\sigma,\mathcal{R}(\Lambda)}$ is the probability density of $\bar{\mathsf{X}}^n = [\mathsf{X}^n] \bmod \mathcal{R}(\Lambda)$, where $\mathsf{X}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.

We define the *discrete Gaussian distribution* over Λ centered at $\mathbf{c} \in \mathbb{R}^n$ as the following discrete distribution taking values in $\lambda \in \Lambda$ [15, 166]:

$$D_{\Lambda,\sigma,\mathbf{c}}(\lambda) = \frac{f_{\sigma,\mathbf{c}}(\lambda)}{f_{\sigma,\mathbf{c}}(\Lambda)} \quad \forall \lambda \in \Lambda,$$

where $f_{\sigma,\mathbf{c}}(\Lambda) \coloneqq \sum_{\lambda \in \Lambda} f_{\sigma,\mathbf{c}}(\lambda)$. Again for convenience, we write $D_{\Lambda,\sigma} = D_{\Lambda,\sigma,\mathbf{0}}$.

It will be useful to define the discrete Gaussian distribution over a coset of Λ, i.e., the shifted lattice $\Lambda - \mathbf{c}$:

$$D_{\Lambda-\mathbf{c},\sigma}(\lambda - \mathbf{c}) = \frac{f_\sigma(\lambda - \mathbf{c})}{f_{\sigma,\mathbf{c}}(\Lambda)} \quad \forall \lambda \in \Lambda.$$

Note the relation $D_{\Lambda-\mathbf{c},\sigma}(\lambda - \mathbf{c}) = D_{\Lambda,\sigma,\mathbf{c}}(\lambda)$, namely, they are a shifted version of each other.

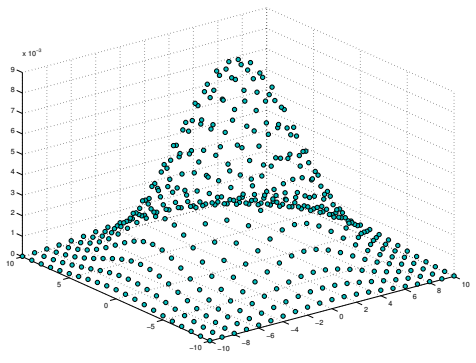


Figure 4.1: The discrete Gaussian distribution over the $\mathbb{Z}^2$ lattice.

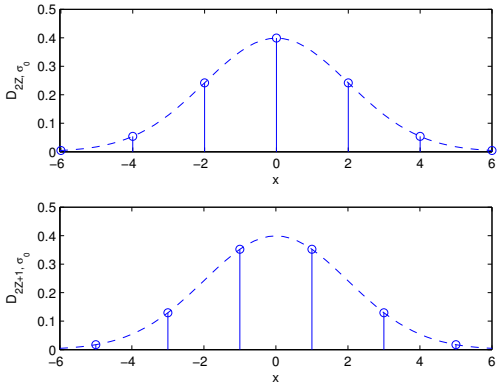


Figure 4.2: Discrete Gaussians over $2\mathbb{Z}$ and its coset $2\mathbb{Z}+1$ for $\sigma_0 = 2$. Note that both distributions are centered at 0.

Remark 4.5 Finding efficient algorithms to sample lattice Gaussians is an important problem in lattice-based cryptography. In particular, it was proven in [96] that Klein's algorithm [128] samples from a distribution which is close to a discrete Gaussian when σ is large enough.

The flatness factor [20] of a lattice Λ measures the L^∞ distance of $f_{\sigma,\Lambda}$ from the uniform distribution on the fundamental region $\mathcal{R}(\Lambda)$:

Definition 4.6 (Flatness factor) *For a lattice Λ with fundamental region $\mathcal{R}(\Lambda)$ and for standard deviation σ, the flatness factor is defined by:*

$$\epsilon_\Lambda(\sigma) \coloneqq \max_{\mathbf{x}\in\mathcal{R}(\Lambda)} \frac{|f_{\sigma,\Lambda}(\mathbf{x}) - 1/V(\Lambda)|}{1/V(\Lambda)}.$$

The flatness factor can be computed from the *theta series* $\Theta_\Lambda(\tau)$ of the lattice Λ [49], which is defined as

$$\Theta_\Lambda(\tau) = \sum_{\lambda\in\Lambda} e^{-\pi\tau\|\lambda\|^2}, \quad \tau > 0. \tag{4.5}$$

Proposition 4.7 (Expression of $\epsilon_\Lambda(\sigma)$) *We have:*

$$\epsilon_\Lambda(\sigma) = \left(\frac{\gamma_\Lambda(\sigma)}{2\pi}\right)^{\frac{n}{2}} \Theta_\Lambda\left(\frac{1}{2\pi\sigma^2}\right) - 1$$

where $\gamma_\Lambda(\sigma) = \frac{V(\Lambda)^{\frac{2}{n}}}{\sigma^2}$ is the volume-to-noise ratio *(VNR).*

Alternatively, the flatness factor can be expressed in terms of the theta series of the dual lattice Λ^* as follows:

$$\epsilon_\Lambda(\sigma) = \Theta_{\Lambda^*}(2\pi\sigma^2) - 1 \tag{4.6}$$

Remark 4.8 One can show that the maxima of both $f_{\sigma,\Lambda}(\mathbf{x})$ and $|f_{\sigma,\Lambda}(\mathbf{x}) - 1/V(\Lambda)|$ are reached when $\mathbf{x} \in \Lambda$.

We note that the flatness factor is equivalent to the notion of smoothing parameter[1] that is commonly used in lattice-based cryptography [166].

Definition 4.9 (Smoothing parameter) *For a lattice Λ and for $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest $\sigma > 0$ such that $\sum_{\boldsymbol{\lambda}^*\in\Lambda^*\setminus\{\mathbf{0}\}} e^{-2\pi^2\sigma^2\|\boldsymbol{\lambda}^*\|^2} \leq \varepsilon$.*

Lemma 4.10 *If $\sigma = \eta_\varepsilon(\Lambda)$, then $\epsilon_\Lambda(\sigma) = \varepsilon$.*

Remark 4.11 The flatness factor is a monotonically decreasing function of σ, i.e., if $\sigma < \sigma'$, then $\epsilon_\Lambda(\sigma') \leq \epsilon_\Lambda(\sigma)$.

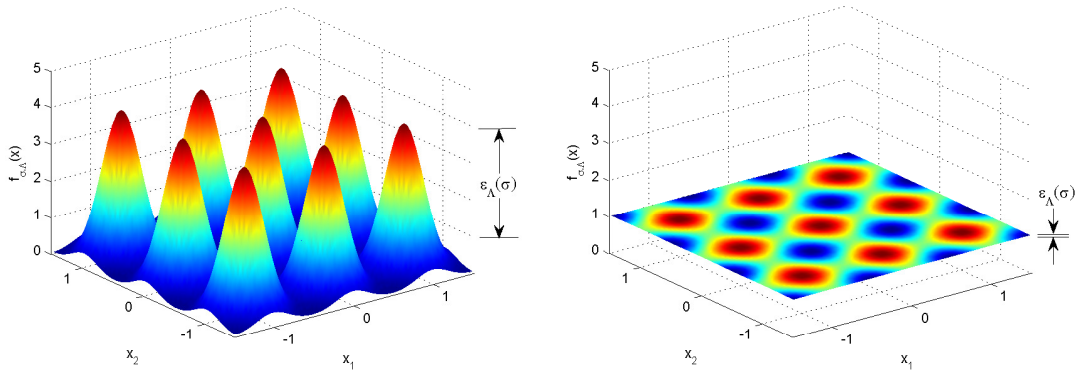
Figure 4.3 illustrates the flatness factor and lattice Gaussian distribution at different VNRs for the $\mathbb{Z}^2$ lattice. When the VNR is high (Figure 4.3(a)), $\epsilon_\Lambda(\sigma)$ is large and the Gaussians are well separated, implying reliable decoding is possible; this scenario is desired in communications. When the VNR is low (Figure 4.3(b)), $\epsilon_\Lambda(\sigma)$ is small and the distribution is nearly uniform, implying reliable decoding is impossible; this scenario is desired in security and will be pursued in following sections.

The next result, which slightly improves upon Lemma 4.3 in [166], shows that the variance per dimension of the discrete Gaussian $D_{\Lambda,\sigma,\mathbf{c}}$ is not far from σ^2 when the flatness factor is small.

Lemma 4.12 (Variance of discrete Gaussian) *Let L be sampled from the Gaussian distribution $D_{\Lambda,\sigma,\mathbf{c}}$. If $\varepsilon \coloneqq \epsilon_\Lambda\left(\sigma/\sqrt{\frac{\pi}{\pi-1/e}}\right) < 1$, then*

$$\left|\mathbb{E}\left[\|\mathsf{L}-\mathbf{c}\|^2\right] - n\sigma^2\right| \leq \frac{2\pi\varepsilon}{1-\varepsilon}\sigma^2. \tag{4.7}$$

[1] This definition differs slightly from the one in [166], where σ is scaled by a constant factor $\sqrt{2\pi}$ (i.e., $s = \sqrt{2\pi}\sigma$).



(a) $\gamma_\Lambda(\sigma) = 8\pi, \epsilon_\Lambda(\sigma) = 3$. (b) $\gamma_\Lambda(\sigma) = \pi, \epsilon_\Lambda(\sigma) = 0.0075$.

Figure 4.3: Lattice Gaussian distribution and flatness factor for $\mathbb{Z}^2$ (a) at high VNR where $\epsilon_\Lambda(\sigma)$ is large and the Gaussians are well separated, and (b) at low VNR where $\epsilon_\Lambda(\sigma)$ is small and the distribution is nearly uniform.

From the maximum-entropy principle [50, Chap. 11], it follows that the discrete Gaussian distribution maximizes the entropy given the average energy and given the same support over a lattice. The following lemma further shows that if the flatness factor is small, the entropy of a discrete Gaussian $D_{\Lambda,\sigma,\mathbf{c}}$ is almost equal to the differential entropy of a continuous Gaussian vector of variance σ^2 per dimension, minus $\log V(\Lambda)$, which corresponds to the entropy of a uniform distribution over the fundamental region of Λ.

Lemma 4.13 (Entropy of discrete Gaussian) *Let* $\mathsf{L} \sim D_{\Lambda,\sigma,\mathbf{c}}$. *If* $\varepsilon \triangleq \epsilon_\Lambda\left(\sigma/\sqrt{\frac{\pi}{\pi-1/e}}\right) < 1$, *then the entropy of* L *satisfies*

$$\left|\mathbb{H}(\mathsf{L}) - \left[n\log(\sqrt{2\pi e}\sigma) - \log V(\Lambda)\right]\right| \le \varepsilon',$$

where $\varepsilon' = -\log(1-\varepsilon) + \frac{\pi\varepsilon}{1-\varepsilon}$.

The following lemma by Regev [193, Claim 3.9] shows that if the flatness factor is small, the sum of a discrete Gaussian and a continuous Gaussian is very close to a continuous Gaussian (see figure 4.4).

Lemma 4.14 (Regev's Lemma) *Let* $\mathbf{c} \in \mathbb{R}^n$ *be any vector, and* $\sigma_0, \sigma > 0$. *Consider the continuous distribution* g *on* $\mathbb{R}^n$ *obtained by adding a continuous Gaussian of variance* σ^2 *to a discrete Gaussian* $D_{\Lambda-\mathbf{c},\sigma_0}$*:*

$$g(\mathbf{x}) = \frac{1}{f_{\sigma,\Lambda}(\mathbf{c})} \sum_{\mathbf{t}\in\Lambda-\mathbf{c}} f_{\sigma_0}(\mathbf{t}) f_\sigma(\mathbf{x}-\mathbf{t}).$$

If $\varepsilon \coloneqq \epsilon_\Lambda\left(\frac{\sigma_0\sigma}{\sqrt{\sigma_0^2+\sigma^2}}\right) < \frac{1}{2}$, *then* $\frac{g(\mathbf{x})}{f_{\sqrt{\sigma_0^2+\sigma^2}}(\mathbf{x})}$ *is uniformly close to* 1*:*

$$\forall \mathbf{x} \in \mathbb{R}^n, \quad \left|\frac{g(\mathbf{x})}{f_{\sqrt{\sigma_0^2+\sigma^2}}(\mathbf{x})} - 1\right| \le 4\varepsilon. \tag{4.8}$$

In particular, the distribution $g(\mathbf{x})$ *is close to the continuous Gaussian density* $f_{\sqrt{\sigma_0^2+\sigma^2}}$ *in* L^1 *distance:*

$$\mathbb{V}\left(g, f_{\sqrt{\sigma_0^2+\sigma^2}}\right) \le 4\varepsilon.$$

Secrecy-good lattices Now, we introduce the notion of secrecy-good lattices. Roughly speaking, a lattice is good for secrecy if its flatness factor is small.

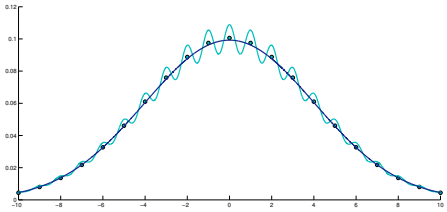


Figure 4.4: An illustration of Regev's Lemma for $\sigma = 4$ and $\sigma_0 = 0.4$. The Gaussian distribution $f_{\sqrt{\sigma_0^2+\sigma^2}}$ is plotted in dark blue, and the distribution g in light blue.

Definition 4.15 (Secrecy-good lattices) *A sequence of lattices $\Lambda^{(n)}$ is* secrecy-good *if for all fixed $\gamma_{\Lambda^{(n)}}(\sigma) < 2\pi$, $\forall c > 0$,*

$$\epsilon_{\Lambda^{(n)}}(\sigma) = o\left(\frac{1}{n^c}\right), \tag{4.9}$$

i.e. the flatness factor vanishes super-polynomially.

The following result guarantees the existence of sequences of lattices whose flatness factor vanish exponentially as the dimension $n \to \infty$, provided that the VNR < 2π.

Theorem 4.16 (Existence of secrecy-good lattices) *For any $\sigma > 0$ and $\delta > 0$, for fixed VNR $\gamma_{\Lambda^{(n)}}(\sigma) < 2\pi$, there exists a sequence of lattices $\Lambda^{(n)}$ such that*

$$\epsilon_{\Lambda^{(n)}}(\sigma) \leq (1+\delta) \cdot \left(\frac{\gamma_{\Lambda^{(n)}}(\sigma)}{2\pi}\right)^{\frac{n}{2}}. \tag{4.10}$$

The proof of this result is based on an average bound for the theta series on Loeliger's ensemble of mod-p lattices built from linear codes over finite fields using Construction A [157].

Remark 4.17 In fact, we can show a concentration result: $\forall \eta > 0$ there exists a lattice ensemble such that lattice sequences from this ensemble are secrecy-good with probability greater than $1 - \eta$.

It is worth mentioning that as soon as the VNR exceeds 2π, the L^∞ flatness factor increases exponentially, as can be seen from Proposition 4.7, since $\Theta_\Lambda(\tau) > 1\ \forall \tau > 0$.

Lattice Gaussian codes achieving strong secrecy and semantic security over the Gaussian wiretap channel We consider the Gaussian wiretap channel depicted in Fig. 4.5, whose outputs Y^n and Z^n at Bob and Eve's end respectively are given by

$$\begin{cases} \mathsf{Y}^n = \mathsf{X}^n + \mathsf{W}_b^n, \\ \mathsf{Z}^n = \mathsf{X}^n + \mathsf{W}_e^n, \end{cases} \tag{4.11}$$

where W_b^n, W_e^n are n-dimensional Gaussian vectors with zero mean and variance σ_b^2, σ_e^2 respectively. We suppose that $\sigma_e > \sigma_b$, i.e., Eve's channel is degraded with respect to Bob's channel (since otherwise the secrecy capacity would be zero). The transmitted codebook $\mathcal{C}$ must satisfy the average power constraint

$$\frac{1}{n}\mathbb{E}\left[\|\mathsf{X}^n\|\right] \leq P. \tag{4.12}$$

We denote this wiretap channel by $W(\sigma_b, \sigma_e, P)$.

Lattice Gaussian Coding We consider a chain of nested lattices $\Lambda_e \subset \Lambda_b$ such that Λ_b is AWGN-good [83], and Λ_e is secrecy good, with nesting ratio $|\Lambda_b/\Lambda_e| = e^{Rn}$. The existence of this chain is proven in Appendix II of [J8].

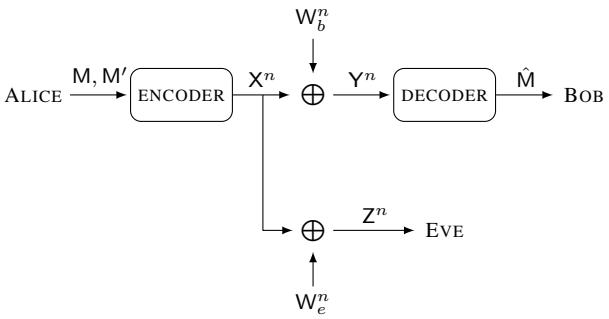


Figure 4.5: The Gaussian wiretap channel.

We consider a message set $\mathcal{M}_n = \{1, \ldots, e^{nR}\}$, and a one-to-one function which associates each message $m \in \mathcal{M}_n$ to a coset $\boldsymbol{\lambda}_m \in \Lambda_b/\Lambda_e$. We can identify the quotient group Λ_b/Λ_e with a set of coset representatives $\Lambda_b \cap \mathcal{R}(\Lambda_e)$ for any fundamental region $\mathcal{R}(\Lambda_e)$. Good choices of fundamental region $\mathcal{R}(\Lambda_e)$ (e.g., the fundamental parallelepiped) can result in low-complexity implementation of the encoder and decoder. Note that we make no *a priori* assumption on the distribution of the message M.

In order to encode the message $m \in \mathcal{M}_n$, Alice samples X_m^n from $D_{\Lambda_e+\boldsymbol{\lambda}_m,\sigma_s}$; equivalently, Alice transmits $\boldsymbol{\lambda} + \boldsymbol{\lambda}_m$ where $\boldsymbol{\lambda} \sim D_{\Lambda_e,\sigma_s,-\boldsymbol{\lambda}_m}$. It is worth mentioning that the distribution $D_{\Lambda_e+\boldsymbol{\lambda}_m,\sigma_s}$ is always centered at $\mathbf{0}$ for all bins (see 4.2). This is key for the conditional output distributions corresponding to different m to converge to the same distribution.

We choose $\sigma_s^2 = P$ in order to satisfy the average power constraint (3.32) asymptotically, thanks to Lemma 4.12, provided that $\epsilon_{\Lambda_e}\left(\sigma_s/\sqrt{\frac{\pi}{\pi-1/e}}\right) \to 0$.

Theorem 4.18 *On the wiretap channel $W(\sigma_b, \sigma_e, P)$, the proposed wiretap coding scheme with $\sigma_s^2 = P$ achieves strong secrecy for any message distribution p_{M} (and thus semantic security) for any secrecy rate (in nats)*

$$R < \frac{1}{2}\log\frac{\sigma_e^2}{\sigma_b^2} - \frac{1}{2}.$$

That is, the achievable rates are $1/2$ nat from the secrecy capacity.

The proof of Theorem 4.18 can be found in Section V of [J8]. We summarize the main ideas here. First, we show that Bob can reliably decode the confidential message using MMSE lattice decoding. Note that since the lattice points are not equally probable a priori in the lattice Gaussian coding, MAP decoding is not the same as standard ML decoding. One can show that MAP decoding is equivalent to Euclidean lattice decoding of Λ_b using a renormalized metric that is asymptotically close to the MMSE metric. Under this metric, the equivalent noise in Bob's channel can be written as the sum of discrete Gaussian variable and Gaussian noise. By Regev's Lemma 4.14, the distribution of this equivalent noise is close in L^1 distance to a Gaussian distribution. Since the fine lattice Λ_b is AWGN good, we have reliability as long as $\gamma_{\Lambda_b}(\tilde{\sigma}_b) > 2\pi e$ [83], where $\tilde{\sigma}_b = \frac{\sigma_s\sigma_b}{\sqrt{\sigma_s^2+\sigma_b^2}}$. Note that *no dither is required* to ensure that the equivalent noise is almost independent of the transmitted vector.

Next, we bound the leakage to the eavesdropper. Lemma 4.14 implies that the conditional output distributions $p_{\mathsf{Z}^n|\mathsf{M}=m}$ are close to the continuous Gaussian distribution $f_{\sqrt{\sigma_s^2+\sigma_e^2}}$ in variational distance if $\varepsilon_n = \epsilon_{\Lambda_e}(\tilde{\sigma}_e)$ is small, where $\tilde{\sigma}_e = \sigma_s\sigma_e/\sqrt{\sigma_s^2+\sigma_e^2}$. An upper bound on the amount of leaked information then follows directly from Lemma 4.4. In order for ε_n to vanish, a sufficient condition is $\gamma_{\Lambda_e}(\tilde{\sigma}_e) < 2\pi$. Lemma 4.13 guarantees that the bin rate is above Eve's channel capacity, in order to obtain a resolvability scheme.

Follow-up work. Lattice Gaussian signalling can also be used for transmission over Gaussian channels *without* secrecy constraints; in this setting, a follow-up work by Ling and Belfiore [152] showed that it achieves capacity for $\mathrm{SNR} > e$. This SNR condition can be removed using a dither [35].

For Gaussian wiretap channels, we conjecture that the $1/2$ nat gap in Theorem 4.18 can be removed by using the L^1 version of the flatness factor, which will be introduced in Section 4.3.

After our paper was submitted, other works have proposed wiretap code constructions for Gaussian channels. Tyagi and Vardy's approach [215] achieves the strong secrecy capacity of the Gaussian wiretap channel using 2-universal hash functions. The *polar lattice* construction built from polar codes using Construction D in [155] also achieves the secrecy capacity.

The flatness factor criterion can be extended to more general fading and MIMO channel models, as will be shown in the next Section.

4.2 Almost universal wiretap codes for MIMO wireless channels

In the follow-up work [J10] in collaboration with Cong Ling and Roope Vehkalahti, we considered the design of lattice codes for fading and multi-antenna wiretap channels, and proposed a new construction of wiretap codes from ideal lattices which builds upon our universal code construction in [J9] (see Section 3.4). Note that ideal lattices from number fields had already been considered for fading wiretap channels under an error probability criterion [18, 127].

In order to tackle the MIMO wiretap channel, we need to extend the definition of flatness factor to the multivariate case, which is related to the extended notion of smoothing parameter in [182].

Let $f_{\sqrt{\Sigma},\mathbf{c}}(\mathbf{z})$ denote the k-dimensional circularly symmetric complex normal distribution with mean $\mathbf{c}$ and covariance matrix Σ:

$$f_{\sqrt{\Sigma},\mathbf{c}}(\mathbf{z}) = \frac{1}{\pi^k \det(\Sigma)} e^{-(\mathbf{z}-\mathbf{c})^\dagger \Sigma^{-1} (\mathbf{z}-\mathbf{c})} \quad \forall \mathbf{z} \in \mathbb{C}^k.$$

We consider lattices of even dimension $n = 2k$ in the Euclidean space $\mathbb{R}^{2k}$, which is identified with the complex space $\mathbb{C}^k$. Given a lattice $\Lambda \subset \mathbb{C}^k$, we consider the Λ-periodic function

$$f_{\sqrt{\Sigma},\Lambda}(\mathbf{z}) = \sum_{\boldsymbol{\lambda} \in \Lambda} f_{\sqrt{\Sigma},\boldsymbol{\lambda}}(\mathbf{z}), \quad \forall \mathbf{z} \in \mathbb{C}^k.$$

Given a positive definite matrix $\Sigma \in M_k(\mathbb{C})$, the flatness factor $\epsilon_\Lambda(\sqrt{\Sigma})$ is defined as

$$\epsilon_\Lambda(\sqrt{\Sigma}) = \max_{\mathbf{z} \in \mathcal{R}(\Lambda)} \left| V(\Lambda) f_{\sqrt{\Sigma},\Lambda}(\mathbf{z}) - 1 \right|.$$

We have the following corollary of a result by Banaszczyk [15], which implies that the smoothing parameter is upper bounded by the minimum distance of the dual lattice [166]. In terms of flatness factor, we can state it as follows.

Lemma 4.19 *Suppose that Λ is an n-dimensional lattice, and let $c > \frac{1}{\sqrt{2\pi}}$, $C = c\sqrt{2\pi e} e^{-\pi c^2} < 1$. If $\tau \geq \frac{\sqrt{n}c}{\sqrt{\pi}\lambda_1(\Lambda)}$, then*

$$\epsilon_{\Lambda^*}(\tau) \leq \frac{C^n}{1 - C^n}. \tag{4.13}$$

Regev's lemma [193, Claim 3.9] (see Lemma 4.14) generalizes to the multivariate case as follows:

Lemma 4.20 *Let X_1 be sampled according to the discrete Gaussian distribution $D_{\Lambda+\mathbf{c},\sqrt{\Sigma_1}}$ and X_2 be sampled according to the continuous Gaussian $f_{\sqrt{\Sigma_2}}$. Let $\Sigma_0 = \Sigma_1 + \Sigma_2$ and $\Sigma^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1}$. Denote by $g(\mathbf{x})$ the density of the random variable $X = X_1 + X_2$. If*

$$\epsilon_\Lambda(\sqrt{\Sigma}) \leq \varepsilon \leq \frac{1}{2}, \tag{4.14}$$

then

$$\mathbb{V}(g, f_{\sqrt{\Sigma_0}}) \leq 4\varepsilon.$$

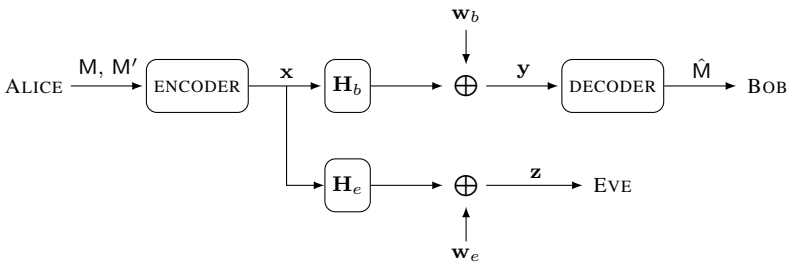


Figure 4.6: The fading wiretap channel.

