UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET INFORMATIQUE

Thèse

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

Spécialité : Mathématiques Pures

présentée et soutenue par

Nicola DI PIETRO

le 31 janvier 2014

On Infinite and Finite Lattice Constellations for the Additive White Gaussian Noise Channel

Thèse dirigée par Gilles ZÉMOR et Joseph J. BOUTROS

JURY :

Erik AGRELL	Professeur, Chalmers University of Technology	Examinateur
Christine BACHOC	Professeur, Université de Bordeaux	Examinateur
Joseph J. BOUTROS	Professeur, Texas A&M University at Qatar	Co-directeur
Loïc BRUNEL	Docteur, Mitsubishi Electric R&D Centre Europe	Examinateur
Uri EREZ	Professeur, Tel Aviv University	Rapporteur
Damien STEHLÉ	Professeur, École Normale Supérieure de Lyon	Rapporteur
Gilles ZÉMOR	Professeur, Université de Bordeaux	Directeur

Abstract

The problem of transmission of information over the AWGN channel using lattices is addressed. Firstly, infinite constellations are considered. A new family of integer lattices built by means of Construction A with non-binary linear codes is introduced. These lattices are called LDA (*Low-Density Construction A*) and are characterised by sparse p-ary parity-check matrices, that put them in direct relation with LDPC codes. Two results about the Poltyrev-capacity-achieving qualities of this family are proved, respectively for logarithmic row degree and constant row degree of the associated parity-check matrices. The second result is based on some expansion properties of the Tanner graphs related to these matrices. Another topic of this work concerns finite lattice constellations. A new proof that general random Construction A lattices achieve capacity under lattice decoding is provided, continuing and improving the work of Erez and Zamir (2004), Ordentlich and Erez (2012), and Ling and Belfiore (2013). This proof is based on Voronoi lattice constellations and MMSE scaling of the channel output. Finally, this approach is adapted to the LDA case and it is shown that LDA lattices achieve capacity with the same transmission scheme, too. Once again, it is necessary to exploit the expansion properties of the Tanner graphs. At the end of the dissertation, an iterative message-passing algorithm suitable for decoding LDA lattices in high dimensions is presented.

Résumé

On étudie le problème de la transmission de l'information à travers le canal AWGN en utilisant des réseaux. On commence par considérer des constellations infinies. Une nouvelle famille de réseaux obtenus par Construction A à partir de codes linéaires non binaires est proposée. Ces réseaux sont appelés LDA (« *Low-Density Construction A* ») et sont caractérisés par des matrices de parité p-aires creuses, qui les mettent en rélation directe avec les codes LDPC. Deux résultats sur leur possibilité d'atteindre la capacité de Poltyrev sont prouvés ; cela est d'abord démontré pour des poids des lignes logarithmiques des matrices de parité associées, puis pour des poids constants. Le deuxième résultat est basé sur certaines propriétés d'expansion des graphes de Tanner correspondants à ces matrices. Un autre sujet de ce travail concerne les constellations finies de réseaux. Une nouvelle preuve est donnée du fait que des réseaux aléatoires obtenus par Construction A générale atteignent la capacité avec décodage de type « lattice decoding ». Cela prolonge et améliore le travail de Erez et Zamir (2004), Ordentlich et Erez (2012), Ling et Belfiore (2013). Cette preuve est basée sur les constellations de Voronoï et la multiplication par le coefficient de Wiener (« MMSE scaling ») du signal en sortie du canal. Finalement, ce résultat est adapté au cas des réseaux LDA, qui eux aussi atteignent la capacité avec le même procédé de transmission. Encore une fois, il est nécessaire d'exploiter les propriétés d'expansion des graphes de Tanner. À la fin de la dissertation, on présente un algorithme de décodage itératif et de type « message-passing » approprié au décodage des LDA en grandes dimensions.

Institut de Mathématiques de Bordeaux
Université de Bordeaux
351, cours de la Libération - F 33405 TALENCE cedex

Acknowledgements

This thesis does not have just an academical value, but it represents the synthesis of four years and a half of work and life in Bordeaux. Many people have contributed to it, someone more directly, someone else less consciously, but still in an important way. Hopefully I will be able to thank them all here, without forgetting anyone.

First of all, I would like to express my gratitude to the members of the jury. I would like to thank the two reviewers, Uri Erez and Damien Stehlé, for having taken on the onerous task of reading this work in detail. I have deeply appreciated their comments and suggestions, without which this thesis would have been less worthy. I am grateful to Christine Bachoc and Erik Agrell, too, who honoured me with their presence at my defense. Last, but far from being least, I express my deepest thankfulness to Gilles Zémor, Loïc Brunel and Joseph Boutros. Thank you, Gilles, for your constant patience and availability and for the countless hours spent assisting me in my work. Thank you, Loïc, for your kindness and the interest that you have always put in listening to my boring mathematical explanations. Thank you, Joseph, for the time that you have devoted to me and our research and for having made me feel at home in Doha. These words will never be enough to say how important the role of supervisors and the human qualities of you all have been to me.

Doing research for three years would not have been possible for me without Mitsubishi Electric. I am grateful to the company for having supported this project and to all my colleagues that have shared with me my days in Rennes. I thank all of them for their kindness and for having always aimed to be more than simple coworkers.

The long days at the IMB would not have been so pleasant without all the friends that I have met there: my first tought goes to Nicolas and Aurélien, who received me as a friend since my first day in our office and with whom I have spent two wonderful years. I will always remember with pleasure all the Ph.D. students encoutered so far, with whom I have shared many nice moments: Alan, Alberto, Alice, Arthur, Aurel, Bruno, Clément, Diomba, Enea, Francesco, Frédéric, Guhan, Jean-Baptiste, Jean-Mathieu, Jocelyn, Louis, Marie, Miguel, Nicolas D., Nicolas M., Pierre C., Pierre L., René, Sagnik, Sophie, Stéphanie, Samuel, Soline, Zoé... I wish you all to succeed in your mathematical careers.

In September 2009, when I left Padova to Bordeaux with Andrea, Alberto and Giovanni, I did not know how meaningful and profound our friendship was going to become. I consider them more than friends. Thank you, brothers, for having been

my French family. I think with gratitude also to all the other friends with whom I have shared the same sense of being a family at home: Dario, Diego, Samuele, Eleonora and Agostino. I feel myself lucky of having had the chance of meeting you every evening, after a tough day of work. I am very grateful to our "cinquième coloc" Mélissa; I owe her much more than I can express here in a few lines.

Life in Bordeaux would not have been the same without sharing with many friends the status of "Italians abroad". Alessio, Daniela, Daniele, Federica, Francesco, Gabriele, Giordano, Nicola, Paola and Dajano (and Danae!), thank you for all the moments spent together, maybe speaking about (and eating!) Italian food.

The tenderest thank is for Marina. Her presence, affection, and support have been more than precious during these months.

Of course, a friendly thought goes also to all the other people that have made these years happy: Anthony, Corinne, Gemma, Ruth, all the people of TalEC (and especially Aurélie, Emmanuel, Eva, Franck, Henriette, Marjorie, Patrice et Xavier), all the lovely members of the group "Parole de Vie" (Anne, Marie, Marie-Anne, Mounette, Nathalie et Benoit, Odile, Patricia et Philippe, Roland, Violaine et Thomas) and all the friends in Toulouse (and especially Alma, Clément, Delphine, Fabio, Guillaume, Henryk, Jérôme, Lieva, Marc, Marjolaine, Pascal, Santiago, Simon, Waël and Yannick).

I would like to mention also all my friends in Italy, that I have always felt very close in spite of the geographical distance that separates us. I feel a special sense of gratefulness for Giacomo, Chiara, Giovanni, Maria and Marta, who have been so brave as to cross the Alps and come and see with their eyes that Bordeaux is a beautiful place.

Finally, all my love goes to my parents and my sister Giulia, who are essential in my life. Their constant support has always motivated me and has been of crucial importance in every moment.

"Pauca, sed matura."
Johann Carl Friedrich Gauss

A Mario di Pietro,
nonno e ingegnere.

Contents

List of Figures

Chapter 1

Introduction

English version

This thesis aims to give a contribution to the domain of information theory and coding theory, addressing the problem of communication over the Additive White Gaussian Noise (AWGN) channel. Coding over finite alphabets for this kind of channel has undergone a huge effort in order to achieve capacity with efficient decoding. While this quest is arguably reaching its conclusion, the similar problem for infinite alphabet coding has received much less attention and has been picking up momentum only recently. In this work, we will deal with lattices and lattice codes, which generalise linear codes over finite fields to the Euclidean space $\mathbb{R}^n$. Lattices are infinite and discrete sets of points, provided with some particular symmetric structure. They possess by themselves an intrinsic theoretical interest and purely mathematical lattice theory was born and developed long before its communication-related branch. The idea of employing lattices for error correction comes mostly from practical reasons, since lattice structures naturally arise in a number of contexts, like for example network communication systems with multiple transmitters and receivers, for which lattice coding schemes have sometimes even better performance than the more classical ones based on random coding. Lattice codes are also useful for other applications, like physical layer security and network coding.

The first notable work on the possibility of sending information with lattices over the AWGN channel with satisfactory performance is due to de Buda and dates back to 1975. More precisely, de Buda investigated the problem of achieving the capacity of the AWGN channel with finite lattices constellations and lattice decoding. This work has been the source of a research flow that has dealt with several different aspects of the problem of communicating through lattice codes:

- achieving the capacity of the AWGN channel with finite lattice constellations and optimal (Maximum Likelihood) decoding;

- achieving the capacity of the AWGN channel with finite lattice constellations and suboptimal decoding (namely, lattice decoding);

- efficiently communicating with infinite lattice constellations and the notion of Poltyrev capacity (a way of analysing the performance of a lattice family independently of the choice of the shaping region);
- finding implementable ML decoding algorithms for lattices (such as the Sphere Decoding algorithm);
- finding low-complexity, eventually suboptimal, decoding algorithms for lattices in high dimensions;
- constructively (and non-randomly) designing lattice families with good performance over the AWGN channel.

It took some years after de Buda's work before other main contributions to the topic appeared and the most consistent results were published in the last two decades. Once the theoretical problem of (non-constructively) achieving capacity with ML decoding was solved, it left the place to the challenge of obtaining the same result with lattice decoding. This turned out to be harder to prove than expected, insomuch as it was also conjectured to be impossible. Finally, less than 10 years ago, Erez and Zamir found the craved solution to the problem. In the mean time, since 2006, some constructive families of lattices adapted to iterative decoding have been proposed, with the intention of translating into concrete evidence the theoretical effort of showing that lattices are adequate to block coding in high dimensions for the AWGN channel.

At the root of this dissertation there is the desire to somehow close the gap between the path traced by Erez and Zamir and the experiments and implementations of well-performing lattice families with low-complexity decoding algorithms. At the present moment and to the best of our knowledge, there is no way of making the theoretical and practical points of view converge completely with finding a constructive family of lattices which is both proved to be capacity-achieving and concretely decodable in high dimensions. Nevertheless, at the end of this thesis, our main results will be of having shown the existence of a lattice family which is

- Poltyrev-capacity-achieving (cf. Theorem 3.1 and Theorem 3.2);
- Shannon-capacity-achieving under lattice decoding (cf. Theorem 4.2);
- decodable with low-complexity in high dimensions and with satisfactory performance under suboptimal, iterative decoding (see the numerical results of Chapter 5).

This is the family of *Low-Density Construction A* lattices, also called LDA lattices for brevity. To build it, we rely on an underlying finite-alphabet linear code structure. Moreover, we depart from the classical binary alphabet choice (common to almost all the constructions considered in the literature) and use codes over non-binary alphabets. Then, we shall call upon the celebrated Construction A technique to obtain lattices from those linear codes. Note that Construction A is not in itself a very restrictive scheme for lattice construction since it has been used, relying

on codes over large alphabets, to produce (non-constructively) capacity-achieving lattices. Construction A also yields some of the best asymptotic sphere-packing densities.

The other main ingredient of the LDA recipe are LDPC codes: indeed, our choice of linear codes for Construction A is restricted to them (after which the name "Low-Density" Construction A lattices). The two features, Construction A structure and LDPC skeleton, summarise the power of the LDA family. In a quite sketchy and approximate way, we can say that

- Construction A allows us to show that LDA lattices are capacity-achieving under lattice decoding, in the sense that non-LDPC Construction A lattices are already known to achieve the same result and this property is not lost when we move to a low-density underlying code structure;
- the LDPC foundation of the family makes it adaptable to iterative (hence low-complexity, practically implementable) decoding and suggests that good performance is obtainable.

Thus, the novelty of this work consists in finding strong information-theoretical and competitive numerical results for a "more constructive and less random" family than it is classically done.

Mathematically, this thesis contains a mix of Euclidean geometry, probability theory and combinatorics. It should be readable by quite a large audience, assuming that they have some familiarity with coding theory and information theory.

Structure of this dissertation

This thesis is organised as follows:

1. Chapter 2 will provide all the background which is needed to understand the sequel. We will start with a number of basic definitions about lattices, which will be mostly useful to fix some notation and to gently introduce the reader to the main objects of our dissertation. Then, we will move to the application of lattices to coding theory and communication over the AWGN channel. We will define with care Construction A and mention other constructions of lattices from linear codes, such as Construction B, C, D and D'. We will present in detail our channel model, without forgetting to give some further historical perspective and to duly present the problem of decoding infinite lattice constellations; as a consequence, we will give the definition of Poltyrev capacity, deeply used in Chapter 3. We will also spend some time talking about some classical coding-related problems involving lattices, for which Construction A provides a good solution. Finally, we will state (and where necessary prove) a certain number of useful lemmas that will be applied many times in the main proofs of Chapter 3 and Chapter 4.
2. Chapter 3 is dedicated to infinite LDA lattice constellations. We will show here the first two main results of this work: namely, that two particular families of

LDA lattices achieve Poltyrev capacity of the AWGN channel (cf. Theorem 3.1 and Theorem 3.2). These families are characterised by the parity-check matrices associated with the LDPC codes underlying Construction A. Theorem 3.1 shows how Poltyrev capacity can be achieved with LDA lattices the parity-check equations of which have degrees logarithmically growing in the dimension of the lattices, while Theorem 3.2 goes a little beyond this result, showing that Poltyrev capacity can be attained also by LDA lattices with constant parity-check equation degrees. Chapter 3 contains also some expansion results about the Tanner graph associated with our LDA lattices. These will be useful for the proofs both of Theorem 3.2 and of Theorem 4.2 in the following chapters.

3. Chapter 4 deals with finite lattice constellations and lattice decoding. It contains two main results: the first one is a proof that general random (hence non-LDA) Construction A lattices can achieve the capacity of the AWGN channel with lattice encoding and decoding (cf. Section 4.2 and Theorem 4.1). This result was actually already shown by other authors, but we give a new proof of it, which has some advantages with respect to the already existing ones. The second result is that LDA lattices can achieve the capacity of the AWGN channel with lattice encoding and decoding, too (cf. Section 4.3 and Theorem 4.2). Summarising, this result is based on Voronoi constellations, MMSE scaling of the AWGN channel output and the expansion properties of the Tanner graphs associated with LDA lattices.

4. Chapter 5 is devoted to the presentation of a practical, iterative decoding algorithm for LDA lattices. We will describe how to decode them and give some detail about the practical implementation. Finally, we will provide some numerical simulations that show the goodness of LDA lattices.

5. Chapter 6 will summarise the main achievements of our work and sketch some perspectives of future development.

   We have always tried to provide extensive and heuristic introductions to our theoretical results, in order to make the reading simple and all proof strategies clear. At this point, we simply wish you all to enjoy the reading.

Version française

Cette thèse se veut une contribution au domaine de la théorie de l'information et de la théorie du codage, étudiant le problème de la communication à travers le canal AWGN (c'est-à-dire soumis à un bruit gaussien blanc additif : « Additive White Gaussian Noise channel »). Dans le cas du codage à alphabets finis pour ce type de canal, il a fallu un effort important pour arriver à atteindre la capacité avec un décodage efficace. Alors que cette quête est probablement proche d'aboutir, le problème similaire pour le codage à alphabets infinis a été beaucoup moins étudié et n'est que depuis peu sujet à une attention croissante. Dans ce travail, nous nous préoccuperons de réseaux (« lattices ») et de constellations de réseaux (« lattice codes »), qui généralisent les codes linéaires sur les corps finis à l'espace euclidien $\mathbb{R}^n$. Les réseaux sont des ensembles infinis et discrets de points, équipés d'une structure symétrique particulière. Ils possèdent un intérêt théorique intrinsèque et la théorie des réseaux purement mathématique est née et s'est développée bien avant sa branche appliquée aux télécommunications. L'idée d'employer les réseaux pour la correction d'erreurs vient principalement de raisons pratiques, étant donné que la structure de réseau surgit naturellement dans plusieurs contextes. Cela se produit par exemple dans les systèmes de communication à émetteurs et récepteurs multiples, pour lesquels les procédés de codage basés sur les réseaux ont parfois des performances meilleures que ceux plus classiques basés sur le codage classique aléatoire. Le codage avec le réseaux est utile aussi pour d'autres applications, comme la sécurité de la couche physique et le « network coding ».

Le premier travail remarquable sur la possibilité d'envoyer de l'information à travers le canal AWGN avec les réseaux et des performances satisfaisantes est dû à de Buda et date de 1975. Plus précisément, de Buda a étudié le problème d'atteindre la capacité du canal AWGN avec des constellations finies de réseaux et un décodage de type « lattice decoding ». Ce travail a été la source d'un courant de recherche qui a traité plusieurs aspects du problème de la communication grâce aux codes de réseaux :

- atteindre la capacité du canal AWGN avec des constellations finies de réseaux et un décodage optimal (au maximum de vraisemblance – « ML decoding ») ;
- atteindre la capacité du canal AWGN avec des constellations finies de réseaux et un décodage sous-optimal (notamment, « lattice decoding ») ;
- communiquer efficacement avec des constellations infinies de réseaux et la notion de capacité de Poltyrev (une façon d'analyser les performances d'une famille de réseaux indépendamment du choix de la région de « shaping ») ;
- trouver des algorithmes implémentables de décodage au maximum de vraisemblance pour les réseaux (tels que l'algorithme « Sphere Decoding ») ;
- trouver des algorithmes de décodage pour les réseaux en grande dimension qui soient à faible complexité, même si sous-optimaux ;

- concevoir d'une façon constructive (et non pas aléatoire) des familles de réseaux qui aient de bonnes performances sur le canal AWGN.