Lattice codes for fading wiretap channels For simplicity, we first present our results in the single antenna case. We consider the channel model illustrated in Figure 4.6, where the outputs $\mathbf{y}, \mathbf{z} \in \mathbb{C}^k$ over k channel uses at Bob and Eve's end respectively are given by

$$\begin{cases} y_i = h_{b,i} x_i + w_{b,i}, \\ z_i = h_{e,i} x_i + w_{e,i}, \end{cases} \qquad i = 1, \ldots, k \tag{4.15}$$

and $w_{b,i}$, $w_{e,i}$ are i.i.d. complex Gaussian vectors with zero mean and variance σ_b^2, σ_e^2 per complex dimension. A confidential message M and an auxiliary message M' with rate R and R' respectively are encoded into $\mathbf{x}$. We denote by $\hat{\mathsf{M}}$ the estimate of the confidential message at Bob's end. We define $H_e = \mathrm{diag}(h_{e,1}, \ldots, h_{e,k})$, $H_b = \mathrm{diag}(h_{b,1}, \ldots, h_{b,k})$. The input $\mathbf{x}$ satisfies the average power constraint

$$\frac{1}{k} \sum_{i=1}^{k} |x_i|^2 \leq P. \tag{4.16}$$

We suppose that $h_{b,i}$, $h_{e,i}$ are isotropically invariant and that Bob and Eve's channel capacities C_b and C_e are well-defined. All rates are expressed in nats per complex channel use.

We assume that the weak law of large numbers (LLN) holds for Bob's channel: $\forall \delta > 0$

$$\lim_{k \to \infty} \mathbb{P} \left\{ \left| \frac{1}{k} \sum_{i=1}^{k} \ln \left(1 + \frac{P |h_{b,i}|^2}{\sigma_b^2} \right) - C_b \right| > \delta \right\} = 0, \tag{4.17}$$

This general setting includes the Gaussian channel, i.i.d. block fading channels where the size of the blocks is fixed and the number of blocks tends to infinity as well as all ergodic fading channels.

Moreover, we require a stricter condition for Eve's channel[2]:, i.e. the asymptotic rate of convergence in the LLN should be faster than $o\left(\frac{1}{k}\right)$: $\forall \delta' > 0$,

$$\lim_{k \to \infty} k \, \mathbb{P} \left\{ \left| \frac{1}{k} \sum_{i=1}^{k} \ln \left(1 + \frac{P |h_{e,i}|^2}{\sigma_e^2} \right) - C_e \right| > \delta' \right\} = 0 \tag{4.18}$$

This condition is satisfied for static channels, i.i.d. fading channels and i.i.d. block fading channels, and ergodic channels whose decay of large deviations is vanishing with rate $o\left(\frac{1}{k}\right)$. We suppose that Bob has perfect CSI of his own channel, and Eve has perfect CSI of both channels. Alice has no instantaneous CSI, apart from partial knowledge of channel statistics. More precisely, the knowledge of C_b and C_e and of the properties (4.17) and (4.18) is sufficient for Alice.

[2]Although a rate of convergence of the order $o\left(\frac{1}{k}\right)$ in the law of large numbers for Eve's channel seems to be necessary for strong secrecy, any rate of convergence is enough to guarantee weak secrecy, see Proposition 3.9 in [J10].

Lattice wiretap coding Let $\Lambda_e^{(k)} \subset \Lambda_b^{(k)}$ be a pair of nested lattices in $\mathbb{C}^k$ with nesting ratio $|\Lambda_b/\Lambda_e| = e^{kR}$, and volumes

$$V(\Lambda_e) = \frac{(\pi e P)^k}{e^{kR'}}, \quad V(\Lambda_b) = \frac{(\pi e P)^k}{e^{k(R+R')}}, \tag{4.19}$$

where $R' > 0$. Let $\mathcal{R}(\Lambda_e)$ be a fundamental region of Λ_e. We consider the secrecy scheme in Section 4.1, where each confidential message $m \in \mathcal{M} = \{1, \ldots, e^{kR}\}$ is associated to a coset leader $\boldsymbol{\lambda}_m \in \Lambda_b \cap \mathcal{R}(\Lambda_e)$. To transmit the message m, Alice samples $\mathbf{x} \in \Lambda_b$ from the discrete Gaussian $D_{\Lambda_e + \boldsymbol{\lambda}_m, \sigma_s}$ with $\sigma_s^2 = P$. We denote this lattice coding scheme by $\mathcal{C}(\Lambda_b, \Lambda_e)$. It follows from Lemma 4.12 and Remark 6 in [J8] that the power constraint (4.16) is verified under suitable conditions on the flatness factor. Similarly, using Lemma 4.13, we can show that the entropy rate $\mathbb{H}(\mathsf{M}')$ of the auxiliary message tends to R' as $k \to \infty$ under suitable flatness factor conditions. We omit details, which can be found in [J10].

We will show that the gap to capacity of our codes is essentially determined by the normalized product distance of the lattices Λ_b and Λ_e^*, which is a special case of the normalized minimum determinant defined in (3.36).

Definition 4.21 (Normalized product distance) *Given a complex lattice $\Lambda \subset \mathbb{C}^k$, its normalized product distance is given by*

$$\mathrm{Np}(\Lambda) = \frac{\inf_{\mathbf{x} \in \Lambda \setminus \{0\}} \prod_{i=1}^k |x_i|}{\sqrt{V(\Lambda)}}.$$

Theorem 4.22 *Consider the wiretap scheme $\mathcal{C}(\Lambda_b, \Lambda_e)$, and suppose that there exist positive constants $\mathrm{t}_b, \mathrm{t}_e$ such that*

$$\liminf_{k\to\infty} \mathrm{Np}(\Lambda_b)^{2/k} \geq \mathrm{t}_b, \quad \liminf_{k\to\infty} \mathrm{Np}(\Lambda_e^*)^{2/k} \geq \mathrm{t}_e. \tag{4.20}$$

If the main channel and the eavesdropper's channel verify the conditions (4.17) and (4.18), then the codes $\mathcal{C}(\Lambda_b, \Lambda_e)$ achieve strong secrecy for any message distribution p_{M}, and thus they achieve semantic security, if

$$R' > C_e + \ln\left(\frac{e}{\pi}\right) - \ln \mathrm{t}_e, \quad R + R' < C_b - \ln\left(\frac{4}{\pi e}\right) + \ln \mathrm{t}_b. \tag{4.21}$$

Thus, any strong secrecy rate

$$R < C_b - C_e - 2\ln\left(\frac{2}{\pi}\right) + \ln \mathrm{t}_b \mathrm{t}_e$$

is achievable with the proposed lattice codes.

Therefore, we established a simple design criterion where the normalized product distance of the lattice and its dual should be maximized simultaneously; in the special case of the Gaussian wiretap channel, the packing density of the lattice and its dual should be maximized (Proposition 3.19 in [J10]).

The key ideas of the proof of Theorem 4.22 are as follows. We focus on the secrecy condition, since the reliability condition can be proven with similar techniques as in [J9]. With CSI at the eavesdropper, the leakage can be expressed as

$$\mathbb{I}(\mathsf{M}; \mathbf{z}, H_e) = \mathbb{I}(\mathsf{M}; H_e) + \mathbb{I}(\mathsf{M}; \mathbf{z}|H_e) = \mathbb{I}(\mathsf{M}; \mathbf{z}|H_e) = \mathbb{E}_{H_e}\left[\mathbb{I}(p_{\mathsf{M}|H_e}; p_{\mathbf{z}|H_e})\right] = \mathbb{E}_{H_e}\left[\mathbb{I}(p_{\mathsf{M}}; p_{\mathbf{z}|H_e})\right].$$

We want to show that the *average* leakage with respect to the fading is small. First, we consider a fixed channel sequence $H_e = \mathrm{diag}(h_{e,1}, \ldots, h_{e,k})$. As in Section 4.1, we bound the leakage through the average variational distance of output distributions corresponding to different confidential messages to a continuous Gaussian.

Using Lemma 4.20 with $\Sigma_1 = H_e H_e^\dagger P$, $\Sigma_2 = \sigma_e^2 I$, we have

$$\mathbb{V}(p_{\mathbf{z}|H_e, \mathsf{M}=m}, f_{\sqrt{\Sigma_0}}) \leq 4\varepsilon_k$$

provided that

$$\epsilon_{H_e\Lambda_e}(\sqrt{\Sigma}) = \epsilon_{\sqrt{\Sigma}^{-1}H_e\Lambda_e}(1) \leq \varepsilon_k \leq \frac{1}{2}, \tag{4.22}$$

where we define $\Sigma_0 = H_e H_e^\dagger P + \sigma_e^2 I$, $\Sigma^{-1} = \frac{(H_e H_e^\dagger)^{-1}}{P} + \frac{I}{\sigma_e^2}$. That is, the flatness factor of the faded lattice should be small.

The upper bound (4.13) in Lemma 4.19 allows us to bound the flatness factor of the faded lattice in terms of the minimum distance of its dual, which should not be too small. This minimum distance can be lower bounded in terms of the product distance as follows:

$$d_{\min}((\sqrt{\Sigma}^{-1} H_e \Lambda_e)^*) = d_{\min}(\sqrt{\Sigma}(H_e^\dagger)^{-1}\Lambda_e^*) = \min_{\mathbf{x}\in\Lambda_e^*\setminus\{0\}} \left\|\sqrt{\Sigma}(H_e^\dagger)^{-1}\mathbf{x}\right\|$$
$$\geq \frac{\sqrt{k}\sqrt{P}}{\prod_{i=1}^k \left(1+\frac{P}{\sigma_e^2}|h_{e,i}|^2\right)^{\frac{1}{2k}}} \prod_{i=1}^k |x_i|^{\frac{1}{k}}$$

where the last step follows from the arithmetic mean - geometric mean inequality.

For a random channel sequence $H_e = \mathrm{diag}(h_{e,1}, \ldots, h_{e,k})$, the conclusion follows from the bound (4.13) by invoking the law of large numbers (4.18). More details can be found in [J10].

MIMO wiretap channel Our results can be generalized to a MIMO fading channel model where Alice is equipped with n antennas, while Bob and Eve have n_b and n_e antennas respectively. For simplicity, we assume that $n_b \geq n$ and $n_e \geq n$. Transmission takes place over k quasi-static fading blocks of delay $T = n$, and the transmitted codeword is of the form $X = (X_1, \ldots, X_k)$, where the matrix $X_i \in M_n(\mathbb{C})$ is sent during the i-th block.

The outputs Y and Z at Bob and Eve's end respectively are given by

$$\begin{cases} Y = H_b X + W_b, \\ Z = H_e X + W_e, \end{cases} \tag{4.23}$$

where $H_b = \mathrm{diag}(H_{b,1}, \ldots, H_{b,k}) \in M_{n_b k \times nk}(\mathbb{C})$, $H_e = \mathrm{diag}(H_{e,1}, \ldots, H_{e,k}) \in M_{n_e k \times nk}(\mathbb{C})$. The coefficients of the noise matrices W_b and W_e are i.i.d. circularly symmetric complex Gaussian with zero mean and variance σ_b^2, σ_e^2 per complex dimension. The input X satisfies the average power constraint (per channel use)

$$\frac{1}{nk}\sum_{i=1}^k \|X_i\|^2 \leq P. \tag{4.24}$$

The average power per symbol is $\sigma_s^2 = \frac{P}{n}$. We denote by $\rho_b = \frac{\sigma_s^2}{\sigma_b^2}$ and $\rho_e = \frac{\sigma_s^2}{\sigma_e^2}$ the signal-to-noise ratios for Bob and Eve respectively. We suppose that $\{H_{b,i}\}$, $\{H_{e,i}\}$ are isotropically invariant channels such that the channel capacities C_b and C_e are well-defined and $\forall \gamma, \gamma' > 0$,

$$\lim_{k\to\infty} \mathbb{P}\left\{\left|\frac{1}{k}\sum_{i=1}^k \ln\det\left(I_{n_b} + \rho_b H_{b,i}^\dagger H_{b,i}\right) - C_b\right| > \gamma\right\} = 0 \tag{4.25}$$

$$\lim_{k\to\infty} k\,\mathbb{P}\left\{\left|\frac{1}{k}\sum_{i=1}^k \ln\det\left(I_{n_e} + \rho_e H_{e,i}^\dagger H_{e,i}\right) - C_e\right| > \gamma'\right\} = 0 \tag{4.26}$$

As before, we suppose that Alice has statistical CSI only, Bob has perfect CSI of his own channel, and Eve has perfect CSI of her channel and of Bob's.

A confidential message M and an auxiliary message M' with rate R and R' respectively are encoded into the multi-block codeword X.

Remark 4.23 For general channels the strong secrecy capacity still appears to be unknown in this setting. In [150] it was shown that the weak secrecy capacity

$$C_s^w = C_b - C_e$$

for i.i.d. fading wiretap channels such that Bob and Eve's fadings are independent.

Multiblock lattice wiretap coding. Let $\Lambda_e \subset \Lambda_b$ be a pair of nested multiblock matrix lattices in $M_{nk\times n}(\mathbb{C})$ such that $\Lambda_e \subset \Lambda_b$ and $|\Lambda_b/\Lambda_e| = e^{nkR}$, with volumes scaling as follows:

$$V(\Lambda_e) = \frac{(\pi e\sigma_s^2)^{n^2k}}{e^{nkR'}}, \quad V(\Lambda_b) = \frac{(\pi e\sigma_s^2)^{n^2k}}{e^{nk(R+R')}}, \tag{4.27}$$

where $R' > 0$. Each message $m \in \mathcal{M} = \{1, \dots, e^{nkR}\}$ is mapped to a coset leader $X^{(m)} \in \Lambda_b \cap \mathcal{R}(\Lambda_e)$, where $\mathcal{R}(\Lambda_e)$ is a fundamental region of Λ_e. In order to transmit the message m, Alice samples X from the discrete Gaussian $D_{\Lambda_e+X^{(m)},\sigma_s}$ where $\sigma_s^2 = \frac{P}{n}$. We denote this coding scheme by $\mathcal{C}(\Lambda_b, \Lambda_e)$.

Let $\{\Lambda^{(n,k)}\}$ be a sequence of n^2k-dimensional multi-block matrix lattices in $M_{nk\times n}(\mathbb{C})$ (see Section 3.4). We consider scaled versions $\Lambda_b = \alpha_b\Lambda^{(n,k)}$, $\Lambda_e = \alpha_e\Lambda^{(n,k)}$ such that $\Lambda_e \subset \Lambda_b$ and $|\Lambda_b/\Lambda_e| = e^{nkR}$. Given rates R, R', we denote the corresponding multi-block lattice coding scheme by $\mathcal{C}(\Lambda^{(n,k)}, R, R')$.
Given a lattice Λ, let $\delta(\Lambda)$ denote its normalized minimum determinant defined in (3.36). Then we have the following multi-antenna extension of Theorem 4.22, where the normalized product distance is replaced by the normalized minimum determinant.

Theorem 4.24 *Consider the multi-block wiretap coding scheme $\mathcal{C}(\Lambda_b, \Lambda_e)$ defined previously, and suppose that*

$$\liminf_{k\to\infty} \delta(\Lambda_e^*)^{\frac{2}{k}} \geq \mathrm{d}_e, \quad \liminf_{k\to\infty} \delta(\Lambda_b)^{\frac{2}{k}} \geq \mathrm{d}_b \tag{4.28}$$

for some positive constants d_e, d_b. If the main channel and the eavesdropper's channel verify the conditions (4.25) and (4.26) respectively, then $\mathcal{C}(\Lambda_b, \Lambda_e)$ achieves strong secrecy and semantic security for all rates

$$R < C_b - C_e - 2n\ln\left(\frac{2n}{\pi}\right) + \ln \mathrm{d}_b\mathrm{d}_e. \tag{4.29}$$

Coding scheme based on division algebras with constant root discriminant Recall that in Section 3.4 we have established the existence of multiblock lattices $\{L_{n,k}\}$ arising from the left regular representation of division algebras over Hilbert class fields, satisfying the volume condition (3.46). Taking two scaled copies of such a lattice for Λ_b and Λ_e, we obtain the following.

Corollary 4.25 *If the main channel and the eavesdropper's channel verify the conditions (4.25) and (4.26), then the multi-block wiretap coding scheme $\mathcal{C}(L_{n,k}, R, R')$ achieves strong secrecy and semantic security for all rates*

$$R < C_b - C_e - 2n\ln\left(\frac{nG\beta^{\frac{n-1}{n}}}{\pi}\right),$$

where $G \approx 92.368$ and $\beta = 23^{1/10}$.

In conclusion, we have proposed an algebraic construction of lattices which achieve strong secrecy and semantic security for all rates $R < C_b - C_e - \kappa$, where C_b and C_e are Bob and Eve's channel capacities respectively, and κ is an explicit constant gap which depends on the geometric invariants of the chosen lattices. Our codes are almost universal in the sense that given C_b and C_e, the same code is good for secrecy for a wide range of fading models. Since for many of the channel models we consider we don't know the actual strong secrecy capacity, the achievable rate $C_b - C_e - \kappa$ provides a lower bound.

Remark 4.26 (Compound channel model) We also considered a compound channel model with the standard definition of compound capacity, and proved that if we consider a more restrictive uncertainty set, then we can guarantee uniform bounds for the error probability and the leaked information, and our codes achieve a constant gap κ to the standard compound capacity. Details can be found in Section VI of [J10].

Open problems and related works Our model assumes perfect CSI of the legitimate channel at the receiver. This assumption is not realistic for a fast fading channel, since in practice most of the available time slots would have to be used to transmit training symbols for channel estimation. However, our channel model is not limited to fast fading, but only assumes the weak law of large numbers for the channel statistics. This includes for example a block fading model, where some fraction of each block can be used for channel estimation and the rest is left for data transmission. We have also provided some results for the arbitrarily varying fading model in Section VI of [J10], where Bob's channel oscillates most of the time above a certain threshold and Eve's channel oscillates mostly below another threshold, without necessarily converging in mean.

Several technical improvements are needed before our lattice code construction can be implemented in practice. As mentioned in Section 3.4, although the proposed families of lattices are deterministic, their construction is not efficient. Moreover, our construction incurs a large gap to the secrecy capacity. This gap might be reduced by taking suitable ideals of the ring of integers in the number field case, or ideals of orders in the division algebra case.

After this paper was first submitted, the Generalized Construction A was extended to a MIMO wiretap setting [37]. The problem of finding well-performing lattice codes for fading wiretap channels in fixed dimension is considered in [62], where it is shown that one can restrict the search to the set of *well-rounded lattices*, i.e. lattices such that all the successive minima are equal to the first minimum. Constructions of ideal lattices from number fields with small average flatness factor are also provided in [62].
It is interesting to note that the same lattices built from Hilbert class field towers have been considered in cryptography [183]. This connection warrants further analysis in view of the recent developments in lattice-based post-quantum cryptography.
The connection between secrecy and the channel coding goodness of dual codes also calls for further investigation. For example, dual codes play a role in the design of LDPC codes for binary erasure wiretap channels [214, 207].

4.3 Secret key generation from Gaussian sources using lattices

Wireless channels in a rich scattering environment provide a high-entropy source of physical randomness which is time-varying, location-dependent and hard to predict. Moreover, it is well-known that the forward and backward channels between a pair of antennas are reciprocal, especially in the time-division duplex (TDD) mode. Thus, the simultaneous measure of a reference signal sent in both directions between Alice and Bob results in highly similar observations[3]. In contrast, an eavesdropper located at a distance of several wavelengths from Alice and Bob will experience an almost uncorrelated realization of the channel. This source of randomness could be used to generate session keys from mobile terminals in a decentralized manner.

The works by Maurer [165] and Ahlswede and Csiszár [4] in information theory have shown that two legitimate users can exploit correlated observations of noisy channels to generate a shared secret key, even in the presence of an adversary who has access to a third sequence of observations and can intercept all the messages exchanged by the users over a public channel. Their analysis relies on the assumption of discrete random sources over countable alphabets. The extension of this framework to continuous sources was considered in [225, 226, 154]. In [174], the authors consider a multiterminal scenario for secret key generation from correlated Gaussian sources in the special case where the eavesdropper has only access to the public channel, and show that the optimal strong secret key rate can be achieved using lattice codes.

In a collaboration with Cong Ling and Matthieu Bloch [C12],[J13], we consider the problem of key generation between two terminals, Alice and Bob, who observe correlated Gaussian sequences X^n and Y^n. Unlike [174], we assume that the eavesdropper, besides observing the exchanges over the public channel, also obtains a correlated sequence Z^n. For simplicity, we suppose that a single round of unidirectional public communication takes place from Alice to Bob.

We consider the source model illustrated in Figure 4.7, in which X^n, Y^n, Z^n are generated by an i.i.d. memoryless source p_{XYZ} whose components are jointly Gaussian with zero mean. The distribution is fully described by

[3] In practice, though, reciprocity is not complete because of thermal noise and imperfect synchronization of the reference signals.

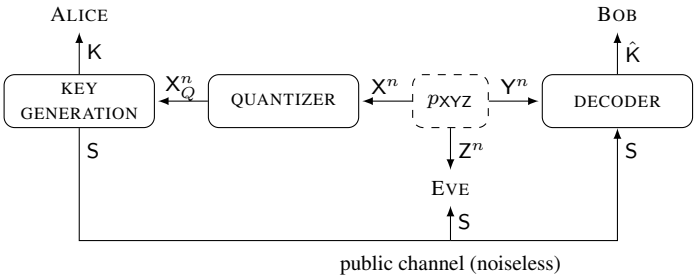


Figure 4.7: Secret key generation in the presence of an eavesdropper with communication over a public channel.

the variances σ_x^2, σ_y^2, σ_z^2 and the correlation coefficients ρ_{xy}, ρ_{xz}, ρ_{yz}. We can write [225, Eq. (6)]:

$$\begin{cases} \mathsf{X}^n = \rho_{xy}\dfrac{\sigma_x}{\sigma_y}\mathsf{Y}^n + \mathsf{W}_1^n, \\ \mathsf{X}^n = \rho_{xz}\dfrac{\sigma_x}{\sigma_z}\mathsf{Z}^n + \mathsf{W}_2^n, \end{cases} \tag{4.30}$$

where W_1^n and W_2^n are i.i.d. zero-mean Gaussian noise vectors of variances

$$\sigma_1^2 = \sigma_x^2(1-\rho_{xy}^2), \quad \sigma_2^2 = \sigma_x^2(1-\rho_{xz}^2), \tag{4.31}$$

respectively, such that $\sigma_2 > \sigma_1$, i.e., the source model is *degraded*. Further, W_1^n is independent of Y^n, and W_2^n is independent of Z^n.

Alice computes a public message S and a secret key K from her observation X^n; she then transmits S over the public channel. From this message and his own observation Y^n, Bob reconstructs a key $\hat{\mathsf{K}}$.

Let $\mathcal{K}_n$ and $\mathcal{S}_n$ be the sets of secret keys and public messages respectively. A *secret key rate - public rate pair* (R_K, R_P) is achievable if there exists a sequence of protocols with

$$\liminf_{n\to\infty} \frac{1}{n}\log|\mathcal{K}_n| \geq R_K, \quad \limsup_{n\to\infty} \frac{1}{n}\log|\mathcal{S}_n| \leq R_P,$$

such that the following properties hold:

$$\lim_{n\to\infty} \log|\mathcal{K}_n| - \mathbb{H}(\mathsf{K}) = 0 \qquad \text{(uniformity)}$$
$$\lim_{n\to\infty} \mathbb{P}\left\{\mathsf{K} \neq \hat{\mathsf{K}}\right\} = 0 \qquad \text{(reliability)}$$
$$\lim_{n\to\infty} \mathbb{I}(\mathsf{K}; \mathsf{S}, \mathsf{Z}^n) = 0 \qquad \text{(strong secrecy).}$$

Following [225], we denote

$$\mathcal{R}(\mathsf{X},\mathsf{Y},\mathsf{Z}) = \{(R_P, R_K) : \ (R_P, R_K) \text{ is achievable}\}.$$

The optimal trade-off between secret key rate and public rate was derived in [225]. For the source model (4.30), given public rate R_P, the secret key rate is upper bounded by

$$R_K \leq \bar{R}_K(R_P) = \frac{1}{2}\log\left(e^{-2R_P} + \frac{\sigma_2^2}{\sigma_1^2}(1-e^{-2R_P})\right). \tag{4.32}$$

The secret key capacity of the Gaussian source model (4.30) is defined as the maximum achievable secret key rate with unlimited public communication and is given by [225]

$$C_s = \sup\{R_K \text{ such that } \exists R_P \geq 0 \ : \ (R_P, R_K) \in \mathcal{R}(\mathsf{X},\mathsf{Y},\mathsf{Z})\} = \frac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2}. \tag{4.33}$$

Additional notation. To simplify notation, we define $\hat{\mathsf{Y}}^n = \rho_{xy}\frac{\sigma_x}{\sigma_y}\mathsf{Y}^n$ and $\hat{\mathsf{Z}}^n = \rho_{xz}\frac{\sigma_x}{\sigma_z}\mathsf{Z}^n$, so that

$$\begin{cases} \mathsf{X}^n = \hat{\mathsf{Y}}^n + \mathsf{W}_1^n, \\ \mathsf{X}^n = \hat{\mathsf{Z}}^n + \mathsf{W}_2^n, \end{cases} \tag{4.34}$$

where $\hat{\mathsf{Y}}^n$ and W_1^n are independent, and $\hat{\mathsf{Z}}^n$ and W_2^n are independent. We denote the variances of $\hat{\mathsf{Y}}^n$ and $\hat{\mathsf{Z}}^n$ by $\hat{\sigma}_y = \rho_{xy}\sigma_x = \sqrt{\sigma_x^2 - \sigma_1^2}$ and $\hat{\sigma}_z = \rho_{xz}\sigma_x = \sqrt{\sigma_x^2 - \sigma_2^2}$ respectively.