Il a fallu quelques années après le travail de de Buda pour que d'autres contributions majeures apparaissent et les résultats les plus importants ont été publiés durant les deux dernières décennies. Une fois que le problème théorique d'atteindre (non constructivement) la capacité avec un décodage optimal fut résolu, il restait le défi d'obtenir le même résultat avec un décodage du type « lattice decoding ». La preuve de cela s'est avérée être plus difficile que prévu, à tel point que son impossibilité fut conjecturée aussi. Finalement, il y a moins de 10 ans, Erez et Zamir ont trouvé la solution si attendue de ce problème. Entre-temps, à partir de 2006, certaines familles constructives de réseaux adaptées au décodage itératif ont été proposées, avec l'intention de traduire en preuves concrètes la recherche théorique de réseaux adéquats au codage par blocs en grandes dimensions pour le canal AWGN.

À la racine de cette dissertation il y a le désir de rapprocher deux aspects : le chemin tracé par Erez et Zamir et les expériences et les implémentations des familles de réseaux qui ont de bonnes performances avec des algorithmes de décodage à faible complexité. À l'heure actuelle et à notre connaissance, il n'existe pas une façon de faire converger complètement les points de vue théorique et pratique. Cela signifierait trouver une famille constructive de réseaux qui atteigne mathématiquement la capacité et qui soit concrètement décodable en grande dimension. Cependant, à la fin de cette thèse, notre résultat principal consistera en avoir montré l'existence d'une famille de réseaux qui

- atteint la capacité de Poltyrev (voir le Théorème 3.1 et le Théorème 3.2) ;
- atteint la capacité de Shannon avec un décodage de type « lattice decoding » (voir le Théorème 4.2) ;
- est décodable en grandes dimensions avec performances satisfaisantes quand on utilise un décodeur itératif sous-optimal à faible complexité (voir les résultats numériques du Chapitre 5).

Il s'agit de la famille des réseaux LDA (*Low-Density Construction A*). Pour la construire, nous nous basons sur une structure de code linéaire sous-jacente aux réseaux. De plus, nous nous séparons du choix classique d'un alphabet binaire (commun à presque toutes les constructions considérées dans la littérature) et nous utilisons des codes sur un alphabet non binaire. Puis, nous invoquons la célèbre Construction A pour obtenir des réseaux à partir de ces codes linéaires. Il faut remarquer que la Construction A n'est pas un procédé très restrictif pour la construction de réseaux, car elle a été utilisée pour produire (non constructivement) des réseaux qui atteignent la capacité à partir de codes sur de grands alphabets. Grâce à la Construction A on obtient aussi certaines des meilleures densités asymptotiques d'empilement de sphères.

L'autre ingrédient principal de la recette LDA sont les codes LDPC (« *Low-Density Parity-Check* ») : en effet, notre choix de codes linéaires pour la Construction

A est restreint à ces derniers (d'ici le « Low-Density » qui fait partie du nom LDA). Les deux caractéristiques, Construction A et squelette LDPC, résument la puissance de la famille LDA. D'une manière sommaire, on peut dire que

- la Construction A nous permet de montrer que les réseaux LDA atteignent la capacité avec un décodage de type « lattice decoding », au sens qu'il est déjà connu que les réseaux construits par Construction A non LDPC ont la même propriété ; celle-ci n'est pas perdue lorsque nous nous concentrons sur une structure de base de type « low-density » ;

- les fondations LDPC de la famille la rendent adaptable à un décodage itératif (donc à faible complexité, pratiquement implémentable) et elles suggèrent que de bonnes performances peuvent être obtenues.

Ainsi, la nouveauté de ce travail consiste à produire des conclusions théoriques fortes du point de vue de la théorie de l'information et de résultats numériques compétitifs pour une famille « plus constructive et moins aléatoire » par rapport à ce qui est fait classiquement.

Mathématiquement, cette thèse contient un mélange de géométrie euclidienne, théorie des probabilités et calcul combinatoire. Elle devrait être lisible par un public plutôt étendu, pourvu qu'il possède une certaine familiarité avec la théorie du codage et de l'information.

Structure de cette dissertation

Cette thèse est organisée de la façon suivante :

1. Le Chapitre 2 fournit le contexte et les notions de base qui sont nécessaires pour comprendre la suite. Nous commencerons avec un certain nombre de définitions sur les réseaux, qui seront surtout utiles à fixer la notation et à introduire progressivement le lecteur aux objets principaux de notre dissertation. Puis, nous nous adresserons à l'application des réseaux à la théorie du codage et à la communication à travers le canal AWGN. Nous définirons avec soin la Construction A et mentionnerons d'autres constructions de réseaux à partir de codes linéaires, telles que la Construction B, C, D et D'. Nous présenterons en détail notre modèle de canal, sans oublier de donner d'autres informations historiques et de présenter dûment le problème du décodage de constellations infinies de réseaux ; par conséquent, nous donnerons la définition de la capacité de Poltyrev, utilisée amplement dans le Chapitre 3. Nous nous attarderons un peu sur certains problèmes classiques liés au codage qui impliquent les réseaux, pour lesquels la Construction A offre une bonne solution. Finalement, nous énoncerons (et prouverons, lorsque cela est nécessaire) un certain nombre de lemmes utiles qui seront appliqués plusieurs fois dans les preuves principales du Chapitre 3 et 4.

2. Le Chapitre 3 est dédié aux constellations infinies de réseaux LDA. Nous montrerons ici les premiers deux résultats principaux de ce travail : notamment,

que deux familles particulières de réseaux LDA atteignent la capacité de Poltyrev du canal AWGN (voir le Théorème 3.1 et le Théorème 3.2). Ces familles sont caractérisées par les matrices de parité associées aux codes LDPC sous-jacents la Construction A. Le Théorème 3.1 montre comment la capacité de Poltyrev peut être atteinte avec des réseaux LDA dont les équations de parité ont un degré qui grandit logarithmiquement en la dimension des réseaux. Le Théorème 3.2 va un peu au delà de ce résultat, en montrant que la capacité de Poltyrev peut être atteinte aussi par des réseaux LDA avec degré d'équations de parité constants. Le Chapitre 3 contient aussi des résultats d'expansion sur les graphes de Tanner associés à nos réseaux LDA. Ces propositions seront utiles pour les preuves du Théorème 3.2 et du Théorème 4.2 dans les chapitres successifs.

3. Le Chapitre 4 traite de constellations finies de réseaux et de décodage de type « lattice decoding ». Il contient deux résultats principaux : le premier est la preuve que les réseaux obtenus avec une Construction A aléatoire et générale (donc non LDA) peuvent atteindre la capacité du canal AWGN avec codage et décodage de type « lattice » (voir la Section 4.2 et le Théorème 4.1). Ce résultat a été déjà montré par d'autres auteurs, mais nous en donnons une nouvelle preuve, qui a certains avantages par rapport à celles déjà existantes. Le deuxième résultat est que les réseaux LDA eux aussi peuvent atteindre la capacité du canal AWGN avec le même type d'encodage et décodage que la Construction A générale (voir la Section 4.3 et le Théorème 4.2). Ce résultat est basé en résumé sur : les constellations de Voronoï, la multiplication par le coefficient de Wiener à la sortie du canal AWGN ainsi que les propriétés d'expansion des graphes de Tanner associés aux réseaux LDA.

4. Le Chapitre 5 est dédié à la présentation d'un algorithme de décodage pratique et itératif pour les réseaux LDA. Nous décrirons comment les décoder et nous donnerons quelques détails concernant l'implémentation pratique. Finalement, nous fournirons quelques simulations numériques qui montrent la qualité des réseaux LDA.

5. Le Chapitre 6 résumera les résultats principaux de notre travail et donnera une idée de développement futur.

   Nous avons toujours essayé d'offrir une introduction complète et heuristique aux résultats théoriques, afin de les simplifier et de rendre plus claires les stratégies de démonstration. Il ne nous reste que vous souhaiter une bonne lecture.

Chapter 2

Background on lattices

The purpose of this chapter is to introduce the reader to the main topics of this thesis. No mathematical novelty appears in this chapter and the experienced reader can move directly to Chapter 3. Nevertheless, here we fix most of the notation that will be employed later on. This chapter is structured in the following way:

- in Section 2.1, we quickly overview basic lattice definitions, with the main intention of giving some simple examples and getting used with some objects which will be useful in the sequel;
- in Section 2.2, we discuss some constructions of lattices from linear codes (with a particular emphasis on Construction A);
- in Section 2.3, we focus our attention on the application of lattices to the transmission of information, providing some history and a state of the art;
- in Section 2.4, we briefly describe four coding-related problems on lattices, for which a particular Construction A lattice family turns out to be a solution;
- in Section 2.5, we list some lemmas that will be useful tools in the next chapters.

2.1 Some basic definitions about lattices

The main object of this dissertation are *lattices*. This first section has the purpose of introducing them, together with basic definitions and some examples. These will be useful to recall the tools needed for reading this thesis and to fix the notation which will be used through all the chapters. A good reference for more detail on what follows is [CS99, Ebe13]. Mathematically, a lattice is defined as a module over a certain ring and embedded in a vector space over a field. For our purposes, we will only consider *real* lattices, that is $\mathbb{Z}$-modules in the Euclidean space. They will simply be discrete, additive subgroups of $\mathbb{R}^n$, according to the following definition:

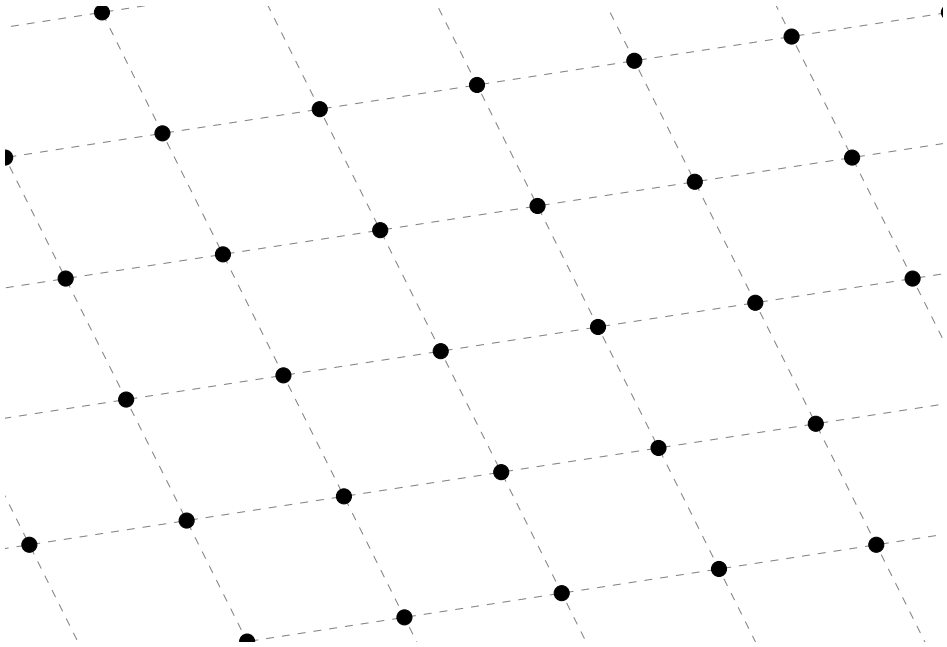

Figure 2.1: A lattice in $\mathbb{R}^2$.

Definition 2.1 (Lattice)**.** *Given m and n two natural numbers, $m \leq n$, and given a set of m linearly independent vectors $\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_m} \in \mathbb{R}^n$, an* m-dimensional lattice *Λ is defined as the set of all integer linear combinations of the $\mathbf{b_i}$'s:*

$$\Lambda = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i=1}^{m} z_i \mathbf{b_i},\ \exists (z_1, z_2, \ldots, z_m) \in \mathbb{Z}^n \right\}.$$

The $\mathbf{b_i}$'s are called a basis *of the lattice and we say that they* generate *it.*

A sublattice *Λ' of Λ is a lattice such that $\Lambda' \subseteq \Lambda$.*

An example of a 2-dimensional lattice in $\mathbb{R}^2$ is shown in Figure 2.1. By definition, the point $\mathbf{0} \in \mathbb{R}^n$ always belongs to a lattice. Furthermore, notice that the same lattice has many different bases and there is no natural choice for one of them. We will see in a while some parameters of lattices that are independent of the choice of the basis.

From now on, we will only deal with *full rank* lattices, that is n-dimensional lattices in an n-dimensional Euclidean space. Let us give some other definitions:

Definition 2.2 (Generator and Gram matrix)**.** *A* generator matrix *G of a lattice $\Lambda \subseteq \mathbb{R}^n$ is a matrix whose rows generate Λ:*

$$G = \begin{pmatrix} \mathbf{b_1} \\ \mathbf{b_2} \\ \vdots \\ \mathbf{b_n} \end{pmatrix} \quad \textit{and} \quad \Lambda = \{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{z}G,\ \exists \mathbf{z} \in \mathbb{Z}^n \}.$$

We have supposed that the lattice is full rank, so the $\mathbf{b_i}$*'s are a basis of the lattice, in the notation of Definition 2.1. The* Gram matrix *A of the lattice is given by*

$$A = GG^T.$$

Let $||\cdot||$ denote the usual Euclidean norm in $\mathbb{R}^n$, that is, if $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, $||\mathbf{y}|| = \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2} = \sqrt{\mathbf{y}\mathbf{y}^T}$. Let $\mathbf{x} = \mathbf{z}G$, for some $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{Z}^n$. Then

$$||\mathbf{x}||^2 = ||\mathbf{z}G||^2 = \mathbf{z}GG^T\mathbf{z}^T = \mathbf{z}A\mathbf{z}^T.$$

The function $f : \mathbb{Z}^n \to \mathbb{Z}^n, f(\mathbf{z}) = \mathbf{z}A\mathbf{z}^T$ is the *quadratic form associated with the lattice* Λ.

We will denote by $\mathrm{Vol}(\cdot)$ the usual Euclidean volume of a set and, for a set S and an element t, let $t + S = \{t + s : \exists s \in S\}$.

Definition 2.3 (Fundamental region)**.** *A set $F \subseteq \mathbb{R}^n$ is called a* fundamental region *for a lattice $\Lambda \subseteq \mathbb{R}^n$ if the following conditions are satisfied:*

1. $\mathbb{R}^n = \bigcup_{\mathbf{x}\in\Lambda}(\mathbf{x} + F)$ *(space covering).*
2. $\mathrm{Vol}\left((\mathbf{x_1} + F) \cap (\mathbf{x_2} + F)\right) = 0$*, for every $\mathbf{x_1}, \mathbf{x_2} \in \Lambda$, $\mathbf{x_1} \neq \mathbf{x_2}$ (no intersection of non-zero measure).*

This definition implies that every point $\mathbf{y}$ of the space can be written as $\mathbf{y} = \mathbf{x}+\mathbf{f}$, for some $\mathbf{x} \in \Lambda$ and some $\mathbf{f}$ belonging to a fundamental region F. Moreover, if $(\mathbf{x_1} + F) \cap (\mathbf{x_2} + F) = \emptyset$ for every $\mathbf{x_1}, \mathbf{x_2} \in \Lambda$, $\mathbf{x_1} \neq \mathbf{x_2}$, the decomposition $\mathbf{y} = \mathbf{x} + \mathbf{f}$ is unique. In this case, we say that the fundamental region is *proper*. If instead $(\mathbf{x_1} + F) \cap (\mathbf{x_2} + F) \neq \emptyset$, the intersection has to be contained in the boundary of the two sets. F can be made proper by appropriately removing a part of its boundary, still guaranteeing the space covering property. This leads to the definition of a *quantiser*:

Definition 2.4 (Quantiser)**.** *Given a lattice $\Lambda \subseteq \mathbb{R}^n$ and a proper fundamental region F for the lattice, the* quantiser with respect to Λ and F *is the function $Q_{\Lambda,F}(\cdot) : \mathbb{R}^n \to \Lambda$ that associates to a $\mathbf{y} \in \mathbb{R}^n$ the unique $\mathbf{x} \in \Lambda$ such that $\mathbf{y} = \mathbf{x} + \mathbf{f}$, for some $\mathbf{f} \in F$.*

Two classical examples of fundamental regions of a lattice are its *fundamental parallelotope* and its *Voronoi region*:

Definition 2.5 (Fundamental parallelotope)**.** *The* fundamental parallelotope *of a lattice $\Lambda \subseteq \mathbb{R}^n$ with basis $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_n}\}$ is the set*

$$P(B) = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i=1}^{n} a_i \mathbf{b_i},\ \exists (a_1, a_2, \ldots, a_n) \in [0, 1)^n \right\}.$$