Main contributions and related works. Our main contribution is to show that, in the case of a degraded source model, the optimal secret key rate can be achieved by a complete lattice-coding scheme considerably different from and much simpler than [174], which also verifies the optimal trade-off in [225].

Typically, secret key generation is composed of two distinct procedures:

- *information reconciliation*: Alice and Bob locate the similarities in their sequences by exchanging error-correction bits. In the case of unidirectional communication, information reconciliation can be seen as a lossy source coding problem with side information (Wyner-Ziv problem) [227].
- *privacy amplification*: Alice and Bob extract from the common sequence a secret key which is independent of Eve's observations.

Information reconciliation and Wyner-Ziv coding. Our strategy for information reconciliation follows the outline of [225, 174]: first, the source X^n is quantized; then, a public message is generated in the manner of Wyner-Ziv coding, so that Bob can decode the quantized variable using the sequence Y^n as side information. The existence of sequences of good nested lattices for Wyner-Ziv coding has been established in [235].

Privacy amplification and randomness extraction. Our privacy amplification strategy is based on the concept of *channel intrinsic randomness*, or the maximum bit rate that can be extracted from a channel output independently of its input [28, 112]. We propose a new technique to extract the randomness, by reducing the source modulo a suitable secrecy-good lattice. We note that nearest-neighbor quantization is not needed, and we only need to implement the mod $\mathcal{R}(\Lambda)$ operation, which can be performed in polynomial time for many fundamental regions $\mathcal{R}(\Lambda)$. In particular, we can choose the fundamental parallelepiped.

The L^1 flatness factor. In the first version of this work [C12], we proposed a simple lattice-based key generation scheme which does not require dithering, and achieves a secret key rate up to half a nat from the optimal. In this scheme, quantization is performed through nearest-neighbor decoding, and privacy amplification is done by reducing modulo a secrecy-good lattice with respect to the L^∞ flatness factor.

In the recent follow-up work [J13], we improve this scheme to achieve the optimal secret key rate and the optimal public rate / secret key rate trade-off. To do so, we introduce an extended notion of flatness factor in which the L^∞ distance is replaced by the L^1 distance. This L^1 flatness condition is satisfied by a wider range of variance parameters, resulting in improved volume conditions for the chain of lattices under consideration, which allows us to achieve the secret key capacity. We prove the existence of lattices with vanishing L^1 flatness factor by leveraging an existence result for resolvability codes for regular channels [113].

We note that the L^1 smoothing parameter was already considered in [48, 60], while L^1 and KL flatness factors were used implicitly earlier in [155, p. 1656]. An upper bound on the L^1 flatness factor based on the Cauchy-Schwarz inequality was given in [168]. The independent work [65] studied the L^1 smoothing parameters both for lattices and for codes. In the case of Gaussian distributions modulo lattices, [65] obtains the same bound on the L^1 smoothing parameter as in our paper, but by a different approach, by decomposing the discrete Gaussian distribution into a convex combination of uniform ball distributions.

Randomized quantization technique. Furthermore, we replace nearest-neighbor decoding by *randomized quantization* [182] with dithering. Essentially, this technique allows to round a continuous Gaussian into a discrete Gaussian distribution with slightly larger variance, provided that the flatness factor of the lattice is small. We partially extend the result of [182] under an L^1 flatness factor criterion. We show that randomized quantization with uniform dithering (where the dither is known by all parties, including the eavesdropper) achieves the optimal trade-off between public communication rate and secret key rate established in [225]. Since the L^1 flatness factor is only an average condition, dithering is required in order to obtain almost uniform keys.

Before describing our key generation protocol, we introduce the new technical tools that will be needed.

The L^1 flatness factor First, we introduce a weaker notion of flatness based on the L^1 distance. We will denote by $f_{\sigma,\mathcal{R}(\Lambda)} = f_{\sigma,\Lambda_{|\mathcal{R}(\Lambda)}}$ the restriction of $f_{\sigma,\Lambda}$ to the fundamental region $\mathcal{R}(\Lambda)$. Note that $f_{\sigma,\mathcal{R}(\Lambda)}$ is the probability density of $\bar{\mathsf{X}}^n = [\mathsf{X}^n] \bmod \mathcal{R}(\Lambda)$.

Definition 4.27 (L^1 flatness factor) *Given a lattice Λ, a fundamental region $\mathcal{R}(\Lambda)$ and $\sigma > 0$, we define the L^1* flatness factor *as follows:*

$$\epsilon^1_\Lambda(\sigma) = \int_{\mathcal{R}(\Lambda)} \left| f_{\sigma,\Lambda}(\mathbf{x}) - \frac{1}{V(\Lambda)} \right| d\mathbf{x} = \mathbb{V}(f_{\sigma,\mathcal{R}(\Lambda)}, \mathcal{U}_{\mathcal{R}(\Lambda)}). \tag{4.35}$$

Remark 4.28 For any lattice Λ, $\forall \sigma > 0$, we have $\epsilon^1_\Lambda(\sigma) \leq \epsilon_\Lambda(\sigma)$.

The L^1 flatness factor is related to the L^1 smoothing parameter, which was discussed in [48, 60].

The following Lemma confirms the intuition that folded additive Gaussian noise with larger variance looks more uniform:

Lemma 4.29 *The L^1 flatness factor is monotonic, i.e. for any lattice Λ, $\forall \sigma' > \sigma$,*

$$\epsilon^1_\Lambda(\sigma') \leq \epsilon^1_\Lambda(\sigma).$$

Definition 4.30 (L^1 secrecy-good lattices) *A sequence of lattices $\{\Lambda^{(n)}\}$ is* L^1 secrecy-good *if for all fixed VNR $\gamma_{\Lambda^{(n)}}(\sigma) < 2\pi e$, $\forall c > 0$, $\epsilon^1_{\Lambda^{(n)}}(\sigma) = o\left(\frac{1}{n^c}\right)$, i.e., the L^1 flatness factor vanishes super-polynomially.*

First, we show that sequences of L^1-secrecy good lattices exist under a less stringent volume condition than the one for L^∞-secrecy goodness in Theorem 4.16.

Theorem 4.31 *If $\gamma_\Lambda(\sigma) < 2\pi e$ is fixed, then there exists a sequence $\{\Lambda^{(n)}\}$ of lattices which are L^1-secrecy good.*

The proof of Theorem 4.31 can be found in Appendix C of [J13]. In order to show the existence of a sequence of lattices $\Lambda^{(n)}$ such that $\epsilon^1_{\Lambda^{(n)}}(\sigma) = \mathbb{V}(f_{\sigma,\mathcal{R}(\Lambda^{(n)})}, \mathcal{U}_{\mathcal{R}(\Lambda^{(n)})}) \to 0$, we actually prove a stronger result, namely that $\mathbb{D}(f_{\sigma,\mathcal{R}(\Lambda^{(n)})} || \mathcal{U}_{\mathcal{R}(\Lambda^{(n)})}) \to 0$. We build the required lattices using Construction A, and their existence follows from the existence of linear resolvability codes in [113].

Randomized rounding Following Peikert [182, Section 4.1], we introduce the notion of randomized rounding with respect to a lattice Λ:[4]

Definition 4.32 (Randomized rounding) *Given an input vector $\mathbf{x} \in \mathbb{R}^n$, we define the random variable*

$$\lfloor \mathbf{x} \rceil_{\Lambda,\sigma} \sim D_{\Lambda,\sigma,\mathbf{x}}. \tag{4.36}$$

Note that $\lfloor \mathbf{x} \rceil_{\Lambda,\sigma}$ is a discrete random variable taking values in Λ.

It was shown in [182] that when X^n is i.i.d. Gaussian with variance σ^2, the randomly rounded variable $\lfloor \mathsf{X}^n \rceil_{\Lambda,\sigma_Q}$ is close in variational distance to the discrete Gaussian $D_{\Lambda,\tilde{\sigma}}$, where $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$, provided that the L^∞ flatness factor $\epsilon_\Lambda(\sigma_Q)$ is small:

Proposition 4.33 (Adapted from Theorem 3.1 of [182]) *Let $\mathsf{X}^n \sim \mathcal{N}(0, \sigma^2 I_n)$ and $\boldsymbol{\mu} \in \mathbb{R}^n$, and consider the rounded variable $\mathsf{X}_Q = \lfloor \mathsf{X}^n + \boldsymbol{\mu} \rceil_{\Lambda,\sigma_Q}$. If $\epsilon_\Lambda(\sigma_Q) < 1/2$, then*

$$\mathbb{V}(p_{\mathsf{X}_Q}, D_{\Lambda,\tilde{\sigma},\boldsymbol{\mu}}) \leq 4\epsilon_\Lambda(\sigma_Q),$$

[4] In essence, randomized rounding consists in sampling from a lattice Gaussian distribution centered at $\mathbf{x}$. There exist several algorithms for this task. In particular, it was proven in [96] that Klein's algorithm [128] samples from a distribution very close to $D_{\Lambda,\sigma,\mathbf{x}}$ when σ is sufficiently large. A new algorithm was given in [224] which overcomes the restriction on σ.

where $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$.

We show a partial generalization of this result under an L^1 flatness factor condition, for randomized rounding with uniform dithering, which may be of independent interest.

Lemma 4.34 *Given a Gaussian random vector* $\mathsf{X}^n \sim \mathcal{N}(0, \sigma^2 I_n)$*, a dither* $\mathsf{U} \sim \mathcal{U}_{\mathcal{R}(\Lambda)}$ *uniform over a fundamental region* $\mathcal{R}(\Lambda)$ *and independent of* X^n*, and a constant* $\boldsymbol{\mu} \in \mathbb{R}^n$*, let* $\mathsf{X}_Q = \lfloor \mathsf{X}^n + \mathsf{U} + \boldsymbol{\mu} \rceil_{\Lambda, \sigma_Q}$*. Then*

$$\mathbb{E}_{\mathsf{U}} \left[\mathbb{V} \left(p_{\mathsf{X}_Q|\mathsf{U}}, D_{\Lambda, \tilde{\sigma}, \mathsf{U}+\boldsymbol{\mu}} \right) \right] \leq 2\epsilon_\Lambda^1(\sigma_Q).$$

Another useful property of discrete Gaussian distributions is that a sample $D_{\Lambda,\sigma,\mathbf{c}}$ is distributed almost uniformly modulo a sublattice $\Lambda' \subset \Lambda$ provided that $\epsilon_{\Lambda'}(\sigma)$ is small [96, Corollary 2.8]:

Proposition 4.35 *Let* $\Lambda' \subset \Lambda$*. Then if* $\epsilon_{\Lambda'}(\sigma) < 1$*,*

$$\left\| D_{\Lambda,\sigma,\mathbf{c}} \bmod \Lambda' - \mathcal{U}_{\Lambda/\Lambda'} \right\|_\infty \leq 4\epsilon_{\Lambda'}(\sigma)$$

In the statement above, with slight abuse of notation, $D_{\Lambda,\sigma,\mathbf{c}} \bmod \Lambda'$ denotes the probability density of the random variable $\mathsf{X}_D \bmod \Lambda'$, where $\mathsf{X}_D \sim D_{\Lambda,\sigma,\mathbf{c}}$.

We can partially generalize this statement in an average sense under an L^1-flatness factor condition:

Lemma 4.36 *Let* $\Lambda' \subset \Lambda$*. Then*

$$\mathbb{E}_{\mathsf{U}} \left[\mathbb{V} \left(D_{\Lambda,\sigma,\mathsf{U}} \bmod \Lambda', \mathcal{U}_{\Lambda/\Lambda'} \right) \right] \leq 2\epsilon_{\Lambda'}^1(\sigma)$$

From Lemma 4.34 and Lemma 4.36, we can immediately deduce the following:

Corollary 4.37 *Consider two nested lattices* $\Lambda' \subset \Lambda$*. Given a Gaussian random vector* $\mathsf{X}^n \sim \mathcal{N}(0, \sigma^2 I_n)$*, a dither* $\mathsf{U} \sim \mathcal{U}_{\mathcal{R}(\Lambda)}$ *uniform over a fundamental region* $\mathcal{R}(\Lambda)$ *and independent of* X^n*, and a constant* $\boldsymbol{\mu} \in \mathbb{R}^n$*, let* $\mathsf{X}_Q = \lfloor \mathsf{X}^n + \mathsf{U} + \boldsymbol{\mu} \rceil_{\Lambda, \sigma_Q}$*. Then*

$$\mathbb{E}_{\mathsf{U}} \left[\mathbb{V} \left(p_{\mathsf{X}_Q|\mathsf{U}} \bmod \Lambda', \mathcal{U}_{\Lambda/\Lambda'} \right) \right] \leq 2\epsilon_\Lambda^1(\sigma_Q) + 2\epsilon_{\Lambda'}^1(\tilde{\sigma}),$$

where $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$.

Secret key generation protocol To define our key generation scheme, we use the lattice chain $\Lambda_1 \supseteq \Lambda_2 \supseteq \Lambda_3$, where

- Λ_1 is L^1-secrecy good with respect to σ_Q, and serves as the "source-code" component of Wyner-Ziv coding;
- Λ_2 is AWGN-good [83] with respect to $\tilde{\sigma}_1 = \sqrt{\sigma_1^2 + \sigma_Q^2}$, and serves as the "channel-code" component in Wyner-Ziv coding;
- Λ_3 is L^1-secrecy-good with respect to $\tilde{\sigma}_2 = \sqrt{\sigma_2^2 + \sigma_Q^2}$, and serves as the extractor of randomness.

The parameter σ_Q controls the quantization rate.
In addition, we assume that U is a uniform dither over a fundamental region $\mathcal{R}(\Lambda_1)$, which is known by Alice, Bob and Eve[5].

The procedure of secret key generation is described as follows:

- Alice quantizes X^n to
$$\mathsf{X}_Q = \lfloor \mathsf{X}^n + \mathsf{U} \rceil_{\Lambda_1, \sigma_Q}, \tag{4.37}$$
according to the randomized rounding operation defined in (4.36). That is, $\mathsf{X}_Q \sim D_{\Lambda_1,\sigma_Q,\mathbf{x}+\mathbf{u}}$ if $\mathsf{X}^n = \mathbf{x}$,

[5] If Alice and Bob already share a secret source of randomness, there is no need for secret key generation. Hence, Eve should know U to avoid trivializing the problem.

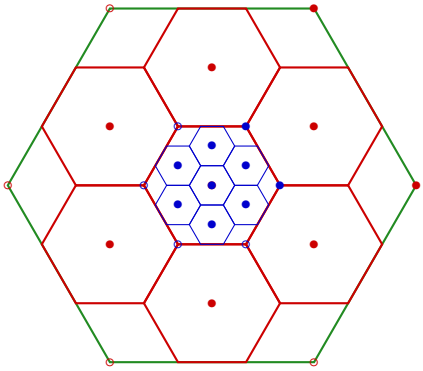


Figure 4.8: A schematic representation of the chain of nested lattices $\Lambda_1 \supset \Lambda_2 \supset \Lambda_3$. The fundamental regions of Λ_1, Λ_2 and Λ_3 are pictured in blue, red and green respectively. The quotient groups Λ_1/Λ_2 and Λ_2/Λ_3 are represented by the blue and red points respectively.

$\mathsf{U} = \mathbf{u}$, or equivalently

$$p_{\mathsf{X}_Q|\mathsf{X}^n,\mathsf{U}}(\mathbf{x}_Q|\mathbf{x},\mathbf{u}) = \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})}. \tag{4.38}$$

She then computes the public message

$$\mathsf{S} = \mathsf{X}_Q^n \bmod \mathcal{V}(\Lambda_2), \tag{4.39}$$

which belongs to a set of coset leaders of Λ_1/Λ_2 in $\mathcal{V}(\Lambda_2)$, and transmits its index to Bob as the public message. Furthermore, Alice computes the key

$$\mathsf{K} = Q_{\Lambda_2}(\mathsf{X}_Q^n) \bmod \mathcal{R}(\Lambda_3), \tag{4.40}$$

which belongs to a set of coset leaders of Λ_2/Λ_3 in $\mathcal{R}(\Lambda_3)$. Note that

$$\mathsf{X}^n = \mathsf{E}_Q^n + \mathsf{S} + \mathsf{K} + \lambda_3 \tag{4.41}$$

for some $\lambda_3 \in \Lambda_3$, where $\mathsf{E}_Q^n = \mathsf{X}^n - \mathsf{X}_Q^n \in \mathcal{V}(\Lambda_1)$ is the quantization error.

- Bob receives S and reconstructs

$$\hat{\mathsf{X}}_Q^n = \mathsf{S} + Q_{\Lambda_2}\left(\rho_{xy}\frac{\sigma_x}{\sigma_y}\mathsf{Y}^n - \mathsf{S}\right).$$

He then computes his version of the key

$$\hat{\mathsf{K}} = Q_{\Lambda_2}(\hat{\mathsf{X}}_Q^n) \bmod \mathcal{R}(\Lambda_3).$$

Let $\bar{\mathsf{X}}_Q = \mathsf{X}_Q \bmod \mathcal{R}(\Lambda_3) \in \Lambda_1/\Lambda_3$, where the quotient Λ_1/Λ_3 is identified with the set of coset representatives $\Lambda_1 \cap \mathcal{R}(\Lambda_3)$. By definition, $\bar{\mathsf{X}}_Q = \mathsf{S} + \mathsf{K}$.

We now state our main result:

Theorem 4.38 *For the Gaussian source model (4.30), there exists a sequence of nested lattices $\Lambda_3^{(n)} \subset \Lambda_2^{(n)} \subset \Lambda_1^{(n)}$ such that for any public rate $R_P > 0$, the previous secret key generation protocol asymptotically achieves the optimal secret key rate $\bar{R}_K(R_P)$ in (4.32). In particular, any secret key rate $R_K < C_s = \frac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2}$ is achievable.*

We summarize here the key steps of the proof. More details can be found in [J13].

Reliability: To show that Bob can decode the key with high probability, we note that a sufficient condition to have $\mathsf{K} = \hat{\mathsf{K}}$ is that $\hat{\mathsf{X}}_Q = \mathsf{X}_Q$, or equivalently, $Q_{\Lambda_2}(\hat{\mathsf{Y}}^n + \mathsf{U} - \mathsf{X}_Q) = 0$, i.e. $\hat{\mathsf{Y}}^n \in \mathsf{X}_Q - \mathsf{U} + \mathcal{V}(\Lambda_2)$. We show that $\mathbb{P}\{\hat{\mathsf{X}}_Q \neq \mathsf{X}_Q\}$ vanishes provided that Λ_1 is L^1 secrecy-good, Λ_2 is AWGN-good and the following volume

conditions hold:

$$\frac{V(\Lambda_1)^{2/n}}{\sigma_Q^2} < 2\pi e, \tag{4.42}$$

$$\frac{V(\Lambda_2)^{2/n}}{\tilde{\sigma}_1^2} > 2\pi e. \tag{4.43}$$

Uniformity: We want to show that the key is asymptotically uniform when $n \to \infty$. Let $\tilde{\sigma}_x^2 = \sigma_x^2 + \sigma_Q^2$. First, we show that the distribution of the key is close to the uniform distribution $\mathcal{U}_\mathcal{K}$ over $\mathcal{K} = \Lambda_2/\Lambda_3$ since

$$\mathbb{V}(p_\mathsf{K}, \mathcal{U}_\mathcal{K}) = \mathbb{V}(p_{\bar{\mathsf{X}}_Q}, \mathcal{U}_{\Lambda_1/\Lambda_3}) \le \mathbb{E}_\mathsf{U}\left[\mathbb{V}\left(p_{\mathsf{X}_Q|\mathsf{U}} \bmod \Lambda_3, \mathcal{U}_{\Lambda_1/\Lambda_3}\right)\right] \overset{(a)}{\le} 2\epsilon^1_{\Lambda_1}(\sigma_Q) + 2\epsilon^1_{\Lambda_3}(\tilde{\sigma}_2) \tag{4.44}$$

where (a) follows from Corollary 4.37 and Lemma 4.29, since $\tilde{\sigma}_2^2 \le \tilde{\sigma}_x^2$. The term (4.44) vanishes as $o\left(\frac{1}{n}\right)$ if both Λ_1 and Λ_3 are L^1-secrecy good and satisfy the volume conditions (4.42) and

$$\frac{V(\Lambda_3)^{2/n}}{\tilde{\sigma}_2^2} < 2\pi e. \tag{4.45}$$

Using [54, Lemma 2.7], we have that if $\mathbb{V}(p_\mathsf{K}, \mathcal{U}_\mathcal{K}) \le \frac{1}{2}$,

$$|\mathbb{H}(p_\mathsf{K}) - \mathbb{H}(\mathcal{U}_\mathcal{K})| \le -\mathbb{V}(p_\mathsf{K}, \mathcal{U}_\mathcal{K}) \log \frac{\mathbb{V}(p_\mathsf{K}, \mathcal{U}_\mathcal{K})}{|\mathcal{K}|}.$$

This vanishes as long as $\mathbb{V}(p_\mathsf{K}, \mathcal{U}_\mathcal{K}) \sim o\left(\frac{1}{n}\right)$, which is indeed the case.

Strong secrecy: Using [53, Lemma 1], we can bound the leakage as follows:

$$\mathbb{I}(\mathsf{K}; \mathsf{S}, \mathsf{Z}^n, \mathsf{U}) = \mathbb{I}(\mathsf{K}; \mathsf{S}, \hat{\mathsf{Z}}^n, \mathsf{U}) \le d_{\text{av}} \log \frac{|\mathcal{K}|}{d_{\text{av}}}, \tag{4.46}$$

where

$$d_{\text{av}} = \sum_{k \in \mathcal{K}} p_\mathsf{K}(k)\, \mathbb{V}(p_{\mathsf{S}\hat{\mathsf{Z}}^n\mathsf{U}|\mathsf{K}=k}, p_{\mathsf{S}\hat{\mathsf{Z}}^n\mathsf{U}}) \tag{4.47}$$

We want to show that d_{av} vanishes as $o\left(\frac{1}{n}\right)$ assuming the conditions (4.42) and (4.45). The proof is rather technical and requires splitting the bound into several terms; a key observation is that on average over the dither and over the eavesdropper's received signal,

$$\int_{\mathbb{R}^n} p_{\hat{\mathsf{Z}}^n}(\mathbf{z}) \mathbb{E}_\mathsf{U}\left[\mathbb{V}\left(p_{\mathsf{X}_Q|\hat{\mathsf{Z}}^n=\mathbf{z},\mathsf{U}} \bmod \Lambda_3, \mathcal{U}_{\Lambda_1/\Lambda_3}\right)\right] d\mathbf{z} \le 2\epsilon^1_{\Lambda_1}(\sigma_Q) + 2\epsilon^1_{\Lambda_3}(\tilde{\sigma}_2),$$

where $\tilde{\sigma}_2^2 = \sigma_2^2 + \sigma_Q^2$. This follows by noticing that $p_{\mathsf{X}_Q|\hat{\mathsf{Z}}^n,\mathsf{U}}(\mathbf{x}_Q|\mathbf{z}, \mathbf{u})$ is the distribution of $\lfloor \mathsf{W}_2^n + \mathbf{z} + \mathbf{u} \rceil_{\Lambda_1, \sigma_Q}$ and by applying Corollary 4.37.

Achievable strong secrecy rate and optimal trade-off: In the previous sections we have imposed the conditions (4.42), (4.43) and (4.45) on the volumes of Λ_1, Λ_2 and Λ_3 respectively, i.e.

$$\frac{V(\Lambda_1)^{2/n}}{\sigma_Q^2} < 2\pi e, \quad \frac{V(\Lambda_2)^{2/n}}{\tilde{\sigma}_1^2} > 2\pi e, \quad \frac{V(\Lambda_3)^{2/n}}{\tilde{\sigma}_2^2} < 2\pi e.$$

Therefore, the achievable secret key rate is upper bounded by

$$R_K = \frac{1}{n} \log \frac{V(\Lambda_3)}{V(\Lambda_2)} < \frac{1}{2} \log \frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2} = \frac{1}{2} \log \frac{\sigma_2^2 + \sigma_Q^2}{\sigma_1^2 + \sigma_Q^2} \tag{4.48}$$

As $\sigma_Q \to 0$,

$$R_K \to \frac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2},$$

which is the optimal secret key rate.

The public communication rate is lower bounded by

$$R_P = \frac{1}{n}\log\frac{V(\Lambda_2)}{V(\Lambda_1)} > \frac{1}{2}\log\frac{\sigma_1^2+\sigma_Q^2}{\sigma_Q^2}.$$

Equivalently, we have $\sigma_Q^2 > \frac{\sigma_1^2}{e^{2R_P}-1}$. Replacing this expression in the bound (4.48) for R_K, and observing that (4.48) is a decreasing function of σ_Q^2, we find

$$R_K < \frac{1}{2}\log\left(e^{-2R_P} + \frac{\sigma_2^2}{\sigma_1^2}(1-e^{-2R_P})\right).$$

which corresponds to the optimal public rate / secret key rate trade-off (4.32).

Remark 4.39 The optimal scaling of the lattice Λ_3 requires the noise variance σ_2 to be known by Alice; if only a lower bound for σ_2 is available, positive secret key rates can still be attained.

Some open problems and perspectives related to this work will be presented in Chapter 6.

4.4 Reconciliation for secret key generation protocols based on Learning With Errors

As seen in the previous sections, in our work on wiretap coding and secret key generation using lattices we have borrowed some technical tools from lattice-based cryptography, such as the flatness factor, which is an equivalent formulation of the smoothing parameter. As part of Charbel Saliba's PhD thesis [197], we investigated reconciliation and error-correction techniques for secret-key generation cryptographic protocols based on lattices.

Currently, lattice-based cryptographic primitives are among the most promising candidates for the new generation of post-quantum secure cryptographic protocols. In particular, many of the key-establishment protocols and digital signature methods submitted to the post-quantum cryptography challenge launched by the U.S. National Institute of Standards and Technology (NIST) since 2016 are based on lattices.
One of the most widely used lattice-based cryptographic primitives is the Learning With Errors (LWE) problem introduced by Regev [193]. The *decision version* of the LWE problem is stated informally as follows:

Definition 4.40 ($\mathsf{LWE}_{q,\chi}$) *Let $q = poly(n)$ be an integer and χ a Gaussian-like distribution over $\mathbb{Z}$. Let $m = poly(n)$, referred sometimes as the number of samples. For a secret vector $\mathbf{s} \in \mathbb{Z}_q^n$ and an $m \times n$ dimensional matrix $\mathbf{A}$ sampled uniformly from $\mathbb{Z}_q^{m\times n}$, consider an error term $\mathbf{e}$ drawn from the χ^m distribution and denote $\mathbf{b} = \mathbf{A}\mathbf{s}+\mathbf{e}$. The problem asks to distinguish between the two samples $(\mathbf{A},\mathbf{b})$ and $(\mathbf{A}',\mathbf{b}')$, where $\mathbf{A}'$ and $\mathbf{b}'$ are uniform samples from $\mathbb{Z}_q^{m\times n}$ and $\mathbb{Z}_q^n$ respectively.*

The LWE problem admits a worst-case quantum reduction from the shortest independent vector problem (SIVP) to within a polynomial approximation factor for generic lattices. Informally speaking, Regev's result can be summarized as follows: given an n-dimensional lattice L and a rounded Gaussian error distribution χ with parameter $\alpha q \geq 2\sqrt{n}$ for $\alpha = \alpha(n) \in (0,1)$, there exists a polynomial-time quantum reduction from $\mathsf{SIVP}_{L,\gamma}$ to $\mathsf{LWE}_{q,\chi}$, where $\gamma \leq \frac{2\sqrt{2}n}{\alpha}\eta_\epsilon(L)$ depends on the *smoothing parameter* $\eta_\epsilon(L)$ (see Definition 4.9), and ϵ is a negligible function of n.
The LWE problem can be used to build a variety of cryptographic algorithms and provides guarantees in terms of IND-CPA (indistinguishability under chosen-plaintext attack) [193] and IND-CCA (indistinguishability under chosen-ciphertext attack) security [184]. Other applications of LWE include fully homomorphic encryption [33].

Although the theoretical results in [193] were proven under the assumption that χ is a *rounded Gaussian* distribution, sampling rounded Gaussians or discrete Gaussians requires a significant algorithmic effort, and many cryptosystems employ other distributions that are easier to implement, such as the centered binomial distribution [9]. This modification does not significantly degrade the secrecy performance as long as the two distributions are close in Rényi divergence.

The public key size for LWE-based encryption is of the order $O(n^2)$, which renders this method impractical compared to RSA. Two structured variants of LWE, the decision *Ring Learning With Errors* (RLWE) and *Module Learning With Errors* (MLWE) were proposed in [161] and [135] respectively to allow for more compact representations and shorter keys of size $O(n)$. Given the ring of integers R of an algebraic number field and its quotient ring $R_q = R/qR$, the M − LWE problem consists in distinguishing uniform samples $(\vec{\mathbf{a}}_i, \mathbf{b}_i) \leftarrow R_q^d \times R_q$ from samples $(\vec{\mathbf{a}}_i, \mathbf{b}_i) \leftarrow R_q^d \times R_q$ where $\vec{\mathbf{a}}_i \leftarrow R_q^d$ is uniform, $\mathbf{b}_i = \langle \vec{\mathbf{a}}_i, \vec{\mathbf{s}} \rangle + \mathbf{e}_i$ with $\mathbf{e}_i$ generated from a "Gaussian-like" distribution Ψ on R_q, and $\vec{\mathbf{s}} \leftarrow \Psi^d$. Solving MLWE was shown to be at least as hard as solving approximate SIVP on module lattices, i.e., lattices corresponding to modules over the ring R.

We present a common setting for LWE-based key encapsulation mechanisms (KEM) in Table 4.1. More precisely, this is a common setting for an LWE-based Public Key Encryption (PKE) scheme. The Fujisaki-Okamoto transform is then used to obtain an IND-CCA secure KEM from an IND-CPA secure PKE [120]. We consider two

Parameters: m, n, q and error distribution χ on $\mathbb{Z}_q$		
Alice (Server)		**Bob (Client)**
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$		$\mathbf{S}' \xleftarrow{\$} \chi^{\bar{m} \times m}, \mathbf{E}' \xleftarrow{\$} \chi^{\bar{m} \times n}$
$\mathbf{S} \xleftarrow{\$} \chi^{n \times \bar{n}}, \mathbf{E} \xleftarrow{\$} \chi^{m \times \bar{n}}$		$\mathbf{E}'' \xleftarrow{\$} \chi^{\bar{m} \times \bar{n}}$
$\mathbf{B} = \mathbf{AS} + \mathbf{E} \in \mathbb{Z}_q^{m \times \bar{n}}$	$\xrightarrow{(\mathbf{A}, \mathbf{B})}$	$\mathbf{U} = \mathbf{S}'\mathbf{A} + \mathbf{E}'$
$\mathbf{V}' = \mathbf{US}$	$\xleftarrow{\mathbf{U}}$	$\mathbf{V} = \mathbf{S}'\mathbf{B} + \mathbf{E}''$

Table 4.1: Common setting for LWE-based KEMs.

terminals, Alice and Bob (Server and Client), whose aim is to generate the same private key. We note that while in physical layer security the randomness inherent in the physical channel can be used to generate secret keys (see Section 4.3), in cryptographic protocols the pseudo-random noise is generated at the legitimate terminals.

Given integer parameters m, n and the modulus q, Alice chooses a uniformly random matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ as well as two random 'small' error matrices $\mathbf{S}, \mathbf{E}$ such that each component is generated from the Gaussian-like error distribution χ, and sends the LWE sample pair $(\mathbf{A}, \mathbf{B})$ as a public key to Bob. Bob generates $\mathbf{S}', \mathbf{E}'$ and $\mathbf{E}''$ and computes the LWE samples

$$\mathbf{U} = \mathbf{S}'\mathbf{A} + \mathbf{E}' \bmod q, \quad \mathbf{V} = \mathbf{S}'\mathbf{B} + \mathbf{E}'' \bmod q.$$

The term $\mathbf{U}$ is sent back to Alice who uses it to compute $\mathbf{V}' = \mathbf{US}$ with her secret key $\mathbf{S}$. Note that

$$\mathbf{V} - \mathbf{V}' = \mathbf{S}'\mathbf{E} + \mathbf{E}'' - \mathbf{E}'\mathbf{S} \bmod q. \tag{4.49}$$

When the distribution χ is chosen appropriately, the term $(\mathbf{S}'\mathbf{E} + \mathbf{E}'' - \mathbf{E}'\mathbf{S})$ is small with high probability, and therefore the values $\mathbf{V}$ and $\mathbf{V}'$ are close to $\mathbf{S}'\mathbf{AS}$ and can be used to generate a common key through an additional exchange of information, as explained in the next sections. For this kind of scheme, IND-CPA security follows from the hardness of decision-LWE [151]. Informally speaking, one should prove that all the information exchanged on the public channel is indistinguishable from uniformly random from the point of view of an adversary. In fact, as seen from Table 4.1, since $\mathbf{B}$ was constructed as an LWE sample, the public key $(\mathbf{A}, \mathbf{B})$ is computationally indistinguishable from uniform $(\mathbf{A}, \mathbf{B}^*)$ assuming the hardness of decision-LWE (see Definition 4.40). For the same reason, the matrices $\mathbf{U}$ and $\mathbf{V}$ are also indistinguishable from uniformly random. We note that if additional reconciliation messages are exchanged, these should also be provably pseudo-random.

The structure of the protocol brings to mind the Diffie-Hellman protocol [66] where the public parameter $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ corresponds to the generator of the group and the noise-free product is analogous to exponentiation.

The presence of noise due to error terms leads to so-called "Noisy Diffie-Hellman" protocols [3, 151]. If the random noise terms are large, an error will occur during the recovery of the private key, affecting the reliability of the scheme. As a result, it is necessary to choose the error distribution in such a way that the failure probability is guaranteed to be exponentially small. This is important not only for reliability, but also for the security parameters when an IND-CPA secure PKE is transformed into an IND-CCA secure KEM. For instance, an insufficiently small error probability can cause a leakage of information due to decryption failure attacks [63] where a failure boosting technique is used to increase the failure rate. To keep the error probability small, one can use either *error correction* or *reconciliation* approaches. These techniques make it possible to agree on an exact shared private key by providing Bob with some additional information.

Encryption-based approach. This first approach is used in many LWE-based encryption schemes, such as [151, 172, 13]. Table 4.2 illustrates this approach. As shown, Bob unilaterally generates a uniform message $\mathbf{m} \in \{0,1\}^{\ell}$

Parameters: $m, n, \bar{m}, \bar{n}, q$ and error distribution χ on $\mathbb{Z}_q$		
Alice (Server)		**Bob (Client)**
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$		$\mathbf{S}' \xleftarrow{\$} \chi^{\bar{m}\times m}, \mathbf{E}' \xleftarrow{\$} \chi^{\bar{m}\times n}$
$\mathbf{S} \xleftarrow{\$} \chi^{n\times\bar{n}}, \mathbf{E} \xleftarrow{\$} \chi^{m\times\bar{n}}$		$\mathbf{E}'' \xleftarrow{\$} \chi^{\bar{m}\times\bar{n}}$
$\mathbf{B} = \mathbf{AS} + \mathbf{E} \in \mathbb{Z}_q^{m\times\bar{n}}$	$\xrightarrow{(\mathbf{A},\mathbf{B})}$	$\mathbf{m} \xleftarrow{\$} \{0,1\}^{\ell}$
		$\mathbf{U} = \mathbf{S}'\mathbf{A} + \mathbf{E}'$
		$\mathbf{V} = \mathbf{S}'\mathbf{B} + \mathbf{E}''$
$\mathbf{V}' = \mathbf{US}$	$\xleftarrow{(\mathbf{U},\mathbf{C})}$	$\mathbf{C} = \mathbf{V} + \text{ENCODE}(\mathbf{m})$
$\mathbf{m}' = \text{DECODE}(\mathbf{C} - \mathbf{V}')$		

Table 4.2: Encryption-based KEM.

and encodes it using some well defined injective function ENCODE that maps $\{0,1\}^{\ell}$ into $\mathbb{Z}_q^{\bar{m}\times\bar{n}}$. He then sends the cyphertext $\mathbf{C} = \mathbf{V} + \text{ENCODE}(\mathbf{m})$ to Alice so that she can recover $\mathbf{m}'$ by applying the decoding function $\text{DECODE}(\mathbf{C} - \mathbf{US})$.

Reconciliation-based approach. The reconciliation method [67] allows two parties who have obtained noisy observations to come to an exact agreement about the value of the key, and consists in sending an auxiliary message $\mathbf{R}$ from Bob to Alice in order to help her recover the private key from her noisy observation $\mathbf{V}'$. As shown

Parameters: $m, n, \bar{m}, \bar{n}, q$ and error distribution χ on $\mathbb{Z}_q$		
Alice (Server)		**Bob (Client)**
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$		$\mathbf{S}' \xleftarrow{\$} \chi^{\bar{m}\times m}, \mathbf{E}' \xleftarrow{\$} \chi^{\bar{m}\times n}$
$\mathbf{S} \xleftarrow{\$} \chi^{n\times\bar{n}}, \mathbf{E} \xleftarrow{\$} \chi^{m\times\bar{n}}$		$\mathbf{E}'' \xleftarrow{\$} \chi^{\bar{m}\times\bar{n}}$
$\mathbf{B} = \mathbf{AS} + \mathbf{E} \in \mathbb{Z}_q^{m\times\bar{n}}$	$\xrightarrow{(\mathbf{A},\mathbf{B})}$	
		$\mathbf{U} = \mathbf{S}'\mathbf{A} + \mathbf{E}'$
		$\mathbf{V} = \mathbf{S}'\mathbf{B} + \mathbf{E}''$
$\mathbf{V}' = \mathbf{US}$	$\xleftarrow{(\mathbf{U},\mathbf{R})}$	$\mathbf{R} = \text{HELPREC}(\mathbf{V})$
$\boldsymbol{\mu}' = \text{REC}(\mathbf{V}', \mathbf{R})$		$\boldsymbol{\mu} = \text{REC}(\mathbf{V}, \mathbf{R})$

Table 4.3: Reconciliation-based KEM.

in Table 4.3, Bob produces a reconciliation message $\mathbf{R}$ using the function HELPREC and uses it to generate the private key $\boldsymbol{\mu}$ by applying the function REC that computes a private key $\boldsymbol{\mu}$ given a noisy observation and the reconciliation information. He then sends the message $\mathbf{R}$ to Alice. This information aims to correct the bias that exists between $\mathbf{V}$ and $\mathbf{V}'$. Alice can now use the function REC to compute the private key $\boldsymbol{\mu}'$.
One example of this technique is the proof-of-concept in [185], which uses the one dimensional lattice $\mathbb{Z}$ in order to agree on one bit of private key, with the priority of keeping a low bandwidth. Other works proposed reconciliation with higher-dimensional lattices; for instance the NIST first-round protocol New Hope based on RLWE [9] considers the four dimensional lattice $\tilde{D}_4$.

We observe that the lattice-based reconciliation steps mentioned previously are essentially of the same form[6] as in our information-theoretic key generation protocol described in Section 4.3[7]. Namely, given a lattice partition chain $\Lambda_3 \subseteq \Lambda_2 \subseteq \Lambda_1$, Bob computes the reconciliation message $\mathbf{R}$ as $\mathbf{R} = Q_{\Lambda_1}(\mathbf{V}) \bmod \Lambda_2$ and transmits it to Alice. Furthermore, Bob computes the private key as $\boldsymbol{\mu} = Q_{\Lambda_2}(\mathbf{V}, \mathbf{R}) \bmod \Lambda_3$. These equations correspond to (4.39) and (4.40).

Main contributions. The aim of Charbel Saliba's thesis [197] was to improve some of the key encapsulation mechanisms proposed for the NIST challenge in terms of security, reliability and bandwidth by introducing new error correction or reconciliation techniques.
The main challenges in this setting are due to the fact that the target error probability is far beyond the range of numerical simulations, and moreover the components of the error distribution are not independent. Some works which use error correcting codes to improve the performance [159, 146] compute the error probability under an independence assumption which does not hold in practice. However, this has been shown to lead to underestimating the error probability by a very large exponential factor [75]. In this thesis, we choose instead to derive rigorous error probability bounds following the example of [9]. Our main inspiration comes from the nested lattice chain method of [C12], but we focus on reconciliation/error correction using a sublattice Λ_2 of small dimensional in order to keep a low complexity.

Error-correction for FrodoKEM [C27]. We first consider the alternative NIST candidate FrodoKEM [172], which is an LWE-based KEM, and propose a modification at the level of the encoding function. We define a new encoder which maps the private key block-wise into the 8-dimensional Gosset lattice E_8. We choose E_8 since it gives the densest 8-dimensional packing, resulting in a more efficient decoding, and admits a low-complexity quantization. We propose three sets of parameters for our modified implementation. Thanks to the improved error correction, the first implementation allows to reduce the bandwidth by 7% by halving the modulus q; the second outperforms FrodoKEM in terms of plausible security by 10 to 13 bits by increasing the error variance, and the third one aims to increase the key size by approximately 50%. In all cases, our scheme can ensure a smaller decryption failure probability compared to the original FrodoKEM.

Reconciliation for KyberKEM [C24]. Next, we focus on the KyberKEM protocol[8] [13]. We propose a modification of KyberKEM featuring a reconciliation mechanism based on E_8. The main technical contribution of this work is a rigorous bound for the decryption failure probability using a polynomial splitting and a union bound over Voronoi-relevant vectors in the Gosset lattice. The final computation of the bound requires extensive numerical simulations. More details can be found in [197, Section 4.4.3].
Similarly to KyberKEM, our scheme generates 256 bits of key and requires 5 or 6 bits of reconciliation per dimension. We show that it can outperform KyberKEM in terms of the modulus q with comparable error probability and similar requirements in terms of bandwidth. For instance, our construction guarantees a smaller error probability than KyberKEM-768's, i.e. $P_e \leq 2^{-174} < 2^{-164}$, with a smaller modulus $q = 2^{11} < 3329$, using 5 bits of reconciliation per dimension. For this choice of q, our scheme achieves 176 bits of post-quantum security compared to 164 bits[9]. A similar improvement can be obtained for KyberKEM-512. Note that unlike KyberKEM, where the modulus q is prime and the Number Theoretic Transform is used for fast polynomial multiplication, we choose q to be a power of two. In this case, efficient polynomial multiplication is still possible using Karatsuba / Toom-Cook algorithms. Moreover, unlike [185, 9], we don't need dithering to obtain a uniform key thanks to the fact that q is even.

[6] For instance, in [9] the functions HELPREC and REC can be written as the above form by considering the product lattices $\Lambda_1 = (q\tilde{D}_4/2^p)^{256}$, $\Lambda_2 = (q\tilde{D}_4)^{256}$ and $\Lambda_3 = q\mathbb{Z}^{1024}$.

[7] Note that the roles of Alice and Bob are exchanged.

[8] We note that after our work was completed, KyberKEM was officially selected by NIST for standardization.

[9] After our work was published, the security estimates for KyberKEM (as well as for other NIST candidates such as SABER) have been questioned due to improved lattice attacks in [104, 164]. However, there is no full consensus in the community about the validity of these attacks which are based on some heuristics [69]. The result quoted above does not take into account these recent developments.

Open problems and perspectives

As part of Charbel Saliba's thesis [197] we also investigated the use of higher-dimensional lattices, such as Barnes-Wall lattices, for reconciliation. However, our results are inconclusive and we were unable to obtain an error probability bound in the range required for post-quantum cryptography applications. In part, this is due to the fact that it is difficult to obtain rigorous and tight bounds for the error probability for high-dimensional lattices. A rather counter-intuitive conclusion is that the use of higher-dimensional lattices does not necessarily bring a gain in terms of minimum distance due to the scaling constraints imposed by the modulo-q integer arithmetic of LWE protocols. After our work was published, [167] revisited the error correction framework by focusing on the limits of this type of protocol in terms of rate (i.e. the ratio of plaintext size to ciphertext size), and showed that compression of the ciphertext is required in order to approach rate 1 asymptotically. One important application for error correction is fully homomorphic encryption based on LWE or RLWE [32, 33], where noise amplification is the main limiting factor for the number of homomorphic operations.

5 | COORDINATION OF AUTONOMOUS AGENTS - INFORMATION THEORETIC BOUNDS AND CODING SCHEMES

The fifth generation of wireless networks (5G) introduced machine to machine communication and the Internet of Things: a unified network of connected objects including embedded sensors, medical devices, smart meters, and autonomous vehicles. In addition to faster communications, it also supports the next wave of technological innovation, from connected cars to factory automation, smart cities, robot-assisted surgery, virtual reality and edge computing.

Future smart decentralized networks must be able to cooperate, to take decisions in a distributed fashion and to reconfigure dynamically by reacting to changes in their environment. In order to achieve such behavior, one must develop efficient techniques to coordinate the actions of different nodes. This problem is of an interdisciplinary nature and brings together various research fields, such as network information theory, distributed control, game theory, and parallel processing. Among the challenges that must be addressed to develop distributed systems, a key aspect is to reach a better understanding of the interplay between coordination and communication.

State of the art A promising framework in network information theory, which analyzes new purposes for communication beyond the traditional transfer of information under a reliability constraint, has been proposed by Cuff, Permuter and Cover [56], related to earlier work on "Shannon's reverse coding theorem" [23] and the compression of probability distributions [134]. This framework also relates to the game-theoretic perspective on coordination [102] which has applications, for example, to power control [137]. In this chapter, we mostly adopt the viewpoint of [56]. The key idea is to model the actions performed by agents in the network by discrete random variables, and to measure the level of coordination by the distance of their joint probability distribution from a target distribution; information enters the network in the form of actions which are assigned to certain nodes by external constraints. The goal is to characterize the minimal communication rate among the nodes that is asymptotically necessary to establish coordination. In particular, it can be shown that the naïve approach of sending explicit messages describing the actions is inefficient in general. In contrast, by exploiting *common randomness* available at the nodes, the communication rates required to coordinate correlated actions can be significantly reduced. In some settings, common randomness (provided for example by a GPS timestamp) may be less expensive to get than communication.

In this chapter, we mostly focus on the simplified scenario of a two-node network. As we will see, even this simple case gives rise to non-trivial problems in information theory.

Consider the setting in Figure 5.1 where two nodes are connected by a one-directional error-free link of rate R. We will assume that a common source of uniform randomness C, of rate R_0, is available at the nodes.

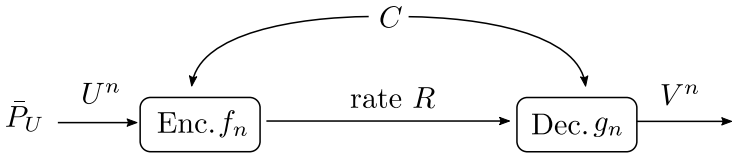


Figure 5.1: Coordination of actions for a two-node network with an error-free link.

At time $i = 1, \dots, n$, the nodes perform the actions U_i and V_i respectively. The source sequence U^n is assigned by nature[1] according to the fixed distribution $\bar{P}_{U^n}$. The encoder generates a message M as a stochastic function of U^n and the common randomness C. The message is sent through the error-free link and the sequence of actions V^n is generated at the second node as a function of M and C.

In [56], two different notions of coordination are proposed, empirical and strong coordination, depending on the choice of metric on the space of joint distributions.

Empirical coordination measures an average behavior over time, requiring that the two sequences of actions be statistically indistinguishable on average, i.e. have the same histogram or type:

Definition 5.1 (Empirical coordination) *A joint distribution $\bar{P}_{UV}$ and a rate of communication R are* achievable for empirical coordination *if there exists a sequence (f_n, g_n) of encoders-decoders such that $\forall \epsilon > 0$, $\exists \bar{n}$ such that $\forall n \geq \bar{n}$,*

$$\mathbb{P}\left\{\mathbb{V}\left(T_{U^nV^n}, \bar{P}_{UV}\right) > \epsilon\right\} < \epsilon,$$

where

$$T_{U^nV^n}(u, v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left\{(u_i, v_i) = (u, v)\right\}$$

is the joint histogram *of the actions induced by the code (f_n, g_n).*

Namely, the distance between the joint histogram and the target distribution converges to zero in probability as $n \to \infty$.

Definition 5.2 (Empirical coordination region) *The* empirical coordination region $\mathcal{R}_e$ *is the closure of the set of achievable pairs $(\bar{P}_{UV}, R)$*[2].

In contrast, *strong coordination* requires the distribution of the sequence of joint actions to converge to the target in total variational distance when the sequence length tends to infinity.

Definition 5.3 (Strong coordination) *A sequence $(\bar{P}_{UV}, R, R_0)$ is achievable for strong coordination if there exists a sequence (f_n, g_n) of encoders-decoders with rate of common randomness R_0, such that*

$$\lim_{n\to\infty} \mathbb{V}\left(P_{U^nV^n}, \bar{P}_{UV}^{\otimes n}\right) = 0,$$

where $P_{U^nV^n}$ is the joint distribution induced by the code (f_n, g_n).

Definition 5.4 (Strong coordination region) *The* strong coordination region $\mathcal{R}$ *is the closure of the set of achievable $(\bar{P}_{UV}, R, R_0)$.*

Empirical coordination turns out to have a close connection to rate-distortion theory. For instance, one can show that any good code for lossy source coding is good for empirical coordination and vice-versa. On the other side, strong coordination is related to a different mechanism, namely channel resolvability [108] (see Chapter 4).

Remark 5.5 We note that strong coordination is to be preferred form a security standpoint. For example, it might be useful to make the sequence of actions appear unpredictable to an adversary who observes the actions of the nodes and tries to anticipate and exploit patterns. Suppose that the adversary performs a statistical test to decide if the distribution P induced by the code is indistinguishable in total variational distance from the i.i.d. distribution $\bar{P}$ (hypothesis H_0). We denote α the probability of Type I error (rejecting H_0 when true) and β the probability of Type II error (accepting H_0 when false). In [147] it is proved that it is possible for the adversary to design blind tests (ignoring his channel observations) that achieve any pair (α, β) such that $\alpha + \beta = 1$, and that the adversary's

[1] Without the assumption that some actions are assigned by nature, the problem becomes trivial. If the two nodes can choose their actions and common randomness is available, no communication is required between the nodes and, if the nodes can agree ahead of time on how they will behave in the presence of common randomness, any conditional distribution $\bar{P}_{V^n|U^n}$ can be generated [57].

[2] This definition allows to avoid boundary complications, see [56].

optimal test satisfies $\alpha + \beta \geq 1 - \mathbb{V}(P, \bar{P})$. By minimizing the variational distance between the two distributions, we ensure that the adversary's best statistical test is not much better than that of a blind test.