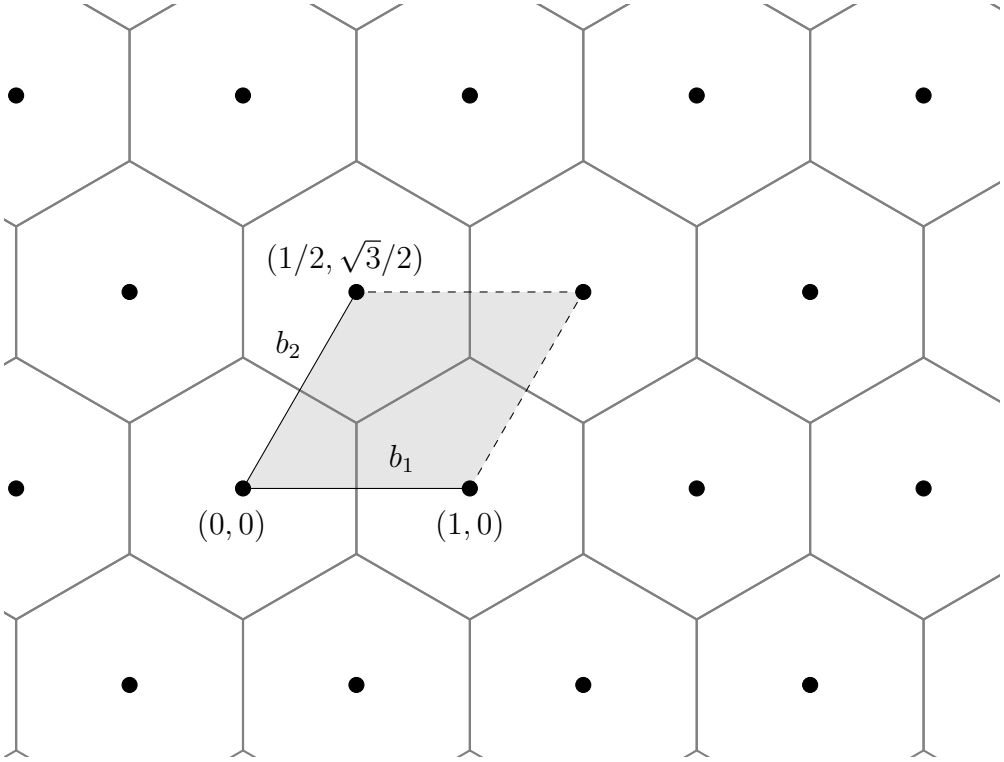

Figure 2.2: The hexagonal lattice $A_2 \subseteq \mathbb{R}^2$ is generated by the two basis vectors $b_1 = (1, 0)$ and $b_2 = (1/2, \sqrt{3}/2)$. It is named after its Voronoi regions, which have the shape of regular hexagons. The fundamental parallelotope associated with the previous basis is shaded in gray.

Definition 2.6 (Voronoi region)**.** *Given a lattice* $\Lambda \subseteq \mathbb{R}^n$ *and a point* $\mathbf{x} \in \Lambda$*, the* Voronoi region of $\mathbf{x}$ *is defined by*

$$\mathcal{V}(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : ||\mathbf{y} - \mathbf{x}|| \leq ||\mathbf{y} - \mathbf{z}||, \ \forall \mathbf{z} \in \Lambda, \mathbf{z} \neq \mathbf{x}\}.$$

Equivalently, $\mathcal{V}(\mathbf{x})$ *is the set of all the real vectors that are closer (or as close) to* $\mathbf{x}$ *than to any other lattice point. The Voronoi region of* $\mathbf{0}$ *is conventionally called* the Voronoi region *of the lattice and it is denoted by* $\mathcal{V}(\Lambda)$*.*

The fundamental parallelotope is a proper fundamental region, while we need to remove from the Voronoi region a part of its boundary to make it proper. The quantiser with respect to Λ and (the proper version of) $\mathcal{V}(\Lambda)$ is simply denoted by $Q_\Lambda(\cdot)$, the Voronoi region being the implicit fundamental region. Figure 2.2 represents a fundamental parallelotope and a Voronoi region for the hexagonal lattice $A_2 \subseteq \mathbb{R}^2$.

Note that, while the definition of fundamental parallelotope depends on the choice of the lattice basis, the Voronoi region does not. The two geometric properties listed in the definition and the lattice symmetry imply that any two fundamental regions have the same volume. In particular, $\mathrm{Vol}(\mathcal{V}(\Lambda)) = \mathrm{Vol}(P(B))$ for any basis

B and it is known that the volume of a parallelotope is given by the absolute value of the determinant of the matrix whose columns are the edges of the parallelotope that meet at any fixed vertex. Explicitly, $\mathrm{Vol}(\mathcal{V}(\Lambda)) = \mathrm{Vol}(P(B)) = |\det(G)|$, where G is the generator matrix of the lattice corresponding to B. This proves that $|\det(G)|$ is a constant for Λ, independently of the choice of the basis, and therefore justifies the following definition:

Definition 2.7 (Volume and determinant)**.** *The* volume *of a lattice* Λ *with generator matrix* G *is defined by*

$$\mathrm{Vol}(\Lambda) = |\det(G)|.$$

If $A = GG^T$ *is the lattice Gram matrix, the* determinant *of the lattice is*

$$\det(\Lambda) = \det(A) = \mathrm{Vol}(\Lambda)^2.$$

We can give another proof of the fact that the two definitions above are independent of the choice of the lattice basis: suppose that $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_n}\}$ and $B' = \{\mathbf{b_1}', \mathbf{b_2}', \ldots, \mathbf{b_n}'\}$ are two different basis of the same lattice Λ, with G and G' the generator matrices associated with B and B' respectively. Then, for every $i \in \{1, 2, \ldots, n\}$ we can write $\mathbf{b_i}' = \mathbf{q_i}G$, for some $\mathbf{q_i} \in \mathbb{Z}^n$. Therefore, if Q is the matrix whose rows are the $\mathbf{q_i}$'s, $G' = QG$ and Q has to be invertible within the matrices with integer coefficients. Equivalently, $\det(Q) = \pm 1$ and $|\det(G')| = |\det(Q) \cdot \det(G)| = |\det(G)|$.

Another parameter of a lattice which deserves investigation is the norm of the smallest non-zero vector; since a lattice is symmetrical by lattice point translations, this also coincides with the minimum distance between any two lattice points.

Definition 2.8 (Minimum distance)**.** *The* minimum distance *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is*

$$\mathrm{d_{min}}(\Lambda) = \min_{\mathbf{x} \in \Lambda} ||\mathbf{x}||.$$

Anticipating definitions and settings that we will deal with later, we remark that this quantity can be guessed to be closely related to the performance of a lattice for the transmission of information over a real channel with Gaussian noise. For this reason, it is interesting to define a way of comparing the minimum distances of two lattices, even when they have different volumes (a priori, a greater volume may trivially imply a greater minimum distance). This motivates the following definition:

Definition 2.9 (Fundamental gain)**.** *The* fundamental gain *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is*

$$\gamma(\Lambda) = \frac{\mathrm{d_{min}}(\Lambda)^2}{\mathrm{Vol}(\Lambda)^{\frac{2}{n}}}.$$

It is also known as the Hermite constant *of the lattice.*

Suppose that we compare two different lattices. For a fixed volume, the lattice with bigger fundamental gain will have a tendency to withstand a stronger noise over a Gaussian channel (under ML decoding, see 2.3.1); similarly, for a fixed minimum

distance, the lattice with bigger fundamental gain will be the one for which the smaller amount of energy is needed to send the same amount of information. This explains why the fundamental gain is used as a lattice performance estimator, when finer approaches are impracticable.

Another parameter which deserves to be mentioned is the *kissing number*; it simply counts the number of lattice points whose norm is the minimum distance of the lattice:

Definition 2.10 (Kissing number)**.** *The* kissing number *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is defined as*

$$\tau(\Lambda) = |\{\mathbf{x} \in \Lambda : ||\mathbf{x}|| = \mathrm{d}_{\min}(\Lambda)\}|.$$

One may also be interested in counting the number of lattice points that have any fixed norm and not only the minimum one. These numbers are collected by the so-called *theta series*:

Definition 2.11 (Theta series)**.** *Let* $q = e^{\pi i z}$ *for some* $z \in \mathbb{C}$ *with* $\mathrm{Im}(z) \geq 0$*; let* N_m *be the number of points of a certain lattice* $\Lambda \subseteq \mathbb{R}^n$ *whose squared Euclidean norm is* m*. Then the* theta series *of* Λ *is defined by*

$$\Theta_\Lambda(z) = \sum_{\mathbf{x} \in \Lambda} q^{\mathbf{x}\mathbf{x}^T} = \sum_{m=0}^{\infty} N_m q^m.$$

Observe that, by the previous definition, we have $\tau(\Lambda) = N_{\mathrm{d}_{\min}}(\Lambda)$.

2.2 Construction of lattices from codes

Lattices can be seen as the generalisation of linear codes over a finite field (Hamming space) to the Euclidean space. In this perspective, we will present some classical ways of obtaining lattices from linear codes. These strategies are very commonly employed in the literature for the achievement of both theoretical and practical results [Loe97, EZ04, ELZ05, GZ07, OE12, SBP06, BC08, SSP12, dPBZB12, dPBZB13, dPBZ13].

Definition 2.12 (Linear code)**.** *Let* $\mathbb{F}_q$ *be a finite field of cardinality* q*. A* linear code $C \subseteq \mathbb{F}_q^n$ *is simply a linear space of* $\mathbb{F}_q^n$*, seen as a vector space over* $\mathbb{F}_q$*. We denote by* $C[n,k]_q$ *a linear code of* $\mathbb{F}_q^n$ *of dimension* k *over* $\mathbb{F}_q$ *and we write* $R = k/n$ *for the* rate *of the code.*

All along this thesis, we will only deal with linear codes over $\mathbb{F}_p$, with p a prime number, not necessarily equal to 2. In the theoretical results of Chapter 3 and Chapter 4, the size of p will tend to infinity with the space dimension n.

2.2.1 Construction A

For our purposes, *Construction A* is the most interesting construction that we present. All the results of the next chapters concern lattices built in this way. We

will start with a quite general definition of Construction A, actually more general than what we will really need. Anyway, this is probably the nicest mathematical approach and it is worth mentioning. Then we will specify the two Construction A approaches that we will use in the rest of the thesis, namely Construction A over $\mathbb{Z}$ and over $\mathbb{Z}[i]$ (the Gaussian integers).