Of course, in the setting of Figure 5.1, a trivial solution to the coordination problem would be to have the first node communicate its randomized actions to the second node using the error-free link, which would require a rate of at least $H(U)$ bits per action. Then, the second node would just have to simulate a discrete memoryless channel $P_{V|U}$ using local randomness. However, it turns out that this strategy is an excessive use of communication. The empirical and strong coordination regions were characterized in [56]:

$$\mathcal{R}_{\text{Cuff},e} = \left\{ (\bar{P}_{UV}, R) \mid \bar{P}_{UV} = \bar{P}_V \bar{P}_{V|U}, R \geq I(U;V) \right\}, \tag{5.1}$$

$$\mathcal{R}_{\text{Cuff}} = \left\{ (\bar{P}_{UV}, R, R_0) \;\middle|\; \begin{array}{l} \bar{P}_{UV} = \bar{P}_V \bar{P}_{V|U} \\ \exists W \text{ taking values in } \mathcal{W},\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{V}| + 1,\ \bar{P}_{UWV} = \bar{P}_U \bar{P}_{W|U} \bar{P}_{V|W} \\ R \geq I(U;W),\ R + R_0 \geq I(UV;W) \end{array} \right\} \tag{5.2}$$

This result characterizes the trade-off between the rate R_0 of available common randomness and the required description rate R for simulating a discrete memoryless channel for a fixed input distribution.

Coding schemes for coordination An important question is how to design practical codes to achieve the coordination regions above. One of the hurdles faced for code design is that the metric to optimize is not a probability of error but a variational distance between distributions. Polar codes [11] prove themselves to be particularly well suited to translate information theoretic properties such as coordination.

A first construction using polar codes for empirical coordination in two-node and three-node cascade networks was proposed in [27], exploiting the connection with rate-distortion theory.

In a collaboration with M. Bloch and J. Kliewer [C11], presented in the next section, we showed that polar codes can achieve strong coordination in a two-node system in the special case when the action imposed by nature is binary and uniform.

5.1 Polar codes for strong coordination of uniform actions over error-free links

Polar codes First, we briefly review the concepts and notation related to polar codes [11] that will be used in the sequel. The key element in the polar coding construction is the decomposition of $n = 2^m$ independent copies of a given binary-input discrete-memoryless channel $W : \{0,1\} \to \mathcal{V}$ into n *bit-channels* which are essentially either error-free or pure noise channels. Specifically, consider the transformation $G_n = G_2^{\otimes n} P_n$ where

$$G_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

is the kernel matrix of polar codes, $\otimes$ is the Kronecker product and P_n is the bit-reversal permutation matrix. The general polarization transform for length $n = 2^m$ proceeds in two steps:

1) *Channel combining:* define a vector channel

$$W_n(\mathbf{v}|\mathbf{u}) = W^n(\mathbf{v}|\mathbf{x}), \quad \mathbf{x} = G_n \mathbf{u}.$$

2) *Channel splitting:* define the *bit channels* $W_n^{(i)} : \{0,1\} \to \mathcal{V}^n \times \{0,1\}^{i-1}$, $i = 1, \ldots, n$ with transition probabilities

$$W_n^{(i)}(\mathbf{v}, (u_1, \ldots, u_{i-1})|u_i) = \sum_{u_{i+1}, \ldots, u_n} \frac{1}{2^{n-1}} W_n(\mathbf{v}|\mathbf{u}).$$

Polar codes for channel resolvability. We now design a simple polar coding scheme for *channel resolvability* (see Definition 4.1) for uniform binary sources and symmetric channels. Instead of considering Arikan's standard notion of good and bad bit channels [11], which is based on the Bhattacharyya parameter, we adopt Mahdavifar and Vardy's notion of "poor bit-channels" which was introduced in the context of wiretap polar codes [162]. More precisely, given $0 < \beta < 1/2$, we define

$$\mathcal{P}_n = \left\{ i \in [\![1, n]\!] : C(W_n^{(i)}) < 2^{-n^\beta} \right\},$$
$$\mathcal{G}_n = \left\{ i \in [\![1, n]\!] : C(W_n^{(i)}) \geq 2^{-n^\beta} \right\}.$$

Let $r = |\mathcal{G}_n|$. Then, it was shown in [162, Proposition 20] that

$$\lim_{n\to\infty} \frac{r}{n} = C(W), \tag{5.3}$$

where the limit is approached from above. Given two vectors $\mathbf{x}^r \in \{0,1\}^r$ and $\mathbf{s}^{n-r} \in \{0,1\}^{n-r}$, we let $(\mathbf{x}^r || \mathbf{s}^{n-r})$ denote the vector $\mathbf{u}^n \in \{0,1\}^n$ such that $\mathbf{u}_{|\mathcal{G}_n} = \mathbf{x}^r$ and $\mathbf{u}_{|\mathcal{P}_n} = \mathbf{s}^{n-r}$.

Our strategy to simulate the i.i.d. output process distributed according to $\bar{P}_V$ is to send random uniform bits on the "good bits" $\mathcal{G}_n$, and fixed bits (for instance, zeros) on the poor bits $\mathcal{P}_n$. We denote by $\mathcal{C}_n$ the corresponding coset code, of rate r/n. The codewords in $\mathcal{C}_n$ will be of the form $\mathbf{x} = G_n(\mathbf{x}^r || \mathbf{0}^{n-r})^T$.

Intuitively, the uniform bits will be preserved by the noiseless bit-channels, while the pure noise bit-channels will produce almost-uniform bits for any input.

Proposition 5.6 *If the channel $W : \mathcal{U} = \{0,1\} \to \mathcal{V}$ is symmetric and $\bar{P}_U \sim \mathcal{B}(\frac{1}{2})$, then $\{\mathcal{C}_n\}_{n\geq 1}$ is a sequence of resolvability codes of resolution rate $C(W)$ for $(W, \bar{P}_U)$.*

Proof. The condition for the resolution rate is verified due to (5.3). Following [162], we consider the composite channel $Q_n : \{0,1\}^{n-r} \to \mathcal{V}^n$, which includes the polar code and the random bits sent on the good bits $\mathcal{G}_n$ as follows:

$$Q_n(\mathbf{v}^n | \mathbf{s}^{n-r}) = \frac{1}{2^r} \sum_{\mathbf{x}^r \in \{0,1\}^r} W^n\left(\mathbf{v}^n \middle| G_n(\mathbf{x}^r || \mathbf{s}^{n-r})^T\right).$$

It was shown in [162, Proposition 13] that Q_n is symmetric and that

$$C(Q_n) \leq \sum_{i \in \mathcal{P}_n} C(W_n^{(i)}) \leq (n-r) 2^{-n^\beta}. \tag{5.4}$$

We now show that this last inequality implies that $\{\mathcal{C}_n\}_{n\geq 1}$ form a sequence of resolvability codes.

The output distribution P_{V^n} induced by the code $\mathcal{C}_n$ coincides with the output distribution $Q_n(\cdot || \mathbf{0}^{n-r})$ of the constant input $\mathbf{s}^{n-r} = \mathbf{0}^{n-r}$. Moreover, since Q_n is symmetric and G_n is full-rank, the output of the channel Q_n for a uniformly distributed input on $\{0,1\}^{n-r}$ has the desired distribution $\bar{P}_V^{\otimes n}$, where $\bar{P}_V \sim \mathcal{B}\left(\frac{1}{2}\right)$. We recall the following property of symmetric channels [91]:

Lemma 5.7 *If $W : \mathcal{X} \to \mathcal{Y}$ is a memoryless symmetric channel and if $\bar{P}_Y$ is the output distribution corresponding to the uniform input distribution $\bar{P}_X$ on $\mathcal{X}$, then*

$$\forall x \in \mathcal{X}, \quad C(W) = \mathbb{D}(W_{\mathsf{Y}|\mathsf{X}=x} || \bar{P}_Y).$$

Applying Lemma 5.7 to the channel Q_n, we find that

$$\mathbb{D}(P_{V^n} || \bar{P}_V^{\otimes n}) = \mathbb{D}(Q_n(\cdot | \mathbf{0}^{n-r}) || \bar{P}_V^{\otimes n}) = C(Q_n),$$

so that $\lim_{n\to\infty} \mathbb{D}\left(P_{V^n} || \bar{P}_V^{\otimes n}\right) = 0$. Pinsker's inequality then ensures that $\lim_{n\to\infty} \mathbb{V}(P_{V^n}, \bar{P}_V^{\otimes n}) = 0$. □

Remark 5.8 The choice of frozen bits set at $\mathbf{0}^{n-r}$ is arbitrary. The choice of a different coset code characterized

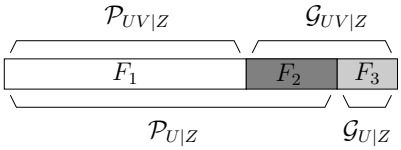


Figure 5.2: Illustration of partition sets F_1, F_2, F_3 (after reordering of indices).

by $\mathbf{u}_F$ in place of $\mathbf{0}^{n-r}$ does not alter the reasoning. In particular, the symmetry of the channel Q_n and (5.7) still hold.

Coding scheme for strong coordination We now get back to the two-node coordination setting illustrated in Figure 5.1. We will restrict our attention to the case where $\mathcal{U} = \{0,1\}$, $\bar{P}_U \sim \mathcal{B}(1/2)$, and the conditional distribution of actions $\bar{P}_{V|U}$ is symmetric.

Let Z be an auxiliary binary random variable satisfying the following conditions:

a) $U - Z - V$ forms a Markov chain, i.e. $\bar{P}_{UZV} = \bar{P}_U \bar{P}_{Z|U} \bar{P}_{V|Z}$;

b) the transition probabilities $\bar{P}_{U|Z}$ and $\bar{P}_{V|Z}$ correspond to symmetric channels.

We first construct polar codes of length $n = 2^m$ for the channel with transition probabilities $\bar{W} = \bar{P}_{UV|Z}$ as follows.

- For the symmetric channel $\bar{W}$, and for $i \in [\![1,n]\!]$, we let $\bar{W}_n^{(i)}$ be the corresponding set of bit channels. We define the sets

$$\begin{aligned} \mathcal{G}_{UV|Z} &= \left\{ i \in [\![1,n]\!] : C(\bar{W}_n^{(i)}) \geq 2^{-n^\beta} \right\}, \\ \mathcal{P}_{UV|Z} &= [\![1,n]\!] \setminus \mathcal{G}_{UV|Z}. \end{aligned} \tag{5.5}$$

- For the symmetric channel $\tilde{W} = \bar{P}_{U|Z}$, and for $i \in [\![1,n]\!]$, we let $\tilde{W}_n^{(i)}$ be the corresponding set of bit channels. We define the sets

$$\begin{aligned} \mathcal{G}_{U|Z} &= \left\{ i \in [\![1,n]\!] : C(\tilde{W}_n^{(i)}) \geq 2^{-n^\beta} \right\}, \\ \mathcal{P}_{U|Z} &= [\![1,n]\!] \setminus \mathcal{G}_{U|Z}. \end{aligned} \tag{5.6}$$

One can show that these sets satisfy the following property:

$$\mathcal{P}_{U|Z} \subset \mathcal{P}_{UV|Z}, \quad \mathcal{B}_{UV|Z} \subset \mathcal{B}_{U|Z}. \tag{5.7}$$

Consequently, the sets F_1, F_2 and F_3 defined as

$$F_1 = \mathcal{P}_{UV|Z}, \quad F_2 = \mathcal{G}_{UV|Z} \cap \mathcal{P}_{U|Z}, \quad F_3 = \mathcal{G}_{UV|Z} \cap \mathcal{G}_{U|Z}.$$

form a partition of $[\![1,n]\!]$, which is illustrated in Figure 5.2.

We now exploit these sets to construct a coordination code, by leveraging results on lossy source coding with polar codes [133]. The bits in positions F_1 are frozen bits with values $\mathbf{u}_{F_1} = \mathbf{0}_{F_1}$ fixed at all times. The encoding and decoding procedures are then the following.

Operation at Node 1: To encode a sequence of binary actions $\mathbf{x} \in \{0,1\}^n$ provided by nature, Node 1 performs successive-cancellation (SC) encoding [133] to determine the value of the bits $\mathbf{u}_{F_3}$ in F_3, using the bits $\mathbf{u}_{F_2}$ from the common randomness in positions F_2 and the frozen bits $\mathbf{u}_{F_1}$ in position F_1.

The bits in F_3 are then transmitted to Node 2. Note that the encoding complexity is that of SC encoding, which is $O(n \log n)$.

Operation at Node 2: To create a sequence of coordinated actions $\mathbf{v} \in \mathcal{V}^n$, Node 2 creates a vector $\mathbf{u}$ with frozen bits $\mathbf{u}_{F_1}$, common randomness bits $\mathbf{u}_{F_2}$, and received bits $\mathbf{u}_{F_3}$ in positions F_1, F_2, F_3, respectively. It then computes the vector $G_n\mathbf{u}$, and simulates its transmission over a memoryless channel with transition probabilities

$\bar{P}_{V|W}$. The resulting vector $\mathbf{v}$ is used as the sequence of coordinated actions. The encoding complexity is again $O(n \log n)$.

The scheme operates at communication rate $R = \frac{|F_3|}{n}$ and requires a rate $R_0 = \frac{|F_2|}{n}$ of common randomness. Our main result is the following.

Proposition 5.9 *For any random variable Z satisfying the conditions a) and b), $(\bar{P}_{UV}, R, R_0)$ is achievable for strong coordination if*

$$R + R_0 > I(UV; Z) \quad \textit{and} \quad R > I(U; Z).$$

Sketch of proof. One can show that the joint distribution $\tilde{P}(\mathbf{x}, \mathbf{v})$ induced by the encoding/decoding procedures is asymptotically close to the target distribution $\bar{P}_{UV}^{\otimes n}(\mathbf{x}, \mathbf{v})$. By recalling that $F_2 \cup F_3 = \mathcal{G}_{UV|Z}$ and that the bits in positions F_2 and F_3 are i.i.d $\mathcal{B}(1/2)$ random bits, Proposition 5.6 guarantees that the coding scheme is a resolvability code for $\bar{W}$ with a resolution rate $R + R_0$ satisfying $\lim_{n\to\infty} \frac{1}{n} \left|\mathcal{G}_{UV|Z}\right| = C(\bar{W}) = I(UV; Z)$. Moreover, since $F_3 = \mathcal{G}_{U|Z}$ and since the bits in position F_3 are i.i.d. $\mathcal{B}(1/2)$ random bits, Proposition 5.6 and Remark 5.8 ensure that $\lim_{n\to\infty} \frac{1}{n} \left|\mathcal{G}_{U|W}\right| = C(\tilde{W}) = I(U; Z)$. □

In general, the achievable coordination region with polar codes given in Proposition 5.9 is strictly smaller than the coordination capacity region $\mathcal{R}_{\text{Cuff}}$ because of the constraints on the random variable Z.

Follow-up work. Our polar coding scheme was restricted to the case where the action $\bar{P}_U$ imposed by nature is binary and uniform, and the action to coordinate is obtained via a symmetric discrete memoryless channel $\bar{P}_{V|U}$. It was extended to general actions in [47] using a block-Markov encoding scheme.

5.2 Coordination of signals and actions over noisy channels

Up to now, we have studied the problem of coordination assuming that an error-free line of communication is available between the agents. In Giulia Cervia's PhD thesis [40], [J11], we considered a more realistic scenario where the two agents communicate through a noisy channel.

In this setting, the signals that are transmitted and received over the physical channel become a part of what can be observed. One may therefore wish to coordinate both behaviors and communication, so that the sequence of both signals and actions follows a prescribed joint distribution [59]. This is particularly interesting if security is required: if for example we require the actions of the agents to appear independent of the communication, a malicious eavesdropper who observes the channel cannot infer anything about the source and the reconstruction without having access to the source of common randomness [201].

This scenario presents two conflicting goals: the encoder needs to convey a message to the decoder to coordinate the actions, while simultaneously coordinating the signals coding the message.

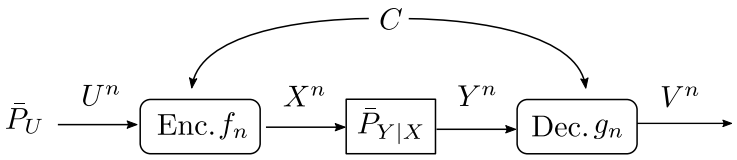


Figure 5.3: Coordination of signals and actions for a two-node network with a noisy channel with non-causal encoder and decoder.

Figure 5.3 illustrates the coordination of signals and actions in a two-node network over a noisy channel. We assume that encoder and decoder are *non-causal*, and that they have access to a shared source of uniform randomness $C \in [1, 2^{nR_0}]$. The encoder observes an i.i.d. source $U^n \in \mathcal{U}^n$ with distribution $\bar{P}_U$, and selects a signal $X^n = f_n(U^n, C)$. The signal X_n is transmitted over a discrete memoryless channel parametrized by the conditional distribution $\bar{P}_{Y|X}$. Upon observing Y_n and C, the stochastic decoder selects an action $V^n = g_n(Y^n, C)$.

The definitions of empirical and strong coordination can be extended to this setting as follows:

Definition 5.10 (Empirical coordination of signals and actions) *A distribution $\bar{P}_{UXYV}$ is* achievable for empirical coordination *if $\forall \epsilon > 0$ there exists a sequence $\{(f_n, g_n)\}$ of encoders-decoders, and $\exists n_0$ such that $\forall n \geq n_0$,*

$$\mathbb{P}\left\{\mathbb{V}\left(T_{U^n X^n Y^n V^n}, \bar{P}_{UXYV}\right) > \epsilon\right\} < \epsilon,$$

where $T_{U^n X^n Y^n V^n}$ is the joint histogram of signals and actions induced by the code. The empirical coordination region $\mathcal{R}_e$ *is the closure of the set of achievable distributions $\bar{P}_{UXYV}$.*

Definition 5.11 (Strong coordination of signals and actions) *A pair $(\bar{P}_{UXYV}, R_0)$ is achievable for strong coordination if there exists a sequence $\{(f_n, g_n)\}$ of encoders-decoders with rate of common randomness R_0, such that*

$$\lim_{n\to\infty} \mathbb{V}\left(P_{U^n X^n Y^n V^n}, \bar{P}_{UXYV}^{\otimes n}\right) = 0$$

where $P_{U^n X^n Y^n V^n}$ is the joint distribution induced by the code. The strong coordination region $\mathcal{R}$ *is the closure of the set of achievable pairs $(\bar{P}_{UXYV}, R_0)$.*

An outer and inner bound for the empirical coordination region for the setting of Figure 5.3 were characterized in [59].

Bounds for the strong coordination region In [J11], as part of Giulia Cervia's thesis [40], we focused on the problem of achieving strong coordination for the same setting. We derive an inner and an outer bound for the strong coordination region by developing a joint source-channel scheme in which an auxiliary codebook allows us to satisfy simultaneously the coordination of signals and of actions.

Theorem 5.12 *Let $\bar{P}_U$ and $\bar{P}_{Y|X}$ be the given source and channel parameters, then $\mathcal{R}'_{in} \subseteq \mathcal{R} \subseteq \mathcal{R}'_{out}$ where:*

$$\mathcal{R}'_{in} := \left\{ \begin{array}{l} (\bar{P}_{UXYV}, R_0) \text{ such that:} \\ \bar{P}_{UXYV} = \bar{P}_U \bar{P}_{X|U} \bar{P}_{Y|X} \bar{P}_{V|UXY} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \ \bar{P}_{UXYWV} = \bar{P}_U \bar{P}_{W|U} \bar{P}_{X|UW} \bar{P}_{Y|X} \bar{P}_{V|WY} \\ I(W;U) \leq I(W;Y),\ R_0 \geq I(W;UXV|Y) \end{array} \right\} \tag{5.8}$$

$$\mathcal{R}'_{out} := \left\{ \begin{array}{l} (\bar{P}_{UXYV}, R_0) \text{ such that:} \\ \bar{P}_{UXYV} = \bar{P}_U \bar{P}_{X|U} \bar{P}_{Y|X} \bar{P}_{V|UXY} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \ \bar{P}_{UXYWV} = \bar{P}_U \bar{P}_{W|U} \bar{P}_{X|UW} \bar{P}_{Y|X} \bar{P}_{V|WY} \\ I(W;U) \leq I(X;Y),\ R_0 \geq I(W;UXV|Y) \\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}| + 4 \end{array} \right\}. \tag{5.9}$$

Remark 5.13 The decomposition of the joint distributions $\bar{P}_{UXYV}$ and $\bar{P}_{UWXYV}$ is equivalently characterized in terms of Markov chains:

$$Y - X - U, \qquad \begin{cases} Y - X - (U, W) \\ V - (Y, W) - (X, U) \end{cases}. \tag{5.10}$$

Note that the information constraint $I(W;U) \leq I(W;Y)$ and the decomposition of the joint probability distribution are the same as for empirical coordination [59, Theorem 1]. The main difference is that strong coordination requires a positive rate of common randomness $R_0 > I(W;UXV|Y)$.

We design an achievability proof for the inner bound by developing a joint source-channel scheme in which an auxiliary codebook allows us to simultaneously coordinate signals and actions. More precisely, we use the *random binning* method introduced by [232]: to prove the existence of coding schemes which induce a certain target joint distribution, we proceed in two steps. First, we define a random binning scheme for the n-letter target i.i.d. distribution. Then, we define a random coding scheme such that the joint distributions induced by the random binning and the random coding scheme are close in variational distance. We first show how to coordinate $(U^n, X^n, Y^n, V^n, W^n)$ using common randomness C and extra randomness F, and then prove that we can fix

the value of the extra randomness if we only require (U^n, X^n, Y^n, V^n) to be coordinated. More details about this proof and the proof of the outer bound can be found in [J11].

Polar code construction A second contribution of Giulia Cervia's thesis for this setting is an explicit polar code construction achieving the inner bound of Theorem 5.12, provided that an error-free channel of negligible rate is available between the encoder and decoder.

For the sake of simplicity, as in Section 5.1, we only focus on the set of achievable distributions in $\mathcal{R}'_{\text{in}}$ for which the auxiliary variable W is binary. The scheme can be extended to the case of a non-binary random variable W using non-binary polar codes [200].

Theorem 5.14 *The subset of the region $\mathcal{R}'_{in}$ defined in (5.8) for which the auxiliary random variable W is binary is achievable using polar codes, provided there exists an error-free channel of negligible rate between the encoder and decoder.*

To convert the information-theoretic achievability proof of Theorem 5.12 into a polar coding proof, we use *source polarization* [12] (rather than channel polarization as in Section 5.1) to induce the desired joint distribution. Inspired by [45], we want to translate the random binning scheme into a polar coding scheme. The key idea is that the information contraints and rate conditions found in the random binning proof directly convert into the definition of the polarization sets. In the random binning scheme we reduced the amount of common randomness F by having the nodes to agree on an instance of F, here we recycle some common randomness using a chaining construction as in [111, 169].

For $n = 2^m$, we note $G_n = G_2^{\otimes n}$ the polarization transform defined in Section 5.1. Let $R^n \coloneqq G_n W^n$ be the polarization of the auxiliary variable W^n. For some $0 < \beta < 1/2$, let $\delta_n = 2^{-n^\beta}$ and define the *very high entropy* and *high entropy* sets:

$$\begin{aligned}
\mathcal{V}_W &\coloneqq \left\{j \in [\![1,n]\!] : H(R_j|R^{j-1}) > 1-\delta_n\right\},\\
\mathcal{V}_{W|Y} &\coloneqq \left\{j \in [\![1,n]\!] : H(R_j|R^{j-1}Y^n) > 1-\delta_n\right\},\\
\mathcal{H}_{W|Y} &\coloneqq \left\{j \in [\![1,n]\!] : H(R_j|R^{j-1}Y^n) > \delta_n\right\}.
\end{aligned} \tag{5.11}$$

One consequence of source polarization is the fact that it is possible to compress the source W^n using Y^n as side information by selecting the high entropy bits $R^n[\mathcal{H}_{X|Y}]$. The reconstruction can be done using successive cancellation encoding [133], with error probability smaller than δ_n. Now define the following disjoint sets:

$$\begin{aligned}
A_1 &\coloneqq \mathcal{V}_{W|U} \cap \mathcal{H}_{W|Y}, \quad A_2 \coloneqq \mathcal{V}_{W|U} \cap \mathcal{H}^c_{W|Y},\\
A_3 &\coloneqq \mathcal{V}^c_{W|U} \cap \mathcal{H}_{W|Y}, \quad A_4 \coloneqq \mathcal{V}^c_{W|U} \cap \mathcal{H}^c_{W|Y}.
\end{aligned}$$

We have:

- $\mathcal{V}_{W|Y} \subset \mathcal{H}_{W|Y}$ and $\lim_{n\to\infty} \frac{|\mathcal{H}_{W|Y} \setminus \mathcal{V}_{W|Y}|}{n} = 0$ [12],
- $\lim_{n\to\infty} \frac{|\mathcal{V}_{W|U}|}{n} = H(W|U)$ [46],
- $\lim_{n\to\infty} \frac{|\mathcal{H}_{W|Y}|}{n} = H(W|Y)$ [12].

Since $H(W|U) - H(W|Y) = I(W;Y) - I(W;U)$, for sufficiently large n, the assumption $I(W;Y) \geq I(W;U)$ directly implies that $|A_2| \geq |A_3|$.

We give a high-level view of the encoder / decoder structure, which is pictured in Figure 5.2.

The key idea is that at the encoder, the bits in A_1 and A_2 are almost uniform given U^n, while the bits in A_3 and A_4 are almost deterministic given U^n and can be generated according to $P_{R_j|R^{j-1}U^n}$ using successive cancellation encoding.