So, let L be a lattice of small dimension m and let L' be a sublattice of L such that the quotient L/L' is finite of *prime* cardinality p. The additive group L/L' injects naturally into the finite field $\mathbb{F}_p$ through an additive group isomorphism, and we assume identification of the two Abelian groups. We can define a lattice Λ of dimension $n = m\ell$ in the following way, which is the general setting for Construction A in Conway and Sloane's terminology [CS99]:

Definition 2.13 (Construction A)**.** *Let $C = C[\ell, k]_p$ be a linear code over $\mathbb{F}_p$ of length ℓ, dimension k, and* rate *$R = k/\ell$ and let $\Pi : L^\ell \to (L/L')^\ell$ be the natural projection. The lattice Λ obtained by* Construction A *is defined as:*

$$\Lambda = \{\mathbf{x} \in L^\ell : \Pi(\mathbf{x}) \in C\}.$$

As anticipated before, in the following chapters we shall mainly focus on two simple cases, namely:

1. $L = \mathbb{Z}$ and $L' = p\mathbb{Z}$.
2. $L = \mathbb{Z}[i]$ and $L' = \phi\mathbb{Z}[i]$, where $(\phi) = (a + bi)$ is a prime ideal of $\mathbb{Z}[i]$ of norm $a^2 + b^2 = p$.

The first case is one of the more classical forms of Construction A, for which we will derive some theoretical results in Chapter 3 and Chapter 4. Numerical experiments of the Chapter 5 will concern both the first and the second construction.

When $L = \mathbb{Z}$ and $L' = p\mathbb{Z}$, then $m = 1, \ell = n$ and, by definition, the lattice Λ is

$$\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x} \bmod p \in C\}.$$

Moreover, let $\Phi : \mathbb{F}_p \to \mathbb{Z}$ be the natural embedding of $\mathbb{F}_p$ into $\mathbb{Z}$, typically with $\Phi(\mathbb{F}_p) = \{-(p-1)/2, \ldots, (p-1)/2\}$, then another way of describing Λ is

$$\Lambda = \Phi(C) + p\mathbb{Z}^n = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x} = \Phi(\mathbf{c}) + p\mathbf{z}, \exists \mathbf{c} \in C, \mathbf{z} \in \mathbb{Z}^n\}. \tag{2.1}$$

To lighten notation, later we might simply write $\Lambda = C + p\mathbb{Z}^n$, identifying $\Phi(C)$ and C itself, since the embedding is trivial.

Notice that all the points of a lattice built in this way have integer coordinates; moreover, $p\mathbb{Z}^n$ is always contained in Λ:

$$p\mathbb{Z}^n \subseteq \Lambda \subseteq \mathbb{Z}^n.$$

This also implies that, if d_C is the minimum Euclidean distance of the linear code C (that is, if $d_c = \min_{\mathbf{c} \in C} ||\Phi(\mathbf{c})||$), then

$$\mathrm{d}_{\min}(\Lambda) = \min\{p, d_C\}.$$

Finally, if H is the parity-check matrix of the code C, another way of defining Λ is

$$\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : H\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\}. \tag{2.2}$$

If $(\mathrm{Id}_k \ B)$ is a $k \times \ell$ generator matrix in systematic form for the code $C[n,k]_p$, then it is easy to show that a generator matrix for Λ is

$$G = \begin{pmatrix} \mathrm{Id}_k & \Phi(B) \\ 0 & p\,\mathrm{Id}_{n-k} \end{pmatrix}$$

This straightforwardly implies that, if $R = k/n$,

$$\mathrm{Vol}(\Lambda) = p^{n-k} = p^{n(1-R)}. \tag{2.3}$$

In the second, Gaussian integer case, $m = 2$ and C has length $\ell = n/2$. Therefore Λ can be seen as a $\mathbb{Z}[i]$-module generated by the $\ell \times \ell$ matrix

$$G' = \begin{pmatrix} \mathrm{Id}_k & \Phi(B) \\ 0 & \phi\,\mathrm{Id}_{\ell-k} \end{pmatrix}$$

where Φ is now an embedding of $\mathbb{F}_p$ into a suitable region of $\mathbb{Z}[i]$ via the isomorphism $\mathbb{F}_p \simeq \mathbb{Z}[i]/(\phi)$: in other words, $\Phi(\mathbb{F}_p)$ is a set of representatives for $\mathbb{Z}[i]/(\phi)$. To obtain a generator matrix for the real lattice Λ of dimension $n = 2\ell$, we simply apply the transformation $x + iy \mapsto \left(\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right)$ to every coordinate of G'. In this case, if $R = k/n = k/2\ell$, we obtain

$$\mathrm{Vol}(\Lambda) = p^{\ell-k} = p^{\frac{1}{2}n(1-R)}.$$

Figure 2.3 shows a suitable representation $\Phi(\mathbb{F}_{41})$ of $\mathbb{Z}[i]/(4+5i)$ as a constellation of $\mathbb{Z}[i] \simeq \mathbb{Z}^2$: a particular family of codes over $\mathbb{Z}[i]/(4+5i)$ and their associated lattices will be experimented with in Chapter 5.

We will see in the next chapters in more detail that it can be very fruitful to build lattices from linear codes with Construction A. As far as we are concerned, this is mainly for two reasons:

- because their structure allows a theoretical analysis, inspired by what is already known about linear codes (see for example [Loe97, EZ04, ELZ05, GZ07, OE12, dPBZB13, dPBZ13]);
- because it is possible to take advantage of practical implementations of good linear codes and generalise them to lattices: see the experimental results of [dPBZB12, SS13] and Chapter 5.

2.2.2 Construction D and D'

Construction D and *D'* involve chains of nested binary linear codes and are employed to build lattices with a low-complexity iterative decoding algorithm (Low-Density Parity-Check (LDPC) Lattices and Turbo Lattices). We refer the reader to [SBP06, BC08, SSP12, SS13] to find all information about these lattice families and here we stick to the simple mathematical presentation of the two constructions. We will not give much information, but just introduce them as a complement to what we have already said on Construction A.

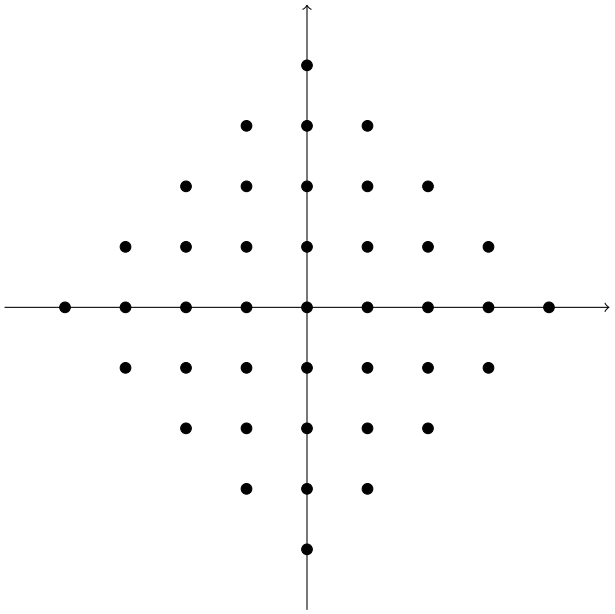

Figure 2.3: A system of representatives for $\mathbb{Z}[i]/(4+5i) \subseteq \mathbb{Z}^2$.

Construction D

Construction D generalises Construction A over $\mathbb{Z}$ with $p = 2$. We keep the notation of [SSP12] for our description, nevertheless, when compared to that paper, our version of the construction will come out scaled; this is because we prefer to have lattices contained in $\mathbb{Z}^n$. A good reference for Construction D is also [CS99].

Definition 2.14 (Construction D)**.** *Consider the chain of nested binary linear codes*

$$C_a \subseteq C_{a-1} \subseteq \ldots \subseteq C_1 \subseteq C_0,$$

where $C_\ell = C_\ell[n, k_\ell]_2$ for $\ell = 0, 1, \ldots, a$ (cf. Definition 2.12) and $C_0 = C_0[n, n]_2$. Denote by $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k_\ell}$ the k_ℓ vectors of $\mathbb{F}_2^n$ that generate the ℓ-th code.

We say that a lattice $\Lambda \subseteq \mathbb{R}^n$ is obtained by Construction D (with $a+1$ levels) *when*

$$\Lambda = \left\{ \mathbf{x} \in \mathbb{Z}^n : \mathbf{x} = \mathbf{z} + \sum_{\ell=1}^{a} \sum_{j=1}^{k_\ell} \beta_j^{(\ell)} 2^{a-\ell} \mathbf{c}_j, \exists \mathbf{z} \in 2^a \mathbb{Z}^n, \beta_j^{(\ell)} \in \{0,1\} \right\}.$$

Let $*$ denote the coordinate-wise product between any two binary codewords $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ and $\mathbf{c}' = (c'_1, c'_2, \ldots, c'_n)$:

$$\mathbf{c} * \mathbf{c}' = (c_1 \cdot c'_1, c_2 \cdot c'_2, \ldots, c_n \cdot c'_n),$$

where $\cdot$ indicates the usual product operation of $\mathbb{F}_2$. If we suppose that for every $\ell < a$,

$$\{\mathbf{c} * \mathbf{c}' : \mathbf{c}, \mathbf{c}' \in C_{\ell+1}\} \subseteq C_\ell.$$

then we can give a more compact formula to define the previously defined Λ:

$$\Lambda = C_a + 2C_{a-1} + \ldots + 2^{a-2}C_2 + 2^{a-1}C_1 + 2^a\mathbb{Z}^n. \tag{2.4}$$

Note that Construction A over $\mathbb{Z}$ with $p = 2$ is a particular case of Construction D with $a = 1$ (compare (2.1) and (2.4); above here we have omitted to specify the trivial embedding Φ of the codes in $\mathbb{Z}^n$).