On the other hand, at the decoder, the bits in A_1 and A_3 can be chosen almost uniformly given Y^n, while the bits in A_2 and A_4 are almost deterministic given Y^n and can be generated according to $P_{R_j|R^{j-1}Y^n}$ using successive cancellation encoding.

Special care is needed to handle the set A_3 of bits which are almost deterministic for the encoder but cannot be recovered reliably at the decoder. The solution is to use a chaining construction as in [47, 169], in order to send the set A_3 for block k in advance in the previous block $k-1$, using a one-time pad.

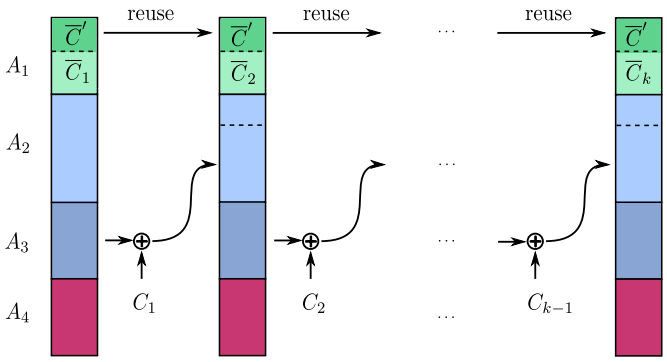


Figure 5.4: Chaining construction for block Markov encoding

5.3 Coordination in two-node networks with two-sided state information

Up to now, we have assumed that the source and the channel follow distributions which are fixed ahead of time and known by the agents. However, this prevents us from modeling situations in which the agent reacts to an external stimulus, and in which the channel statistics depend on the environment. For instance, the actions of an agent might be constrained by obstacles that prevent it from making certain choices. In this case the probability distributions given by nature could change with time and might be partially or completely unknown to some of the agents. To include such situations in the coordination framework, as part of Giulia Cervia's thesis we extended the model to take into account the uncertainty about the source and channel distribution. This setting has already been taken into consideration in [140, 138, 139, 142] for empirical coordination.

We consider the model depicted in Figure 5.5, where we introduce a state in the description of the behavior. It consists of a state-dependent i.i.d. source (U^n, S^n, Z^n) generated according to $\bar{P}_{USZ}$ and a state-dependent discrete memoryless channel $\bar{P}_{Y|XS}$. The encoder selects a signal $X^n = f_n(U^n, C)$, and transmits it over the channel. The decoder then selects an action $V^n = g_n(Y^n, Z^n, C)$.

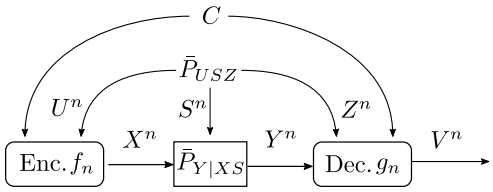


Figure 5.5: Coordination of signals and actions for a two-node network with a noisy channel with state and side information at the decoder.

The channel state information and side information at the decoder are represented by the random variables S^n and Z^n respectively, and we make no assumptions on the correlation of (U^n, S^n, Z^n). This includes scenarios where the encoder has access to partial, perfect or noisy channel state information, since the variables U^n and S^n are possibly correlated. Moreover, the decoder side information Z^n can contain partial, perfect or noisy information on the channel state, on the source, or on both of them.

In the case of non-causal encoder and decoder, the problem of characterizing the strong coordination region $\mathcal{R}_{\text{state}}$ for the model in Figure 5.5 is still open, but we establish the following inner and outer bounds [J11].

Theorem 5.15 *Let $\bar{P}_{USZ}$ and $\bar{P}_{Y|XS}$ be the given source and channel parameters. Then $\mathcal{R}_{in} \subseteq \mathcal{R} \subseteq \mathcal{R}_{out}$ where:*

$$\mathcal{R}_{in} := \left\{ \begin{array}{l} (\bar{P}_{USZXYV}, R_0) \text{ such that:} \\ \bar{P}_{USZXYV} = \bar{P}_{USZ}\bar{P}_{X|U}\bar{P}_{Y|XS}\bar{P}_{V|UXYSZ} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \bar{P}_{USZWXYV} = \bar{P}_{USZ}\bar{P}_{W|U}\bar{P}_{X|UW}\bar{P}_{Y|XS}\bar{P}_{V|WYZ} \\ I(W;U) \leq I(W;YZ),\ R_0 \geq I(W;USXV|YZ) \end{array} \right\}, \tag{5.12}$$

$$\mathcal{R}_{out} := \left\{ \begin{array}{l} (\bar{P}_{USZXYV}, R_0) \text{ such that:} \\ \bar{P}_{USZXYV} = \bar{P}_{USZ}\bar{P}_{X|U}\bar{P}_{Y|XS}\bar{P}_{V|UXYSZ} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \bar{P}_{USZWXYV} = \bar{P}_{USZ}\bar{P}_{W|U}\bar{P}_{X|UW}\bar{P}_{Y|XS}\bar{P}_{V|WYZ} \\ I(W;U) \leq \min\{I(XUS;YZ), I(XS;Y) + I(U;Z)\} \\ R_0 \geq I(W;USXV|YZ) \\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{S} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}| + 5 \end{array} \right\}. \tag{5.13}$$

Remark 5.16 Observe that the decomposition of the joint distributions $\bar{P}_{USZXYV}$ and $\bar{P}_{USZWXYV}$ is equivalently characterized in terms of Markov chains:

$$\left\{ \begin{array}{l} Z - (U,S) - (X,Y) \\ Y - (X,S) - U \end{array} \right. , \quad \left\{ \begin{array}{l} Z - (U,S) - (X,Y,W) \\ Y - (X,S) - (U,W) \\ V - (Y,Z,W) - (X,S,U) \end{array} \right. . \tag{5.14}$$

5.3.1 Strong coordination region for special cases

Although the inner and outer bounds of Theorem 5.15 do not match in general, we are able to characterize the strong coordination region for some special cases: perfect channel, lossless decoding and separation between the channel and the source. In all these cases, the achievability proof is merely a consequence of the general achievability proof of Theorem 5.15. The converse proofs, on the other hand, rely of the specifics of each setting, and need to be proven separately.

The empirical coordination region for these three settings was derived in [142, 141]. For strong coordination we recover the same information constraints, but we show that a positive rate of common randomness is required. This corroborates the conjecture that, given enough common randomness, the strong coordination capacity region is the same as the empirical coordination capacity region for any network setting [56].

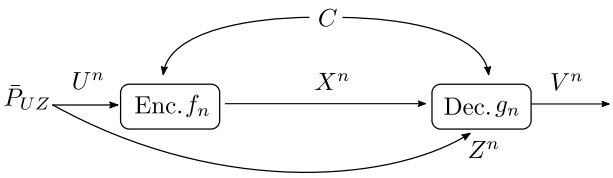


Figure 5.6: Coordination of signals and actions for a two-node network with a perfect channel with side information at the decoder.

Perfect channel The first special setting corresponds to the case of a perfect channel as in Figure 5.6. In this case $X^n = Y^n$ and Z^n plays the role of side information at the decoder.

Theorem 5.17 *In the setting of Theorem 5.15, suppose that $\bar{P}_{Y|XS}(\mathbf{y}|\mathbf{x},\mathbf{s}) = \mathbb{1}_{X=Y}\{\mathbf{x} = \mathbf{y}\}$. Then the strong*

coordination region is

$$\mathcal{R}_{PC} \coloneqq \left\{ \begin{array}{l} (\bar{P}_{UZXV}, R_0) \text{ such that:} \\ \bar{P}_{UZXV} = \bar{P}_{UZ}\bar{P}_{X|U}\bar{P}_{V|UXZ} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \ \bar{P}_{UZWXV} = \bar{P}_{UZ}\bar{P}_{W|U}\bar{P}_{X|UW}\bar{P}_{V|WXZ} \\ I(WX;U) \leq H(X) + I(W;Z|X) \\ R_0 \geq I(W;UV|XZ) \\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{V}| + 4 \end{array} \right\} \tag{5.15}$$

Lossless decoding Here, we investigate a special case where the decoder wants to reconstruct the source losslessly, i.e., $V = U$ as in Figure 5.7.

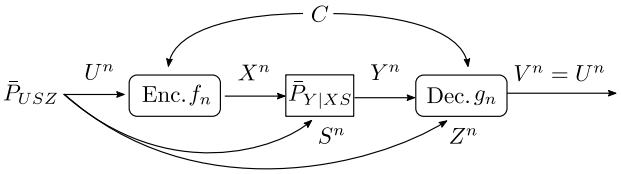


Figure 5.7: Coordination of signals and actions for a two-node network with a noisy channel and a lossless decoder.

The following theorem characterizes the strong coordination region $\mathcal{R}_{\text{LD}}$.

Theorem 5.18 *Consider the setting of Theorem 5.15 and suppose that* $\bar{P}_{V|USXYZ}(\mathbf{v}|\mathbf{u},\mathbf{s},\mathbf{x},\mathbf{y},\mathbf{z}) = \mathbb{1}_{V=U}\{\mathbf{u}=\mathbf{v}\}$. *Then the strong coordination region is*

$$\mathcal{R}_{LD} \coloneqq \left\{ \begin{array}{l} (\bar{P}_{USZXY}, R_0) \text{ such that:} \\ \bar{P}_{USZXYV} = \bar{P}_{USZ}\bar{P}_{X|U}\bar{P}_{Y|XS}\mathbb{1}_{V=U} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \ \bar{P}_{USZWXYV} = \bar{P}_{USZ}\bar{P}_{W|U}\bar{P}_{X|UW}\bar{P}_{Y|XS}\mathbb{1}_{V=U} \\ I(W;U) \leq I(W;YZ) \\ R_0 \geq I(W;USX|YZ) \\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{S} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{Y}| + 3 \end{array} \right\} \tag{5.16}$$

Remark 5.19 An equivalent characterization of the region is:

$$\mathcal{R}_{\text{LD}} \coloneqq \left\{ \begin{array}{l} (\bar{P}_{USZXY}, R_0) \text{ such that:} \\ \bar{P}_{USZXY} = \bar{P}_{USZ}\bar{P}_{X|U}\bar{P}_{Y|XS} \\ \exists\, W \text{ taking values in } \mathcal{W},\ \bar{P}_{USZWXY} = \bar{P}_{USZ}\bar{P}_{W|U}\bar{P}_{X|UW}\bar{P}_{Y|XS} \\ H(U) \leq I(WU;YZ) \\ R_0 \geq I(W;USX|YZ) + H(U|WYZ) \\ |\mathcal{W}| \leq |\mathcal{U} \times \mathcal{S} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{Y}| + 1 \end{array} \right\} \tag{5.17}$$

The constraint for the mutual information in (5.17) is the same as for the empirical coordination region [142, Section IV.B].

Independence between source and channel Suppose that the channel state P_{S^n} is independent of the source and side information $P_{U^nZ^n}$, and that the target joint distribution is of the form $\bar{P}_{UZV}^{\otimes n}\bar{P}_{SXY}^{\otimes n}$. For simplicity, we will suppose that the encoder has perfect state information (see Figure 5.8).

In this case the coordination requirements are three-fold: the random variables (U^n, Z^n, V^n) should be coordinated, the random variables (S^n, X^n, Y^n) should be coordinated and finally (U^n, Z^n, V^n) should be independent

of (S^n, X^n, Y^n). We introduce two auxiliary random variables W_1 and W_2, where W_2 is used to accomplish the coordination of (U^n, Z^n, V^n), while W_1 has the double role of ensuring the independence of source and state as well as coordinating (S^n, X^n, Y^n).

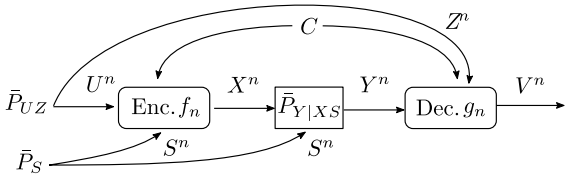


Figure 5.8: Coordination of signals and actions for a two-node network with a noisy channel where the source is separated from the channel.

Theorem 5.20 *Consider the setting of Theorem 5.15 and suppose that $\bar{P}_{USXYZV} = \bar{P}_{UZV}\bar{P}_{SXY}$. Then, the strong coordination region is*

$$\mathcal{R}_{IND} \coloneqq \left\{ \begin{array}{l} (\bar{P}_{USZXY}, R_0) \text{ such that:} \\ \bar{P}_{USZXYV} = \bar{P}_{UZ}\bar{P}_{V|UZ}\bar{P}_S\bar{P}_{X|S}\bar{P}_{Y|XS} \\ \exists\ (W_1, W_2) \text{ taking values in } \mathcal{W}_1 \times \mathcal{W}_2 \\ \bar{P}_{USZW_1W_2XYV} = \bar{P}_{UZ}\bar{P}_{W_2|U}\bar{P}_{V|ZW_2}\bar{P}_S\bar{P}_{X|S}\bar{P}_{W_1|SX}\bar{P}_{Y|XS} \\ I(W_1;S) + I(W_2;U) \leq I(W_1;Y) + I(W_2;Z) \\ R_0 \geq I(W_1;SX|Y) + I(W_2;UV|Z) \\ (|\mathcal{W}_1|, |\mathcal{W}_2|) \leq |\mathcal{U} \times \mathcal{S} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}| + 4. \end{array} \right\} \tag{5.18}$$

Coordination under secrecy constraints It turns out that in the separation setting of the previous section, strong coordination offers additional security guarantees "for free". In this context, the common randomness is not only useful to coordinate signals and actions of the nodes but plays the role of a secret key shared between the two legitimate users.

To simplify the notation, we do not consider channel state and side information at the decoder (however, one can show that the result holds more generally). Suppose that an eavesdropper observes the signals sent over the channel (see Figure 5.9). We will show that not knowing the common randomness, the eavesdropper cannot infer any information about the actions.

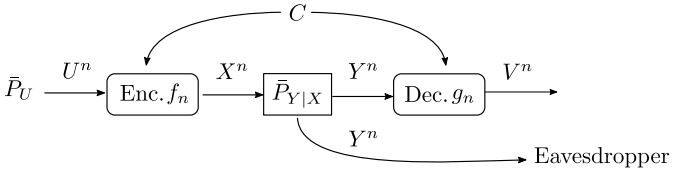


Figure 5.9: Wiretap channel: strong coordination implies secrecy.

Lemma 5.21 *In the setting of Theorem 5.20, without state and side information at the decoder, suppose that there is an eavesdropper that receives the same sequence Y^n as the decoder but has no knowledge of the common randomness. Then there exists a sequence (f_n, g_n) of strong coordination codes achieving the pair $(\bar{P}_{UV}\bar{P}_{XY}, R_0) \in \mathcal{R}_{IND}$ such that the induced joint distribution $P^{RC}_{U^nV^nX^nY^n}$ satisfies the strong secrecy condition*

$$\lim_{n\to\infty} \mathbb{D}(P^{RC}_{U^nV^nY^n} \| P^{RC}_{U^nV^n} P^{RC}_{Y^n}) = \lim_{n\to\infty} I(U^nV^n; Y^n) = 0. \tag{5.19}$$

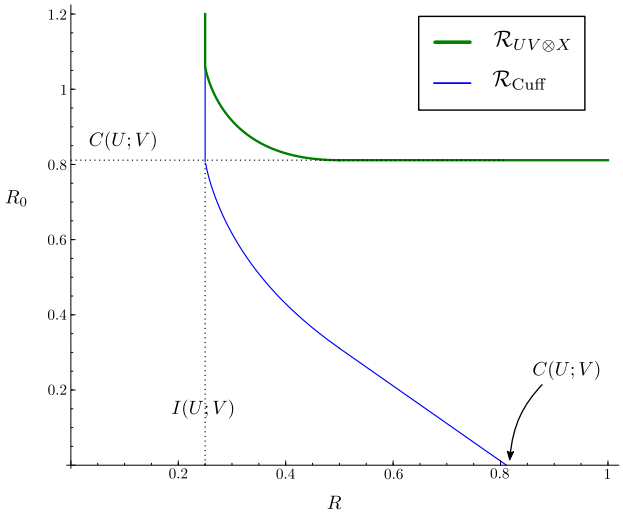


Figure 5.10: Comparison of the joint coordination region $\mathcal{R}_{UV\otimes X}$ with $\mathcal{R}_{\text{Cuff}}$ [58, 55]: boundaries of the regions for a binary erasure channel with erasure probability $p_e = 0.75$ and a Bernoulli-half input.

Validity of the separation principle When extending the analysis of coordination from error-free to noisy channels, it is natural to ask whether some form of joint source-channel coordination version of Shannon's source-channel separation theorem [204] holds. In this section, we show that unlike the case of empirical coordination [143], separation does not hold for strong coordination.

If the separation principle is still valid for strong coordination, by concatenating the strong coordination of the source and its reconstruction with the strong coordination of the input and output of the channel we should retrieve the same mutual information and rate constraints. We will show that this is not the case.

As a counterexample, we compare the optimal region for strong coordination of actions with our result on joint coordination of signals and actions in the case in which the channel is perfect and the target joint distribution is of the form $\bar{P}_{UV}^{\otimes n}\bar{P}_{X}^{\otimes n}$. The choice of a perfect channel might appear counterintuitive; as a matter of fact, if the separation principle holds for any noisy link, it should in particular hold for a perfect one. We consider the two-node network with fixed source $\bar{P}_U$ and an error-free link of rate R in Figure 5.1, and the corresponding optimal region $\mathcal{R}_{\text{Cuff}}$ in (5.2) for strong coordination of actions.

We compare this region to our results when the requirement to coordinate the signals X^n and Y^n in addition to the actions U^n and V^n is relaxed. We consider, in the simpler scenario with no state and no side information, the intersection $\mathcal{R}_{UV\otimes X} \coloneqq \mathcal{R}_{\text{PC}} \cap \mathcal{R}_{\text{IND}}$.

Figure 5.10 shows the difference of the two regions in the case of a Bernoulli(1/2) source U, where V is the output of a binary erasure channel with input U.

Observe that, while the information constraint is the same in the two regions $\mathcal{R}_{UV\otimes X}$ and $\mathcal{R}_{\text{Cuff}}$, the rate of common randomness R_0 required in $\mathcal{R}_{UV\otimes X}$ is larger than the rate of common randomness in $\mathcal{R}_{\text{Cuff}}$. In fact, in the setting of Figure 5.1 both X^n and the pair (U^n, V^n) achieve coordination separately (i.e. P_X^n is close to $\bar{P}_X^{\otimes n}$ and $P_{U^nV^n}$ is close to $\bar{P}_{UV}^{\otimes n}$ in total variation distance), but there is no extra constraint on the joint distribution $P_{U^nX^nV^n}$. On the other hand, the setting in $\mathcal{R}_{UV\otimes X}$ requires the control of the joint distribution $P_{U^nX^nV^n}$, which requires more common randomness.

Moreover, note that as $R = H(X)$ tends to infinity, similarly to [136] the minimum rate of common randomness R_0 needed for strong coordination is Wyner's common information $C(U;V)$. In particular to achieve joint strong coordination of (U, X, V) a positive rate of common randomness is required. More details can be found in [J11].

5.4 Coordination of signals and actions with strictly causal encoder

Until now we have considered joint source-channel coordination in the presence of a non-causal encoder and non-causal decoder. In [C23], we also examined the case in which the encoder is strictly causal, which has the benefit of shortening the transmission delay. For simplicity, we considered the setting without state and side information. For empirical coordination, [59] provides a complete characterization of the region. Although the strong coordination region is still unknown, we provide an inner and an outer bound that differ only in the amount of common randomness needed to strongly coordinate signals and actions. The achievability proof relies on a random binning argument, but the nature of this setting presents some extra difficulties. In fact, the information about the source at time i is needed for the reconstruction, but is observed by the encoder only at time $i+1$. So this information must be recovered by the decoder at a later time. In order to ensure coordination, we use a block-Markov scheme and a one-time pad. Finally, in Giulia Cervia's thesis [40, Chapter 6] we also proved that polar codes provide a constructive alternative to random binning proofs and we described an explicit scheme for strong coordination.

Open problems and follow-up work Despite the fact that we have fully characterized the region of strong coordination for signals and actions in some special cases, inner and outer bound differ in general on the information constraint. Closing this gap is left for future study.

Though we provided an explicit polar coding construction, our scheme relies on chaining over several blocks, which is not practical for delay-constrained applications. This is another issue that may be studied further.

One of the most interesting consequences of strong coordination is the fact that under particular circumstances it offers security "for free": by coordinating signals and actions, the synthesized sequences would appear to be statistically indistinguishable from i.i.d. to an outside observer. The related problem of secure strong coordination was studied by Cervia, Bassi and Skoglund in [41].

Moreover, our results could be extended to a strategic coordination setting. This represents a scenario where the objectives of the two agents are not necessarily aligned, and has been investigated for empirical coordination in [144].

6 | OPEN PROBLEMS AND PERSPECTIVES

This section presents some open problems and perspectives for future work in the areas of coding for physical layer security and lattice and code-based post-quantum cryptography. Despite the diversity of the applications, it is interesting to note that similar tools (the smoothing parameter, which characterizes the performance of average-case to worst-case reductions, and has recently been extended to codes [65]) can be used to provide both computational security and information-theoretic secrecy[1]. These connections are still largely unexplored and call for a deeper understanding that would bring together the cryptography and information theory communities.

6.1 Physical layer security

Physical layer security techniques are seriously being envisaged for future 6G networks [170, 44] as a first line of defense with low latency and low computational cost. A key requirement for their practical implementation is the guarantee of an asymmetry in the signal quality between legitimate terminals and attackers. New technologies, such as sharp beamforming using massive MIMO as well as distributed MIMO, channel engineering using Reconfigurable Intelligent Surfaces, as well as transmission in the THz range with high directivity for short range scenarios, could provide such security advantages, allowing in particular to make the case for wiretap coding.
In the short term, I plan to address some open questions related to wiretap coding, covert communication and key generation / fuzzy extractors. This research will be led in part in the context of national (PEPR 5G) and international (EU HORIZON-SNS) projects on security for 6G, both in collaboration with Arsenia Chorti.

Design of short packet wiretap codes

The potential use cases for wiretap codes include Internet of things (IoT) networks featuring short packet payloads, and ultra-reliable low latency communications (URLLC). In this context, the development of short blocklength wiretap codes could have high practical impact.

It is therefore important to obtain precise bounds for the information leakage of wiretap codes in finite blocklength. The first non-asymptotic bounds for the leakage in wiretap channels were obtained by Hayashi [114]. Building on the theoretical breakthrough by Polyanskiy, Poor and Verdù in the analysis of finite-length channel coding rates [189], Yang, Schaefer and Poor [230] proved tight bounds on the second-order coding rate for discrete memoryless and Gaussian wiretap channels. In place of the mutual information, they consider the average total variation distance $S(\mathsf{M}|\mathsf{Z}^n) = \mathbb{V}(p_{\mathsf{M}\mathsf{Z}^n}, \mathcal{U}_{\mathcal{M}} p_{\mathsf{Z}^n})$ as a measure of leakage[2], where M denotes the message belonging to a finite set $\mathcal{M}$, and Z^n the eavesdropper's observation. For degraded discrete memoryless wiretap channels and Gaussian wiretap channels, they show that the maximal secrecy rate $R^*(n, \epsilon, \delta)$ for block length n,

[1] In particular, it was already noted in [161, Chapter 4] that when the noise variance is too large compared to the smoothing parameter, the LWE problem becomes information-theoretically intractable.

[2] An alternative metric considered in [230] is the maximum total variation distance, which is related to distinguishing security and semantic security.

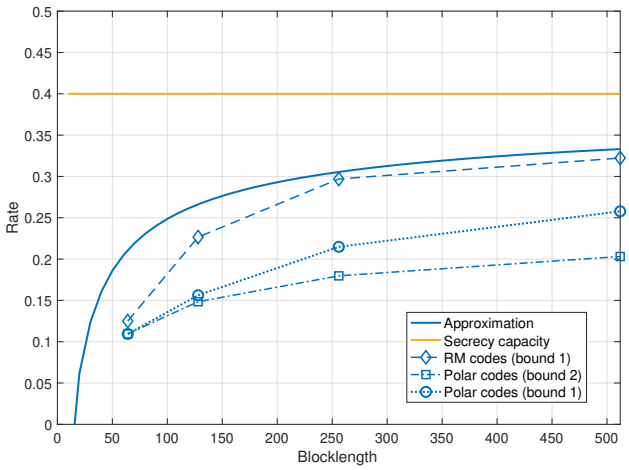


Figure 6.1: Comparison of the lower bound on achievable secrecy rates of Reed-Muller and polar codes for $p = 0.4$ and $\delta = 0.001$, with the second order approximation secrecy rate in (6.2) [C26]. Bound 1 is computed by Monte-Carlo simulations and bound 2 is obtained using the closed form expression for the bit-channel capacity.

error probability ϵ and under the constraint $S(\mathsf{M}|\mathsf{Z}^n) \leq \delta$ is bounded by

$$C_s - \sqrt{\frac{V_1}{n}}Q^{-1}(\epsilon) - \sqrt{\frac{V_2}{n}}Q^{-1}(\delta) + \mathcal{O}\left(\frac{\log n}{n}\right) \leq R^*(n,\epsilon,\delta) \leq C_s - \sqrt{\frac{V_3}{n}}Q^{-1}(\epsilon+\delta) + \mathcal{O}\left(\frac{\log n}{n}\right), \tag{6.1}$$

where C_s is the secrecy capacity, Q denotes the Q-function, and the constants V_1, V_2, V_3 depend on the distributions of the main channel and eavesdropper's channel. A simpler expression can be obtained for *semi-deterministic* wiretap channels, where the output of the main channel is a deterministic function of the input. In this case, assuming that $\epsilon + \delta < 1$, the maximal secrecy rate is

$$R^*(n,\epsilon,\delta) = C_s - \sqrt{\frac{V_S}{n}}Q^{-1}\left(\frac{\delta}{1-\epsilon}\right) + \mathcal{O}\left(\frac{\log n}{n}\right), \tag{6.2}$$

where V_S is the conditional variance of the information density of p_{XZ}. In particular, equation (6.2) evidences a trade off between reliability and secrecy.