Construction D'

Construction D' is dual to Construction D. To describe it, we will mostly use the same notation of [SBP06]; another reference to learn about it is [CS99]. Let $\alpha \in \{1, 2\}$ and consider, as for Construction D, the chain of nested binary linear codes

$$C_a \subseteq C_{a-1} \subseteq \ldots \subseteq C_1 \subseteq C_0, \tag{2.5}$$

with $C_\ell = C_\ell[n, k_\ell]_2$ for $\ell = 0, 1, \ldots, a$. Every one of the codes is generated by $r_\ell = n - k_\ell$ parity-check equations. Let $\mathbf{h}_1, \mathbf{h}_2 \ldots, \mathbf{h}_{r_a} \in \mathbb{F}_2^n$ be the equations that generate the smallest code C_a and suppose that C_ℓ is generated by $\mathbf{h}_1, \mathbf{h}_2 \ldots, \mathbf{h}_{r_\ell}$. This guarantees that the inclusions in (2.5) are respected.

Definition 2.15 (Construction D')**.** *With the previous notation and imposing* $r_{-1} = 0$*, we say that a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is built by* Construction D' (with $a + 1$ levels) *when*

$$\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{h}_j\mathbf{x}^T \equiv 0 \bmod 2^{\ell+1}, \forall \ell = 0, 1, \ldots, a \text{ and } r_{a-\ell-1} + 1 \leq j \leq r_{a-\ell}\}.$$

Also in this case, Construction A over $\mathbb{Z}$ with $p = 2$ can be seen as a 1-level Construction D' (compare (2.2) and (2.15)).

2.2.3 Other constructions

For the sake of completeness, we integrate the presentation of infinite constellation constructions from codes recalling that there exist also Construction B and C. These are less employed in the literature that we refer to, so we do not give much detail on them. We simply report that:

- Construction B employs a binary linear code C with minimum distance equal to 8 to obtain a lattice whose points $\mathbf{x}$ are such that:
  1. $\mathbf{x}$ is congruent modulo 2 to a codeword of C.
  2. The sum of all the coordinates of $\mathbf{x}$ is divisible by 4.
- Construction C generates non-lattice constellations. Construction D is a modification of Construction C.

A deeper discussion of these constructions can be found in [CS99].

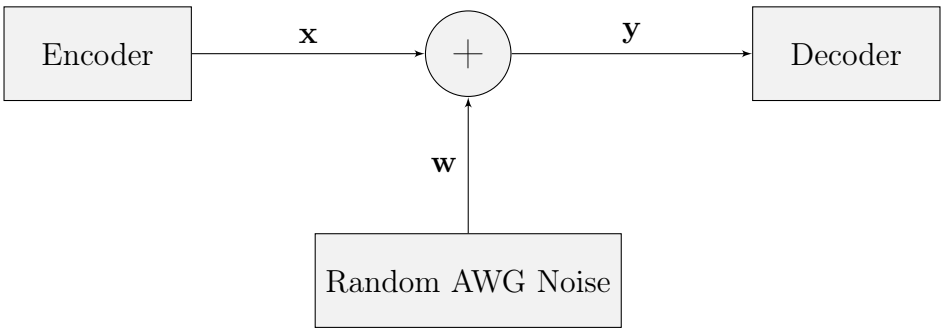

Figure 2.4: The AWGN channel model: the output $\mathbf{y}$ equals the coded input $\mathbf{x}$ plus the random noise $\mathbf{w}$.

2.3 Lattices for the AWGN channel

2.3.1 The AWGN channel and Maximum Likelihood decoding

This thesis in entirely dedicated to the application of lattices to channel coding. In particular, the channel model that we deal with is the so-called *Additive White Gaussian Noise* channel (or, briefly, *AWGN* channel). It is a real channel, meaning that its effect is to add to the channel entry $\mathbf{x}$ a random noise vector $\mathbf{w}$ (see Figure 2.4). The coordinates of $\mathbf{w}$ are n i.i.d. random variables which follow a Gaussian distribution with average 0 and fixed variance σ^2. Moreover, the noise is independent of the channel input $\mathbf{x}$.

The motivation for investigating the AWGN channel lies in the fact that it is a model for some very common communication channels, for example satellite links or telephone channels (both wired and wireless). Of course, for any fixed noise variance, it is easy in principle to design a code which allows reliable communication through the AWGN channel. It suffices to take all codewords far apart one from another and almost every instance of the random noise will be correctly decoded. Nevertheless, this may a priori imply that the input ensemble has very large energy, which is not a good assumption for practical reasons. The model becomes more interesting if we add some power constraint: we assume that every codeword $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is such that

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \leq P, \tag{2.6}$$

that is, we impose a maximum squared Euclidean norm of nP to all the codewords, for some fixed value P, called the *average power per dimension*. With this hypothesis, it is a well-known result of Shannon (well presented in [CT91]) that the capacity of the AWGN channel is

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}\right) = \frac{1}{2}\log_2\left(1 + \mathrm{SNR}\right) \quad \text{bits per transmission.} \tag{2.7}$$

Let X, Y and W be the random variables that represent respectively the channel input, the channel output and the random noise. Correspondingly, let $p_X(\mathbf{x}), p_Y(\mathbf{y})$ and $p_W(\mathbf{w})$ be their probability density functions. In the case of AWG noise, the channel is represented by the following transition probabilities (with standard notation for marginal and conditional):

$$\begin{aligned} p_{Y|X}(\mathbf{y}|\mathbf{x}) &= \prod_{i=1}^{n} p_{Y_i|X_i}(y_i|x_i) \\ &= \prod_{i=1}^{n} p_{W_i|X_i}(w_i|x_i) \\ &= \prod_{i=1}^{n} p_{W_i}(w_i) \\ &= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{w_i^2}{2\sigma^2}\right), \end{aligned} \tag{2.8}$$

where the first equality holds because of the coordinate-wise independence of the noise, the second one from the fact that $y_i = x_i + w_i$, the third one from the noise independence of $\mathbf{x}$ and the fourth one from the definition of the noise. The decoder tries to guess from the output $\mathbf{y}$ what the transmitted message $\mathbf{x}$ is. *Maximum A Posteriori (MAP)* decoding, which is optimal, consists in finding the codeword $\hat{\mathbf{x}}(\mathbf{y})$ that maximises $p_{X|Y}(\mathbf{x}|\mathbf{y})$:

$$\begin{aligned} \hat{\mathbf{x}}(\mathbf{y}) &= \arg \max_{\text{codewords } \mathbf{x}} p_{X|Y}(\mathbf{x}|\mathbf{y}) \\ &= \arg \max_{\text{codewords } \mathbf{x}} p_{Y|X}(\mathbf{y}|\mathbf{x}) \frac{p_X(\mathbf{x})}{p_Y(\mathbf{y})} \\ &= \arg \max_{\text{codewords } \mathbf{x}} p_{Y|X}(\mathbf{y}|\mathbf{x}) p_X(\mathbf{x}). \end{aligned}$$

Notice that, if $\mathbf{x}$ belongs to a finite set, if $p_X(\mathbf{x})$ is uniform and there is no distinction among the input probabilities, then

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \max_{\text{codewords } \mathbf{x}} p_{Y|X}(\mathbf{y}|\mathbf{x})$$

and this decoding rule is also known as *Maximum Likelihood (ML)* decoding. In this case, writing $w_i = y_i - x_i$ and using (2.8), we get:

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \max_{\text{codewords } \mathbf{x}} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i - x_i)^2}{2\sigma^2}\right).$$

The decreasing nature of the exponential functions clearly implies that the codeword that maximises the product on the right is simply the one for which $|y_i - x_i|$ is minimised for every i. In other words, we have just proved that for the AWGN channel with equiprobable inputs, *the optimal decoder just looks for the codeword closest to the received point* $\mathbf{y}$. Unless differently specified, we will always suppose that the inputs are equally distributed and talk without distinction of "optimal" and "ML" decoding.

2.3.2 Lattices and lattice codes

The problem that we want to investigate is how good can lattices be for the transmission of information over the AWGN channel. In other words, we suppose that the set of codewords consists of a subset of a lattice Λ. Typically, every message is associated with a lattice point and we give the following definition:

Definition 2.16 (Lattice code)**.** *Given a lattice $\Lambda \subseteq \mathbb{R}^n$ and a bounded region $S \subseteq \mathbb{R}^n$, a* lattice code *(or* lattice constellation*) $\mathcal{C}$ is the intersection of Λ and S, used as input set for the AWGN channel:*

$$\mathcal{C} = S \cap \Lambda.$$

S is called the shaping region *and, if M is the cardinality of the lattice code, its* rate *is defined as*

$$R_{\mathcal{C}} = \frac{\log_2(M)}{n}.$$

Desirable features of the lattice Λ and the shaping region are:

1. We would like Λ to be "geometrically good" in order to intrinsically be a "good" noise corrector, independently of the shaping region.