In spite of these theoretical advances, few works have considered practical wiretap code constructions in finite blocklength. In [187], Monte-Carlo simulations were used to evaluate the equivocation of short wiretap codes over a BEC. Algebraic criteria to compute exact equivocation expressions have been found in [109]. In [175] randomized Reed-Muller code constructions were considered for Gaussian wiretap channels under a mutual information criterion for extremely short blocklengths ($n = 16$). This limitation is due to the complexity of providing a tight estimate of the leakage, which is done using neural networks. Other recent works [88, 25, 191] have used deep learning to design wiretap codes. However, the complexity of estimating the leakage remains a significant challenge, which currently limits their use to very short blocklengths or low rates.

For the polar wiretap code construction of Mahdavifar and Vardy [162], finite-length bounds for the leakage can be obtained directly in terms of the sum of the capacities of the "poor" bit-channels, as in equation (5.4). However, ensuring reliability is not straightforward [199].

In a collaboration with Arsenia Chorti and Mahdi Shakiba-Herfeh at ETIS [C26], we derived lower bounds on the secrecy performance of polar and Reed-Muller codes over a wiretap channel where where the main channel W_b is noiseless and the eavesdropper's channel W_e is a binary erasure channel, and compared them to the second-order coding rate (6.2) in the semi-deterministic case. A recent study by Abbe and Ye [1] has evidenced that Reed-Muller codes also have polarization properties, although this mechanism is still not well-understood. This allows to estimate their leakage through Mahdavifar and Vardy's bound.

As seen in Figure 6.1, in this simple wiretap channel model there is a significant gap between the lower bounds on the achievable secrecy rates of polar codes and the second order approximation secrecy rate. On the other hand, Reed-Muller codes show a promising performance. At blocklength 256 and for $\delta = 0.001$, the lower bound on achievable secrecy rate of Reed-Muller codes is less than 3% away from the second order approximation secrecy rate. This difference might be due to the speed of polarization of Reed-Muller codes, which is apparently faster than for polar codes [1][3].

Reed-Muller codes have very recently been shown to achieve capacity[4] on any DMC [192, 2] relying on the weight enumerator bounds in [202]. On the BEC and BSC, it is also known that they have optimal finite-length scaling [110]. However, they still suffer from the absence of an efficient decoding algorithm, in spite of recent advances; a new type of decoder, the recursive projection-aggregation decoder [233], seems to have good performance in the low and high rate regimes.

The evaluation of the achievable secrecy rate (or good approximations) for other codes and more general wiretap models is a timely open problem. I plan to pursue this topic in collaboration with Charles Pillet (École de Technologie Supérieure, Montreal) and Valerio Bioglio (Università di Torino).

In the case of polar codes, some methods providing good approximations of the bit-channels for other channel models are available in the literature [210]. One possible extension of our work could be the secrecy analysis of polar codes with large kernels, which achieve optimal scaling for the BEC [86]. A new polynomial formalism introduced by Bardet, Dragoi, Otmani and Tillich [16] shows that both polar and Reed-Muller codes can be viewed as decreasing monomial codes, whose properties are related to the size of their permutation groups. This new framework allowed Yao, Fazeli and Vardy to develop an efficient algorithm to compute the complete weight distribution of all decreasing monomial codes [231], which could be a useful tool to derive bounds for the average total variation distance metric in finite length.

Covert communication

As part of the INEX Ambition project PHEBE "Physical Layer Security for Beyond-5G", I am co-supervising the PhD thesis of Cécile Bouette in collaboration with Ligong Wang (formerly at ETIS, now with ETH Zurich) on the topic of covert communication or "communication with low probability of detection".

This term refers to a scenario where the legitimate transmitter and receiver want to hide from a potential eavesdropper the very fact that communication is taking place. Indeed, in some cases, even if the content of the message is not disclosed, just knowing meta-data such as *who* the communicating parties are, and *at what time* and *where* the communication is happening, might leak sensitive information.

Covert communication has a wide range of applications, and has been previously studied in the context of steganography and spread-spectrum techniques. The approach considered by the information theory community differs from the former in terms of the metrics being considered. It was first introduced by Bash, Goeckel and Towsley [17], who assumed that the eavesdropper or "warden" performs an optimal statistical test (such as the Neyman-Pearson test) to detect whether a transmission is ongoing (hypothesis H_0) or not (alternative H_1). The sum of the probabilities α and β of Type I and Type II errors for the optimal test is $\alpha + \beta = 1 - \mathbb{V}(Q_0, Q_1)$, where Q_0 and Q_1 are the output distributions of the eavesdropper's observation without / with communication respectively. Thus, in order for the communication to be almost undetectable, the variational distance $\mathbb{V}(Q_0, Q_1)$ should be small. In alternative, one can choose the stronger requirement that the Kullback-Leibler divergence $\mathbb{D}(Q_0||Q_1)$ should be small. It is assumed that the legitimate parties share a secret key to help with the transmission.

In [17, 42], it was observed that the maximum amount of information that can be transmitted under these requirements scales like the square root of the blocklength, both for Gaussian channels and for binary symmetric channels. Namely, the capacity in this setting is equal to zero. Wang, Wornell and Zheng [223] further showed that the "square root law" holds for a wide class of DMCs and for the Gaussian channel, and characterized its fundamental scaling constant

$$L = \lim_{\epsilon \to 0} \varliminf_{n \to \infty} \frac{K_n(\delta, \epsilon)}{\sqrt{n\delta}}, \tag{6.3}$$

[3]For polar codes, [196] shows that the rates of polarization of good and bad channels must be the same, so that the secrecy performance and error correction performance are linked.

[4]Note that this also implies that they achieve strong secrecy over the BSC [196].

where $K_n(\delta, \epsilon)$ is the maximum of $\ln |\mathcal{C}|$ for which there exists a code $\mathcal{C}$ of length n satisfying the covertness condition

$$\mathbb{D}(Q_0||Q_1) \leq \delta,$$

and whose average probability of decoding error for the legitimate receiver is at most ϵ. Independently, Bloch [31] studied this problem from a resolvability perspective, and characterized the minimum length of the key.

Channels with memory and non-Gaussian noise channels As part of Cécile Bouette's PhD thesis, our goal is to extend the results of [223] and to investigate whether the square root law holds for more general continuous channels, including non-Gaussian additive noise channels and channels with memory.
Some preliminary results are presented in our conference paper [C28]. The first scenario we consider is the Gaussian channel with memory, where the noise sequence is a Gaussian vector with an arbitrary invertible covariance matrix. We show that the fundamental limit for covert communication over such a channel is the same as in the memoryless case, that is, $L = 1$.
The second type of channel we consider is one with memoryless generalized Gaussian noise. For this family of noise distributions with shape parameter p [74], we prove the general upper bound $L \leq \sqrt{2/p}$ for covert communication over n channel uses. When $p \in (0, 1]$, we also prove a matching lower bound. The key property that we use to establish this lower bound is the self-decomposability of generalized Gaussian distributions in this range [73], which guarantees the existence of a suitable input distribution such that the output of the channel is a scaled version of the noise.

We are currently working on the characterization of the scaling constant L for a wide class of additive noise channels under mild integrability conditions. An interesting question is whether the square-root law still holds for heavy-tailed noise distributions such as α-stable distributions [77] and for general channels with memory.

We also plan to bound the size of the secret key shared between the legitimate users which gives the receiver the required advantage in order to decode the message reliably. This key size was characterized in [31] for covert communication over continuous channels when the covertness constraint is based on the variational distance; we plan to adapt this technique to the KL divergence metric.

Coding for covert communication. Codes for covert communication differ inherently from traditional error-correcting codes. For example, it is known that linear codes are strictly suboptimal for discrete-input channels [125]. While there has been progress on coding for discrete channels [237, 125], few works have considered the Gaussian channel, which is arguably more relevant in practice. The recent work [124] proposes a sparse signalling scheme combining pulse position modulation and multilevel coding, which is not entirely explicit. However, [222] showed that binary phase-shift keying (BPSK) with amplitude scaling as $O(n^{-1/4})$ is first-order optimal for covert communication, which allows the use of linear binary codes. The challenge for the Gaussian channel is that the rate of the code must tend to zero as the block-length tends to infinity, which is not the case for traditional linear codes. Thus, one needs to find new methods to analyse linear binary codes in the regime where the rate of the code vanishes. In the case of discrete memoryless channels, polar codes do not work well for covertness [87] essentially due to their suboptimal finite-length scaling [106], and we believe that a similar issue would occur over Gaussian channels. An intriguing question is whether Reed-Muller codes could be adapted to provide covertness and reliability thanks to their polarization properties [1] and whether they could be efficiently decoded in this zero-rate regime.

Secret key generation and fuzzy extractors

The distillation of symmetric keys from correlated observations of wireless channels is a promising technique for forthcoming 6G applications that require lightweight key agreement protocols, as a standalone solution or to complement existing cryptographic algorithms. In this context, I plan to continue investigating lattice-based secret key generation protocols, following the framework described in Section 4.3, in collaboration with Cong Ling at Imperial College London.

Practical coding schemes. An immediate step for future work is to turn our approach in [J13], which is based on the notion of flatness factor of a lattice, into a practical key generation scheme. A promising option is to instantiate the lattices using polar codes. *Polar lattices* [156] have been shown to be good for quantization, channel coding and secrecy. Their encoding and decoding complexity is quasi-linear in the blocklength n. In particular, they achieve the secrecy capacity of Gaussian wiretap channels [155] and exhibit a vanishing L^1 flatness factor. One open problem is how to implement the randomized rounding algorithm over a polar lattice.

More general models. Other aspects we may investigate include identifying whether is is possible to remove dithering and/or randomized quantization, characterizing the second-order asymptotics of our scheme [115], extending it to multi-terminal systems and vector Gaussian sources [226, 154], as well as considering its performance under bidirectional interactive communication. Another open problem is the generalization to other continuous source distributions, which is far from immediate since the flatness factor relies on the properties of Gaussian distributions.

Relation to fuzzy extractors. A related question is whether our secret key generation scheme can be modified to yield a *fuzzy extractor*, which would require redesigning the lattices with respect to other entropy measures. The notion of fuzzy extractor [68] was proposed to solve the problem of converting noisy physical measurements into uniformly random strings which can be reproduced reliably. Given a measurement, the fuzzy extractor outputs a secret key and public helper data. Instead of storing the key, a server can just store the public data, which is safe since the key remains close to uniform even given the helper data. Upon receiving another noisy measurement from the same source, the extractor can use the helper data to recover the key.

In the case of fuzzy extractors for continuous sources, which is our focus, the measurement needs to be discretized first, in such a way as to extract most of the entropy while reducing the noise. Note that in contrast to the discrete case, there exist no universal fuzzy extractors for continuous sources [219]. Nevertheless, we plan to investigate whether lattice-based techniques could be used to extract randomness from general continuous random variables that exhibit certain properties, such as circular symmetry or small tails.

Another aspect to consider is the robustness of the fuzzy extractors with respect to imperfect estimation of the source. In fact, in practice the source distribution has to be estimated empirically, and it is usually infeasible to build an accurate model of a high-entropy distribution just by sampling from it. In the discrete case, [90] defined a new notion of *fuzzy min-entropy* that corresponds to the maximal key size which can be extracted with perfect knowledge of the source, but also showed that when the distribution is uncertain, there exist families of distributions with positive fuzzy min-entropy such that no fuzzy extractor is secure for most distributions in the family. A similar negative result holds in the continuous case [89]. An open problem is to find suitable conditions for a family of distributions to admit a fuzzy extractor.

Variants of the smoothing parameter and their application to physical layer security and cryptography

As we have seen in Section 4.3, the notion of secrecy-good lattices based on the L^1 version of the flatness factor / smoothing parameter [48, 155, 65] results in optimal rates for secret key generation, and merits further investigation. As an immediate application, it should be possible to use L^1*-secrecy good wiretap codes* to remove the $\frac{1}{2}$-nat gap to the secrecy capacity in Theorem 4.18, which is associated to the use of the L^∞ flatness factor, and possibly achieve optimal error exponents.

In collaboration with Cong Ling, we are also studying other variants of the flatness factor based on *Rényi divergence*, which seem to have an interesting algebraic characterization. Rényi divergence has been used in lattice-based cryptography to obtain tighter security reductions and smaller parameters [190, 14]. In information theory, leakage measures based on Rényi divergence have been considered in [234].

6.2 Post-quantum cryptography

As shown in this manuscript, lattice-theoretic tools lie at the crossroads of communications and cryptography: the smoothing parameter can provide both computational hardness and information-theoretic security; discrete Gaussian sampling can be used to generate error distributions in cryptography, but also to achieve capacity over Gaussian channels; embedding techniques can be employed to solve the closest vector problem for MIMO decoding but also to attack lattice-based cryptosystems; the structure of unit groups can be exploited to decode algebraic space-time codes, but also for the cryptanalysis of structured lattices.

As a long-term perspective, I plan to explore these connections in view of the rise of post-quantum cryptography, which is of critical importance in order to face the disruptive consequences of quantum computing. While I'm relatively new to this area, I have gained some familiarity with lattice-based cryptosystems in the context of Charbel Saliba's thesis.

In the short term, I will pursue this study through the collaboration with Kévin Carrier at ETIS, by co-supervising Valérian Hatey's PhD thesis as part of the ANR JCJC DECODE project. This project aims to contribute to the cryptanalysis and calibration of post-quantum cryptosystems, and in particular of digital signature schemes[5]. It has the wider objective of transposing techniques for generic decoding and near collision search from codes to lattices and vice-versa. Such connections have proven fruitful in recent works: for instance, the smoothing parameter can be extended from lattices to codes [65]; statistical decoding can be seen as the code-based counterpart of dual lattice attacks [38]; LLL reduction can be adapted to codes [64]; a lattice version of the McEliece scheme, based on the "lattice isomorphism problem", has been proposed [70].

Analysis of dual lattice attacks The concrete security of lattice-based cryptosystems is measured by their resistence to known attacks. Typically, the strategies to solve the (decision or search) LWE problem fall into two general categories. Given LWE samples $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e})$, *primal attacks* aim at recovering $(\mathbf{s}, \mathbf{e})$ directly by constructing a uSVP instance from the LWE problem, by embedding the lattice spanned by $\mathbf{A}$ into a higher-dimensional lattice (similarly to the embedding technique described in Section 3.2.2). The uSVP solution can then be found using an efficient lattice reduction algorithm such as BKZ. On the other hand, *dual attacks* aim at solving the decision version of LWE by finding many dual vectors $\mathbf{v}$ such that $\mathbf{v}^T\mathbf{A} = \mathbf{0} \bmod q$, or, more generally, such that $\mathbf{v}^T\mathbf{A}$ is "short", and using them as distinguishers by running a hypothesis test on the distribution of $\langle \mathbf{v}, \mathbf{b} \rangle$, which should be uniformly distributed if $(\mathbf{A}, \mathbf{b})$ is a uniform sample, and approximately Gaussian-like if $(\mathbf{A}, \mathbf{b})$ is an LWE sample. Such lists of short vectors can be generated using the BKZ algorithm with an enumeration or sieving [6] subroutine.

In July 2022, NIST announced the first candidates for standardization[6], which include the KYBERKEM key encapsulation protocol. Currently, the concrete security estimates for KYBERKEM and other related cryptosystems such as SABER are being called into question due to improved dual lattice attacks in [104, 164], which employ faster distinguishers based on batching together many samples using a Fast Fourier Transform. However, there is no full consensus in the community about the scope of these dual attacks, which rely on some unproven heuristics, requiring further analysis [69]. We note that coding-theoretic techniques can be used in dual attacks to find short vectors: inspired by the coded-BKW method to find collisions [105], Carrier *et al.* [39] proposed to use non-binary polar codes for lossy source-coding to efficiently reduce the dimension of the search space.

Provably secure lattice-based protocols It must be mentioned that designing lattice-based protocols which are at the same time efficient and provably secure is still an open problem. In fact, the parameters chosen for the current lattice-based cryptosystems, which were selected mainly for efficiency purposes, do not satisfy the hypotheses of the worst-case to average-case reduction theorems, although they are heuristically assumed to be secure. This problem was studied in [94, 93] for the FRODOKEM protocol based on plain LWE, but the proposed parameters lead to an inefficient cryptosystem.

[5] A new call for signatures has been started by NIST in 2023, in order to bring more diversity to the existing candidates for standardization.
[6] https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8413-upd1.pdf

Computational hardness of lattice problems for structured lattices KYBERKEM and its companion lattice signature scheme DILITHIUM are based on MLWE, a structured variant of the LWE problem based on module lattices which enjoys a reduction to a worst-case instance of Module-SIVP, i.e. approximate SIVP on module lattices (see Section 4.4). Compared to plain LWE, MLWE allows for a more efficient implementation and shorter key sizes, which are deemed essential by NIST to achieve acceptable performance for widely used applications, such as the TLS Internet protocol.

However, in principle the additional algebraic structure of MLWE (and RLWE) might make these variants more vulnerable to attacks. Although currently there are no specific attacks directly targeting RLWE or MLWE, a recent series of works seems to suggest that some problems in algebraic number theory can be efficiently solved with quantum computers, calling into question the computational hardness of SIVP in ideal lattices.
Biasse and Song [26] exhibited a quantum algorithm for computing class groups and solving the principal ideal problem in number fields. Building on this result and exploiting the structure of the log-unit lattice, Cramer *et al.* showed that given an ideal lattice in a cyclotomic ring, there is a quantum algorithm which finds a short generator of the ideal in polynomial time [51, 52]. Nevertheless, these results do not immediately threaten the security of RLWE as they require additional hypotheses, such as the existence of an unusually short vector, or are limited to large approximation factors.

In [186], Pellet-Mary, Hanrot and Stehlé introduced the PHS algorithm, which solves SVP_γ on ideal lattices with $\gamma = 2^{O(\sqrt{n})}$ in quantum polynomial time, or for polynomial γ in time $2^{O(\sqrt{n})}$, at the cost of an exponential (offline) preprocessing phase. A "twisted" version of the PHS algorithm [24] seems to have much better approximation factors in practice. The main idea is that SIVP can be reduced to CVP in the log-S-unit lattice, which can be computed once and for all for each number field. However, both [186] and [24] rely on some unproven heuristics, and numerical simulations are feasible only for relatively small cyclotomic number fields.
These results seem to suggest that approx-SVP for ideal lattices may be easier to solve than RLWE[7] , since there are reductions from MLWE to RLWE [8]; on the other hand the reductions from Module-SIVP to MLWE are sharp for modules of rank greater than 2.

Other works have focused on efficiently retrieving short vectors in module lattices. In this direction, there has been progress in developing variants of the LLL algorithm for module lattices [145]. Like the previous works, this also relies on heuristics about the log-unit lattice which are difficult to prove. The results of this work seem to indicate that finding short vectors in module lattices of rank 2 could be easier than in general lattices of the same dimension, pointing out a potential weakness of cryptosystems based on module lattices.

[7] We also note that the recent work [132] argues that the worst-case to average-case reduction of Ideal-SVP to RLWE is not tight in finite dimension, since large constant factors are hidden in the proof of [161].

REFERENCES

[1] E. Abbe and M. Ye, "Reed-Muller codes polarize," *IEEE Transactions on Information Theory*, vol. 66, no. 12, pp. 7311–7332, 2020.

[2] E. Abbe and C. Sandon, "A proof that Reed-Muller codes achieve Shannon capacity on symmetric channels," *arXiv preprint arXiv:2304.02509*, 2023.

[3] C. Aguilar, P. Gaborit, P. Lacharme1, J. Schrek, and G. Zémor, "Noisy Diffie-Hellman protocols, Recent results session at Post-Quantum Cryptography-3rd international workshop, PQCrypto," 2010, https://www.yumpu.com/en/document/read/53051354/noisy-diffie-hellman-protocols.

[4] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[5] M. Ajtai, "Generating hard instances of lattice problems," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, 1996, pp. 99–108.

[6] M. Ajtai, R. Kumar, and D. Sivakumar, "A sieve algorithm for the shortest lattice vector problem," in *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, 2001, pp. 601–610.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] M. R. Albrecht and A. Deo, "Large modulus ring-LWE $\geq$ module-LWE," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2017, pp. 267–296.

[9] E. Alkim, L. Ducas, T. Pöppelmann, and P. Schwabe, "Post-quantum key exchange-a new hope." in *USENIX Security Symposium*, 2016.

[10] C. Alves and J.-C. Belfiore, "Lattices from maximal orders into quaternion algebras," *Journal of Pure and Applied Algebra*, vol. 219, no. 4, pp. 687–702, 2015.

[11] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[12] E. Arıkan, "Source polarization," in *IEEE International Symposium on Information Theory (ISIT)*, 2010, pp. 899–903.

[13] R. Avanzi *et al.*, "CRYSTALS-Kyber algorithm specifications and supporting documentation," *NIST PQC Round*, 2020, https://pq-crystals.org/kyber/data/kyber-specification-round3.pdf".

[14] S. Bai, T. Lepoint, A. Roux-Langlois, A. Sakzad, D. Stehlé, and R. Steinfeld, "Improved security proofs in lattice-based cryptography: using the Rényi divergence rather than the statistical distance," *Journal of Cryptology*, vol. 31, pp. 610–640, 2018.

[15] W. Banaszczyk, "New bounds in some transference theorems in the geometry of numbers," *Math. Ann.*, vol. 296, pp. 625–635, 1993.

[16] M. Bardet, V. Dragoi, A. Otmani, and J.-P. Tillich, "Algebraic properties of polar codes from a new polynomial formalism," in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 230–234.

[17] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE journal on selected areas in communications*, vol. 31, no. 9, pp. 1921–1930, 2013.

[18] J.-C. Belfiore and F. Oggier, "Lattice code design for the Rayleigh fading wiretap channel," in *IEEE International Conference on Communications (ICC)*, 2011.

[19] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden code: a 2×2 full-rate space-time code with nonvanishing

determinants," *IEEE Transactions on information theory*, vol. 51, no. 4, pp. 1432–1436, 2005.

[20] J.-C. Belfiore, "Lattice codes for the compute-and-forward protocol: The flatness factor," in *Proc. ITW 2011*, Paraty, Brazil, 2011.

[21] J.-C. Belfiore and F. Oggier, "Secrecy gain: A wiretap lattice code design," in *2010 International Symposium On Information Theory & Its Applications*, 2010, pp. 174–178.

[22] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology*, ser. Lecture Notes in Computer Science, vol. 7417. Springer-Verlag, 2012, pp. 294–311.

[23] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Transactions on Information Theory*, vol. 48, no. 10, pp. 2637–2655, 2002.

[24] O. Bernard and A. Roux-Langlois, "Twisted-PHS: Using the Product Formula to Solve Approx-SVP in Ideal Lattices," in *ASIACRYPT 2020*. Springer, 2020, pp. 349–380.

[25] K.-L. Besser, P.-H. Lin, C. R. Janda, and E. A. Jorswieck, "Wiretap code design by neural network autoencoders," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3374–3386, 2020.

[26] J.-F. Biasse and F. Song, "Efficient quantum algorithms for computing class groups and solving the principal ideal problem in arbitrary degree number fields," in *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 2016, pp. 893–902.

[27] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar codes for coordination in cascade networks," in *Proc. of International Zurich Seminar on Communications*, 2012, pp. 55–58.

[28] M. Bloch, "Channel intrinsic randomness," in *Proc. Int. Symp. Inform. Theory (ISIT 2010)*, June 2010, pp. 2607 –2611.

[29] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, Dec 2013.

[30] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[31] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[32] Z. Brakerski and V. Vaikuntanathan, "Fully homomorphic encryption from ring-LWE and security for key dependent messages," in *Annual cryptology conference*. Springer, 2011, pp. 505–524.

[33] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," *SIAM Journal on Computing*, vol. 43, no. 2, pp. 831–871, 2014.

[34] C. J. Bushnell and I. Reiner, "Solomon's conjectures and the local functional equation for zeta functions of orders," *Bulletin of the American Mathematical Society*, vol. 2, no. 2, pp. 306–310, 1980.

[35] A. Campello, D. Dadush, and C. Ling, "AWGN-goodness is enough: Capacity-achieving lattice codes based on dithered probabilistic shaping," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1961–1971, 2018.

[36] A. Campello, C. Ling, and J.-C. Belfiore, "Universal lattice codes for MIMO channels," *IEEE Transactions on Information Theory*, vol. 64, no. 12, pp. 7847–7865, 2018.

[37] A. Campello, C. Ling, and J.-C. Belfiore, "Semantically secure lattice codes for compound mimo channels," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1572–1584, 2019.

[38] K. Carrier, T. Debris-Alazard, C. Meyer-Hilfiger, and J.-P. Tillich, "Statistical decoding 2.0: reducing decoding to LPN," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2022, pp. 477–507.

[39] K. Carrier, Y. Shen, and J.-P. Tillich, "Faster dual lattice attacks by using coding theory," *Cryptology ePrint Archive*, 2022.

[40] G. Cervia, "Coordination of autonomous devices over noisy channels: capacity results and coding techniques," Ph.D. dissertation, Université Paris Seine, 2018. [Online]. Available: https://www.theses.fr/2018CERG0960

[41] G. Cervia, G. Bassi, and M. Skoglund, "Secure strong coordination," in *2020 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2020, pp. 1–6.

[42] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *2013 IEEE International Symposium on Information Theory*, 2013, pp. 2945–2949.

[43] L.-C. Choo, C. Ling, and K.-K. Wong, "Achievable rates for lattice coding over the Gaussian wiretap channel," in *ICC 2011 Physical Layer Security Workshop*, 2011.

[44] A. Chorti, A. N. Barreto, S. Köpsell, M. Zoli, M. Chafii, P. Sehier, G. Fettweis, and H. V. Poor, "Context-aware security for 6G wireless: The role of physical layer security," *IEEE Communications Standards Magazine*, vol. 6, no. 1, pp. 102–108, 2022.

[45] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages," in *IEEE Information Theory Workshop (ITW)*, 2015, pp. 1–5.

[46] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inform. Theory*, vol. 61, no. 11, pp. 6213–6237, 2015.

[47] R. A. Chou, M. R. Bloch, and J. Kliewer, "Empirical and strong coordination via soft covering with polar codes," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 5087–5100, 2018.

[48] K.-M. Chung, D. Dadush, F.-H. Liu, and C. Peikert, "On the lattice smoothing parameter problem," in *IEEE Conference on Computational Complexity*, 2013.

[49] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*. Springer Science & Business Media, 2013, vol. 290.

[50] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[51] R. Cramer, L. Ducas, C. Peikert, and O. Regev, "Recovering short generators of principal ideals in cyclotomic rings," in *Advances in Cryptology–EUROCRYPT*. Springer, 2016, pp. 559–585.

[52] R. Cramer, L. Ducas, and B. Wesolowski, "Mildly short vectors in cyclotomic ideal lattices in quantum polynomial time," *Journal of the ACM (JACM)*, vol. 68, no. 2, pp. 1–26, 2021.

[53] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, pp. 40–47, 1996.

[54] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[55] P. Cuff, "Communication requirements for generating correlated random variables," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, 2008, pp. 1393–1397.

[56] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4181–4206, 2010.

[57] P. Cuff, "Communication in networks for coordinating behavior," Ph.D. dissertation, Stanford University, 2009.

[58] P. Cuff, "Distributed Channel Synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7071–7096, Nov. 2013.

[59] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *Allerton Conference on Communication, Control and Computing*, 2011, pp. 235–239.

[60] D. Dadush and O. Regev, "Towards strong reverse Minkowski-type inequalities for lattices," in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2016, pp. 447–456.

[61] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Generalised sphere decoder for asymmetrical space-time communication architecture," *Electronics Letters*, vol. 36, no. 2, pp. 166–167, 2000.

[62] M. T. Damir, A. Karrila, L. Amoros, O. W. Gnilke, D. Karpuk, and C. Hollanti, "Well-rounded lattices: Towards optimal coset codes for Gaussian and fading wiretap channels," *IEEE Transactions on Information Theory*, vol. 67, no. 6, pp. 3645–3663, 2021.

[63] J.-P. D'Anvers, Q. Guo, T. Johansson, A. Nilsson, F. Vercauteren, and I. Verbauwhede, "Decryption failure attacks on IND-CCA secure lattice-based schemes," in *IACR International Workshop on Public Key Cryptography*. Springer, 2019, pp. 565–598.

[64] T. Debris-Alazard, L. Ducas, and W. P. van Woerden, "An algorithmic reduction theory for binary codes: Lll and more," *IEEE Transactions on Information Theory*, vol. 68, no. 5, pp. 3426–3444, 2022.

[65] T. Debris-Alazard, L. Ducas, N. Resch, and J.-P. Tillich, "Smoothing codes and lattices: Systematic study and new bounds," *IEEE Transactions on Information Theory*, vol. 69, no. 9, pp. 6006 – 6027, Sep. 2023.

[66] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[67] J. Ding, X. Xie, and X. Lin, "A simple provably secure key exchange scheme based on the learning with errors problem." *IACR Cryptol. ePrint Arch.*, vol. 2012, p. 688, 2012.

[68] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *Advances in Cryptology-EUROCRYPT 2004: International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004. Proceedings 23*. Springer, 2004, pp. 523–540.

[69] L. Ducas and L. Pulles, "Does the dual-sieve attack on learning with errors even work?" *Cryptology ePrint Archive*, 2023.

[70] L. Ducas and W. van Woerden, "On the lattice isomorphism problem, quadratic forms, remarkable lattices, and cryptography," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2022, pp. 643–673.

[71] W. Duke, Z. Rudnick, and P. Sarnak, "Density of integer points on affine homogeneous varieties," *Duke mathematical journal*, vol. 71, no. 1, pp. 143–179, 1993.

[72] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 618–629, 2016.

[73] A. Dytso, R. Bustin, H. V. Poor, and S. Shamai, "Analytical properties of generalized Gaussian distributions," *Journal of Statistical Distributions and Applications*, vol. 5, no. 1, pp. 1–40, 2018.

[74] A. Dytso, R. Bustin, H. V. Poor, and S. Shamai Shitz, "On additive channels with generalized Gaussian noise," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 426–430.

[75] J.-P. D'Anvers, F. Vercauteren, and I. Verbauwhede, "The impact of error dependencies on Ring/Mod-LWE/LWR based schemes," in *International Conference on Post-Quantum Cryptography*. Springer, 2019.

[76] A. Edelman and N. R. Rao, "Random matrix theory," *Acta Numerica*, vol. 14, pp. 233–297, 2005.

[77] M. Egan, "Isotropic and non-isotropic signaling in multivariate α-stable noise," *Frontiers in Communications and Networks*, vol. 2, p. 718945, 2021.

[78] K. Eisenträger and S. Hallgren, "Algorithms for ray class groups and Hilbert class fields," in *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 2010, pp. 471–483.

[79] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 968–985, Jun. 2004.

[80] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869–3884, Sept. 2006.

[81] D. Epstein and C. Petronio, "An exposition of Poincaré's Polyhedron Theorem," *L'Enseignement Mathématique*, no. 40, pp. 113–170, 1994.

[82] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct 2005.

[83] U. Erez and R. Zamir, "Achieving 1/2 log(1+SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[84] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, 2005.

[85] A.-M. Ernvall-Hytönen and C. Hollanti, "On the eavesdropper's correct decision in Gaussian and fading wiretap channels using lattice codes," in *2011 IEEE Information Theory Workshop*. IEEE, 2011, pp. 210–214.

[86] A. Fazeli, H. Hassani, M. Mondelli, and A. Vardy, "Binary linear codes with optimal scaling: Polar codes with large kernels," *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 5693–5710, 2020.

[87] G. Frèche, M. R. Bloch, and M. Barret, "Polar codes for covert communications over asynchronous discrete memoryless channels," *Entropy*, vol. 20, no. 1, p. 3, 2017.

[88] R. Fritschek, R. F. Schaefer, and G. Wunder, "Deep learning based wiretap coding via mutual information estimation," in *Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning*, 2020, pp. 74–79.

[89] B. Fuller and L. Peng, "Continuous-source fuzzy extractors: source uncertainty and insecurity," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 2952–2956.

[90] B. Fuller, L. Reyzin, and A. Smith, "When are fuzzy extractors possible?" *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5282–5298, 2020.

[91] R. G. Gallager, *Information theory and reliable communication*. Wiley, 1968.

[92] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2701–2710, 2009.

[93] J. Gärtner, "Concrete security from worst-case to average-case lattice reductions," *Cryptology ePrint Archive*, 2023.

[94] F. Gates, "Reduction-respecting parameters for lattice-based cryptosystems," Ph.D. dissertation, McMaster University, Ontario, Canada, 2018.

[95] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the forty-first annual ACM symposium on Theory of computing*, 2009, pp. 169–178.

[96] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*, 2008, pp. 197–206.

[97] S. Goldwasser and S. Micali, "Probabilistic encryption," *J. Comput. Syst. Sci.*, vol. 28, no. 2, pp. 270–299, 1984.

[98] E. S. Golod and I. R. Shafarevich, "On the class field tower," *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 28, no. 2, pp. 261–272, 1964.

[99] A. Gorodnik and B. Weiss, "Distribution of lattice orbits on homogeneous varieties," *GAFA Geometric And Functional Analysis*, vol. 17, no. 1, pp. 58–115, 2007.

[100] A. Gorodnik and A. Nevo, "Counting lattice points," *Journal für die reine und angewandte Mathematik*, vol. 2012, no. 663, pp. 127–176, 2012.

[101] A. Gorodnik, H. Oh, and N. Shah, "Strong wavefront lemma and counting lattice points in sectors," *Israel Journal of Mathematics*, vol. 176, no. 1, pp. 419–444, 2010.

[102] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, pp. 1603–1636, 2006.

[103] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*. Amsterdam: Elsevier, 1987.

[104] Q. Guo and T. Johansson, "Faster dual lattice attacks for solving LWE with applications to CRYSTALS," in *Advances in Cryptology–ASIACRYPT*. Springer, December 2021, pp. 33–62.

[105] Q. Guo, T. Johansson, and P. Stankovski, "Coded-BKW: Solving LWE using lattice codes," in *Annual Cryptology Conference*. Springer, 2015, pp. 23–42.

[106] V. Guruswami and P. Xia, "Polar codes: Speed of polarization and polynomial gap to capacity," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 3–16, 2014.

[107] F. Hajir and C. Maire, "A symptotically good towers of global fields," *PROGRESS IN MATHEMATICS-BOSTON-*, vol. 202, pp. 207–218, 2001.

[108] T. Han and S. Verdu, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[109] W. K. Harrison, "Exact equivocation expressions for wiretap coding over erasure channel models," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2687–2691, 2020.

[110] H. Hassani, S. Kudekar, O. Ordentlich, Y. Polyanskiy, and R. Urbanke, "Almost optimal scaling of Reed-Muller codes on BEC and BSC channels," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 311–315.

[111] S. H. Hassani and R. Urbanke, "Universal polar codes," in *IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 1451–1455.

[112] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans.*

Inf. Theory, vol. 57, no. 6, pp. 3989 –4001, Jun. 2011.

[113] M. Hayashi and R. Matsumoto, "Secure multiplex coding with dependent and non-uniform multiple messages," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2355–2409, 2016.

[114] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1562–1575, 2006.

[115] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3796–3810, 2016.

[116] X. He and A. Yener, "Providing secrecy with lattice codes," in *2008 46th annual Allerton conference on communication, control, and computing*. IEEE, 2008, pp. 1199–1206.

[117] X. He and A. Yener, "The Gaussian many-to-one interference channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2730–2745, 2011.

[118] X. He and A. Yener, "Strong secrecy and reliable Byzantine detection in the presence of an untrusted relay," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 177–192, 2012.

[119] B. M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[120] D. Hofheinz, K. Hövelmanns, and E. Kiltz, "A modular analysis of the Fujisaki-Okamoto transformation," in *Theory of Cryptography Conference*. Springer, 2017, pp. 341–371.

[121] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4765–4780, 2010.

[122] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1474–1484, 2005.

[123] J. Jaldén and P. Elia, "Sphere decoding complexity exponent for decoding full-rate codes over the quasi-static MIMO channel," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5785–5803, 2012.

[124] I. A. Kadampot, M. Tahmasbi, and M. R. Bloch, "Codes for covert communication over additive white Gaussian noise channels," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 977–981.

[125] I. A. Kadampot, M. Tahmasbi, and M. R. Bloch, "Multilevel-coded pulse-position modulation for covert communications over binary-input discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6001–6023, 2020.

[126] R. Kannan, "Minkowski's convex body theorem and integer programming," *Mathematics of Operations Research*, vol. 12, no. 3, pp. 415–440, 1987.

[127] D. Karpuk, A.-M. Ernvall-Hytönen, C. Hollanti, and E. Viterbo, "Probability estimates for fading and wiretap channels from ideal class zeta functions," *Adv. Math. Commun.*, vol. 9, no. 4, pp. 391–413, 2015.

[128] P. Klein, "Finding the closest lattice vector when it's unusually close," *Proc. ACM-SIAM Symposium on Discrete Algorithms*, pp. 937–941, 2000.

[129] E. Kleinert, *Units of classical orders: a survey*. L'Enseignement Math., 1994, vol. 40.

[130] E. Kleinert, *Units in skew fields*. Birkäuser, Basel, 2000.

[131] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B.-J. Kwak, "LDPC codes for the Gaussian wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 532–540, 2011.

[132] N. Koblitz, S. Samajder, P. Sarkar, and S. Singha, "Concrete analysis of approximate ideal-SIVP to decision ring-LWE reduction," *Cryptology ePrint Archive*, 2022.

[133] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.

[134] G. Kramer and S. A. Savari, "Communicating probability distributions," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 518–525, 2007.

[135] A. Langlois and D. Stehlé, "Worst-case to average-case reductions for module lattices," *Designs, Codes and Cryptography*, vol. 75, no. 3, pp. 565–599, 2015.

[136] A. Lapidoth and M. Wigger, "Conditional and relevant common information," in *Proc. of IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, 2016, pp. 1–5.

[137] B. Larrousse, S. Lasaulce, and M. R. Bloch, "Coordination in distributed networks via coded actions with application to power control," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3633–3654, 2018.

[138] B. Larrousse, S. Lasaulce, and M. Wigger, "Coordinating partially-informed agents over state-dependent networks," in *IEEE Information Theory Workshop (ITW)*, 2015.

[139] B. Larrousse, S. Lasaulce, and M. Wigger, "Coordination in state-dependent distributed networks: The two-agent case," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 979–983.

[140] M. Le Treust, "Correlation between channel state and information source with empirical coordination constraint," in *IEEE Information Theory Workshop (ITW)*, 2014, pp. 272–276.

[141] M. Le Treust, "Coding theorems for empirical coordination," ETIS (UMR 8051), Tech. Rep., 2015. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01865569v1

[142] M. Le Treust, "Empirical coordination with two-sided state information and correlated source and state," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 466–470.

[143] M. Le Treust, "Joint empirical coordination of source and channel," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5087–5114, 2017.

[144] M. Le Treust and T. Tomala, "Persuasion with limited communication capacity," *Journal of Economic Theory*, vol. 184, p. 104940, 2019.

[145] C. Lee, A. Pellet-Mary, D. Stehlé, and A. Wallet, "An LLL algorithm for module lattices," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2019, pp. 59–90.

[146] E. Lee *et al.*, "Modification of FrodoKEM using Gray and error-correcting codes," *IEEE Access*, vol. 7, 2019.

[147] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

[148] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Znnalen*, vol. 261, no. ARTICLE, pp. 515–534, 1982.

[149] Y. Liang, H. V. Poor, S. Shamai *et al.*, "Information theoretic security," *Foundations and Trends® in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2009.

[150] S.-C. Lin and C.-L. Lin, "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3293–3306, June 2014.

[151] R. Lindner and C. Peikert, "Better key sizes (and attacks) for LWE-based encryption," in *Cryptographers' Track at the RSA Conference*. Springer, 2011, pp. 319–339.

[152] C. Ling and J.-C. Belfiore, "Achieving AWGN channel capacity with lattice Gaussian coding," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 5918–5929, 2014.

[153] B. Linowitz, M. Satriano, and R. Vehkalahti, "A non-commutative analogue of the Odlyzko bounds and bounds on performance for space-time lattice codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1971–1984, Apr. 2015.

[154] J. Liu, P. Cuff, and S. Verdú, "Key capacity for product sources with application to stationary Gaussian processes," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 984–1005, 2016.

[155] L. Liu, Y. Yan, and C. Ling, "Achieving secrecy capacity of the Gaussian wiretap channel with polar lattices," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1647–1665, 2018.

[156] L. Liu, Y. Yan, C. Ling, and X. Wu, "Construction of capacity-achieving lattice codes: Polar lattices," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 915–928, 2019.

[157] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.

[158] H.-F. Lu, "Constructions of multiblock space-time coding schemes that achieve the diversity-multiplexing tradeoff," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3790–3796, Aug. 2008.

[159] X. Lu, Y. Liu, Z. Zhang, D. Jia, H. Xue, J. He, B. Li, K. Wang, Z. Liu, and H. Yang, "LAC: Practical ring-LWE based public-key encryption with byte-level modulus." *IACR Cryptol. ePrint Arch.*, p. 1009, 2018.

[160] V. Lyubashevsky and D. Micciancio, "On bounded distance decoding, unique shortest vectors, and the minimum distance problem," in *Crypto'09*, Aug. 2009, pp. 577–594.

[161] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors," *Journal ACM*, vol. 60, no. 6, Nov 2013.

[162] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, Oct 2011.

[163] J. Martinet, "Tours de corps de classes et estimations de discriminants," *Inventiones Mathematicae*, no. 44, pp. 65–73, 1978.

[164] MATZOV, "Report on the security of LWE: Improved dual lattice attack," The Center of Encryption and Information Security, Tech. Rep., April 2022. [Online]. Available: https://zenodo.org/record/6493704

[165] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[166] D. Micciancio and O. Regev, "Worst-case to average-case reductions based on Gaussian measures," in *Proc. Ann. Symp. Found. Computer Science*, Rome, Italy, Oct. 2004, pp. 372–381.

[167] D. Micciancio and M. Schultz, "Error correction and ciphertext quantization in lattice cryptography," *Cryptology ePrint Archive*, 2023.

[168] H. Mirghasemi and J. Belfiore, "The semantic secrecy rate of the lattice Gaussian coding for the Gaussian wiretap channel," in *IEEE Information Theory Workshop (ITW)*, Nov 2014, pp. 112–116.

[169] M. Mondelli, S. H. Hassani, I. Sason, and R. Urbanke, "Achieving Marton's region for broadcast channels using polar codes," *IEEE Transactions on Information Theory*, vol. 61, no. 2, pp. 783–800, 2015.

[170] L. Mucchi, S. Jayousi, S. Caputo, E. Panayirci, S. Shahabuddin, J. Bechtold, I. Morales, R.-A. Stoica, G. Abreu, and H. Haas, "Physical-layer security in 6g networks," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1901–1914, 2021.

[171] A. D. Murugan, H. El Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: Rediscovering the sequential decoder," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 933–953, 2006.

[172] M. Naehrig *et al.*, "FrodoKEM. tech. rep." in *National Institute of Standards and Technology*, 2020, https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[173] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.

[174] S. Nitinawarat and P. Narayan, "Secret key generation for correlated Gaussian sources," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3373 –3391, June 2012.

[175] A. Nooraiepour, S. R. Aghdam, and T. M. Duman, "On secure communications over Gaussian wiretap channels via finite-length codes," *IEEE Communications Letters*, vol. 24, no. 9, pp. 1904–1908, 2020.

[176] A. M. Odlyzko, "Bounds for discriminants and related estimates for class numbers, regulators and zeros of zeta functions: a survey of recent results," *Journal de théorie des nombres de Bordeaux*, vol. 2, no. 1, pp. 119–141, 1990.

[177] F. Oggier, P. Solé, and J.-C. Belfiore, "Lattice codes for the wiretap Gaussian channel: Construction and analysis," *IEEE Trans. Inform. Theory*, vol. 62, no. 10, pp. 5690–5708, Oct. 2016.

[178] F. E. Oggier, J.-C. Belfiore, and E. Viterbo, *Cyclic division algebras: A tool for space-time coding*. Foundations and Trends in Communications and Information Theory, 2007, vol. 4, no. 1.

[179] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space–time block codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3885–3902, Sept. 2006.

[180] O. Ordentlich and U. Erez, "Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 323–340, Jan. 2015.

[181] A. Page, "Computing fundamental domains for arithmetic kleinian groups," Master's thesis, Université Paris 7 - Diderot, 2010.

[182] C. Peikert, "An efficient and parallel Gaussian sampler for lattices," in *Proc. CRYPTO*, vol. 6223. Springer-Verlag, 2010, pp. 80–97.

[183] C. Peikert and A. Rosen, "Lattices that admit logarithmic worst-case to average-case connection factors," in *Proc. of the 39-th ACM Symp. on the Theory of Computing (STOC)*, 2007, pp. 478–487.

[184] C. Peikert, "Public-key cryptosystems from the worst-case shortest vector problem," in *Proceedings of the forty-first annual ACM symposium on Theory of computing*, 2009, pp. 333–342.

[185] C. Peikert, "Lattice cryptography for the internet," in *International Workshop on Post-Quantum Cryptography*. Springer, 2014, pp. 197–219.

[186] A. Pellet-Mary, G. Hanrot, and D. Stehlé, "Approx-SVP in ideal lattices with pre-processing," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2019, pp. 685–716.

[187] J. Pfister, M. A. Gomes, J. P. Vilela, and W. K. Harrison, "Quantifying equivocation for finite blocklength wiretap codes," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6.

[188] G. Poltyrev, "On coding without restrictions for the awgn channel," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 409–417, 1994.

[189] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, p. 2307, 2010.

[190] T. Prest, "Sharper bounds in lattice-based cryptography using the Rényi divergence," in *Advances in Cryptology–ASIACRYPT*. Springer, December 2017, pp. 347–374.

[191] V. Rana and R. A. Chou, "Short blocklength wiretap channel codes via deep learning: Design and performance evaluation," *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1462–1474, 2023.

[192] G. Reeves and H. D. Pfister, "Reed–Muller codes on BMS channels achieve vanishing bit-error probability for all rates below capacity," *IEEE Transactions on Information Theory*, 2023.

[193] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *J. ACM*, vol. 56, no. 6, pp. 34:1–34:40, 2009.

[194] I. Reiner, *Maximal Orders*. Academic Press, New York, 1975.

[195] G. Rekaya, J.-C. Belfiore, and E. Viterbo, "A very efficient lattice reduction tool on fast fading channels," in *Proceedings of ISITA*, vol. 2004, 2004.

[196] J. M. Renes, "Duality of channels and codes," *IEEE Transactions on Information Theory*, vol. 64, no. 1, pp. 577–592, 2018.

[197] C. Saliba, "Error correction and reconciliation techniques for lattice-based key generation protocols," Ph.D. dissertation, CY Cergy Paris Université, 2022.

[198] A. Sanderovich, M. Peleg, and S. Shamai, "LDPC coded MIMO multiple access with iterative joint decoding," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1437–1450, 2005.

[199] E. Şaşoğlu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *IEEE International Symposium on Information Theory*, 2013, pp. 1117–1121.

[200] E. Şaşoğlu, "Polar codes for discrete alphabets," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, 2012, pp. 2137–2141.

[201] S. Satpathy and P. Cuff, "Secure cascade channel synthesis," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6081–6094, Nov. 2016.

[202] O. Sberlo and A. Shpilka, "On the performance of Reed-Muller codes with respect to random errors and erasures," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2020, pp. 1357–1376.

[203] C.-P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," in *Fundamentals of Computation Theory: 8th International Conference, Gosen, Germany*. Springer, 1991, pp. 68–85.

[204] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[205] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[206] C. L. Siegel, "A mean value theorem in geometry of numbers," *Annals of Mathematics*, pp. 340–347, 1945.

[207] A. T. Suresh, A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy for erasure wiretap channels," in *2010 IEEE Information Theory Workshop*. IEEE, 2010, pp. 1–5.

[208] M. Taherzadeh and A. K. Khandani, "On the limitations of the naive lattice decoding," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4820–4826, 2010.

[209] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4801–4805, 2007.

[210] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6562–6582, 2013.

[211] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE transactions on information theory*, vol. 44, no. 2, pp. 744–765, 1998.

[212] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, July 2006.

[213] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[214] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, 2007.

[215] H. Tyagi and A. Vardy, "Explicit capacity-achieving coding scheme for the Gaussian wiretap channel," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 956–960.

[216] R. Vehkalahti, C. Hollanti, J. Lahtonen, and K. Ranto, "On the densest MIMO lattices from cyclic division algebras," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3751–3780, Aug. 2009.

[217] R. Vehkalahti, C. Hollanti, H.-F. Lu, and J. Lahtonen, "Some simple observations on MISO codes," in *Proc. 2010 IEEE Int. Symp. Inf. Theory and its Appl.*, Oct. 2010, pp. 537–541.

[218] R. Vehkalahti, C. Hollanti, and F. Oggier, "Fast-decodable asymmetric space-time codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2362–2384, Apr. 2012.

[219] E. A. Verbitskiy, P. Tuyls, C. Obi, B. Schoenmakers, and B. Skoric, "Key extraction from general nondiscrete signals," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 269–279, 2010.

[220] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639–1642, 1999.

[221] S. Vituri, "Dispersion analysis of infinite constellations in ergodic fading channels," Master's thesis, Tel Aviv University, 2013. [Online]. Available: http://arxiv.org/abs/1309.4638

[222] L. Wang, "On Gaussian covert communication in continuous time," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–10, 2019.

[223] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[224] Z. Wang and C. Ling, "On the geometric ergodicity of Metropolis-Hastings algorithms for lattice Gaussian sampling," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 738–751, Feb. 2018.

[225] S. Watanabe and Y. Oohama, "Secret key agreement from correlated Gaussian sources by rate limited public communication," *IEICE Trans. Fundamentals*, vol. E93-A, pp. 1976–1983, Nov. 2010.

[226] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited public communication," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 541–550, 2011.

[227] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[228] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, pp. 1355–1387, Oct. 1975.

[229] S. Yang and J. C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel," in *International Zurich Seminar on Communications*, 2006, pp. 122–125.

[230] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, 2019.

[231] H. Yao, A. Fazeli, and A. Vardy, "A deterministic algorithm for computing the weight distribution of polar code," *IEEE Transactions on Information Theory*, 2023.

[232] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability Proof via Output Statistics of Random Binning," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 6760–6786, Nov. 2014.

[233] M. Ye and E. Abbe, "Recursive projection-aggregation decoding of Reed-Muller codes," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 4948–4965, 2020.

[234] L. Yu and V. Y. Tan, "Rényi resolvability and its applications to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1862–1897, 2018.

[235] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1250–1276, Jun. 2002.

[236] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.

[237] Q. Zhang, M. Bakshi, and S. Jaggi, "Computationally efficient deniable communication," in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 2234–2238.

[238] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

INDEX