2. Among all subsets of Λ that allow to encode our set of messages, we would like to choose one which has the lowest total energy (in the sense of minimising P in (2.6)).

The first point is stated in a quite vague way and for now we just would like to point out the fact that not all lattices have the same performance. We will give this concept proper treatment in a while, when we will talk about the unconstrained AWGN channel and Poltyrev capacity (see also Section 2.4.4).

The second aspect can be translated into looking for the lattice that allows to send the biggest number of messages for fixed volume and shaping region; equivalently, we search for the lattice of given volume for which the biggest possible number of lattice points lie inside the given shaping region. A complementary point of view is to fix the shaping region and the rate of the lattice code and look for the lattice whose intersection with the shaping region has the desired cardinality and which has the greatest minimum distance. One can also remark that the problem of finding a "good" lattice for the AWGN channel with ML decoder is also related to the problem of finding a lattice with "good" fundamental gain (cf. Definition 2.9).

Some history

We would like to go into some more detail about the state of the art of lattice-based information theory.The first work that is definitely worth mentioning is the analysis made by de Buda in [dB75], dating 1975. He showed how lattice codes whose shaping region is a sphere can be asymptotically reliably decoded at any rate up to $1/2 \log_2(P/\sigma^2)$ (in the notation of (2.7)) under *lattice decoding*.

Lattice decoding is a decoding strategy that does not take into account the shaping region defining the constellation. In other words, a lattice decoder simply returns the closest lattice point to the decoder input, regardless of the fact that it belongs to the constellation or not. As a consequence, the decoding decision regions are all equivalent and coincide with the Voronoi regions of the lattice points (cf. Definition 2.6). Of course, this method is suboptimal with respect to the real ML decoder (also referred to as a *nearest-neighbour* or *nearest-codeword* decoder). Nevertheless, its easier algorithmic nature, due to the fact that the shaping region boundary does not affect it, makes it appealing for both theoretical analysis and practical implementation.

The work by de Buda continued in [dB89] and was partially corrected by Linder, Schlegel and Zeger in [LSZ93]. They were able to prove that lattice codes can attain the capacity of the AWGN channel under optimal decoding, with shaping determined by "thin" spherical shells. This peculiar shaping region actually makes the code lose most of its lattice structure and look similar to a random code on a sphere. Urbanke and Rimoldi filled this gap with the proof that lattice codes made up of the intersection between a ball and a lattice are capacity-achieving under nearest-codeword decoding (cf. [UR98]).

Thus, the theoretical problem of showing that lattice codes are capacity-achieving was solved. Nonetheless, the question whether this result can be obtained under (a priori non-optimal) lattice decoding remained answerless. It resisted the attempt of Magalhães de Oliveira and Battail, too, whose proof in [MdOB90] contained a mistake. Finally, once again they could not go beyond the limit of $1/2\log_2(P/\sigma^2)$ for the maximum rate of a reliably decodable lattice code. In 1997, Loeliger proved the achievability of $1/2\log_2(P/\sigma^2)$ with Construction A lattices over $\mathbb{F}_p$ and conjectured that this limit could not be overcome with lattice decoding.

It has been necessary to wait for Erez and Zamir's solution to the problem, based on the MLAN (Modulo-Lattice Additive Noise) channel and Voronoi constellations [EZ04] with Construction A lattices (cf. Definition 2.13). We invite the reader to go to Section 4.1 and Section 4.2 to find a more detailed discussion on this strategy and to see how we have been able to improve this result (cf. Theorem 4.1). More recently, Belfiore and Ling have proposed a solution that involves a non-uniform distribution on the channel inputs and an infinite (but probabilistically finite) codebook (cf. [LB13]).

Poltyrev capacity

The particular interest sparked by lattice decoding and the intention of separating the problem of finding a good shaping region from the problem of finding intrinsically "good" lattices, motivated Poltyrev to carry out his analysis of the *unconstrained* AWGN channel [Pol94]. It consists of the AWGN channel without the power constraint (2.6). The codebook is then made by the whole unbounded lattice, for which lattice decoding turns out to be optimal. Poltyrev showed once again that the rate $1/2\log_2(P/\sigma^2)$ is achievable and he also extended his analysis to non-lattice infinite constellations. He provided an exponential random coding bound for

the decoding error probability and established the notion of *normalised logarithmic density*. This concept allows one to compare the performance of different lattices in this unbounded setting for which capacity equals infinity and the notion of coding rate becomes meaningless.

Definition 2.17 (Normalised logarithmic density)**.** *Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice. The* normalised logarithmic density (NLD) *of Λ is*

$$\delta = \frac{\ln\left(\mathrm{Vol}(\Lambda)^{-1}\right)}{n}.$$

We report now Poltyrev's result, in a weaker version with respect to its statement in [Pol94]; however, this will be sufficient for our purposes. A notational remark: we denote by P_e the probability of making a decoding error, since it does not depend on the sent lattice point because of the lattice symmetry and the unboundedness of the codebook.

Theorem 2.1 (Poltyrev, 1997)**.** *For the unconstrained AWGN channel with noise variance σ^2, let $\delta^* = 1/2\ln(1/(2\pi e\sigma^2))$. Then, the following statements hold:*

1. *There exists a sequence of lattices $\Lambda_n \subseteq \mathbb{R}^n$ of NLD equal to $\delta < \delta^*$ such that P_e decreases to 0 exponentially in n.*

2. *For every sequence of lattices with NLD strictly greater than δ^*, the decoding error probability P_e does not tend to 0 when n tends to infinity.*

The theorem suggests the following definition:

Definition 2.18 (Generalised capacity)**.** *We give the name of* generalised capacity *of the unconstrained AWGN channel to the quantity*

$$\delta^* = \frac{1}{2}\ln\frac{1}{2\pi e\sigma^2}.$$

It is the threshold that separates the values for which reliable lattice decoding is possible from values for which no reliably decodable sequence of lattices exists.

Note that the inequality

$$\frac{\ln\left(\mathrm{Vol}(\Lambda)^{-1}\right)}{n} = \delta < \delta^* = \frac{1}{2}\ln\frac{1}{2\pi e\sigma^2}$$

puts in relation the noise variance of the channel with the volume of the lattices among which a reliably decodable sequence can be found. Symmetrically, for a fixed lattice volume, we can make explicit the maximum noise variance value for which reliable decoding of a sequence of lattices with that volume exists:

$$\sigma^2 < \frac{\mathrm{Vol}(\Lambda)^{2/n}}{2\pi e} = \sigma^2_{\max}. \tag{2.9}$$

Definition 2.19 (Poltyrev-capacity-achieving families)**.** *We say that a family of lattices $\Lambda_n \subseteq \mathbb{R}^n$ with fixed volume V_n* achieves Poltyrev capacity *if for all $\sigma^2 < \sigma^2_{\max}$, the probability of a decoding error for the unconstrained AWGN channel of noise variance σ^2 tends to 0 when n tends to infinity.*

Observe that there is a slight abuse in this definition, due to the fact that it is more usual to define the capacity as an intrinsic feature of the channel. Here, instead, we prefer to define it as a maximum noise variance value, given a lattice volume. This is the notion of Poltyrev capacity to which we will refer in Theorem 3.1 and Theorem 3.2. The probability of decoding error is meant with respect to the randomness related to both the lattice family and the error distribution. It is in this sense that the defintion above should be read.

We conclude the section recalling that Ingber, Zamir and Feder have very recently revisited Poltyrev's asymptotic results and extended them to finite-dimension infinite constellations [IZF13].

2.4 Some problems involving lattices

In this section, we would like to introduce some problems that classically involve lattices and that are related to coding theory in Euclidean space. Far from providing an extensive summary of all the classical problems concerning lattices, we will only face the problems of *sphere packing, sphere covering, channel coding and quantisation*. We present them because they are useful for the general background and because they will be recalled in state of the art part of Chapter 4. Furthermore, we want to mention that a simultaneous solution to all the four problems exists and it will turn out to be of great relevance to us. Indeed, it has been shown by Erez, Litsyn and Zamir [ELZ05] that there exists a random family of lattices that solve all of these problems at the same time. This family is based on Construction A and extensively used all along this dissertation, assuming a dominating role in our work.

Of course, these problems naturally arise from their binary Hamming space analogue. The literature on this subject is extensively rich and many famous results exist, showing how random codes can solve the different problems (we refer the reader to [ELZ05] for a precise and concise introduction on that). As typically happens, the shortage of structure of fully random codes, despite its practicality for probabilistic theoretical results, pushed the scientific community to look for linear-code-based solutions. The step from linear codes in the Hamming space to lattices in the Euclidean space is then conceptually brief and leads to the formulation of the problems that we will present below here. We will substantially follow the descriptive flow of [ELZ05], without adding any mathematical novelty.

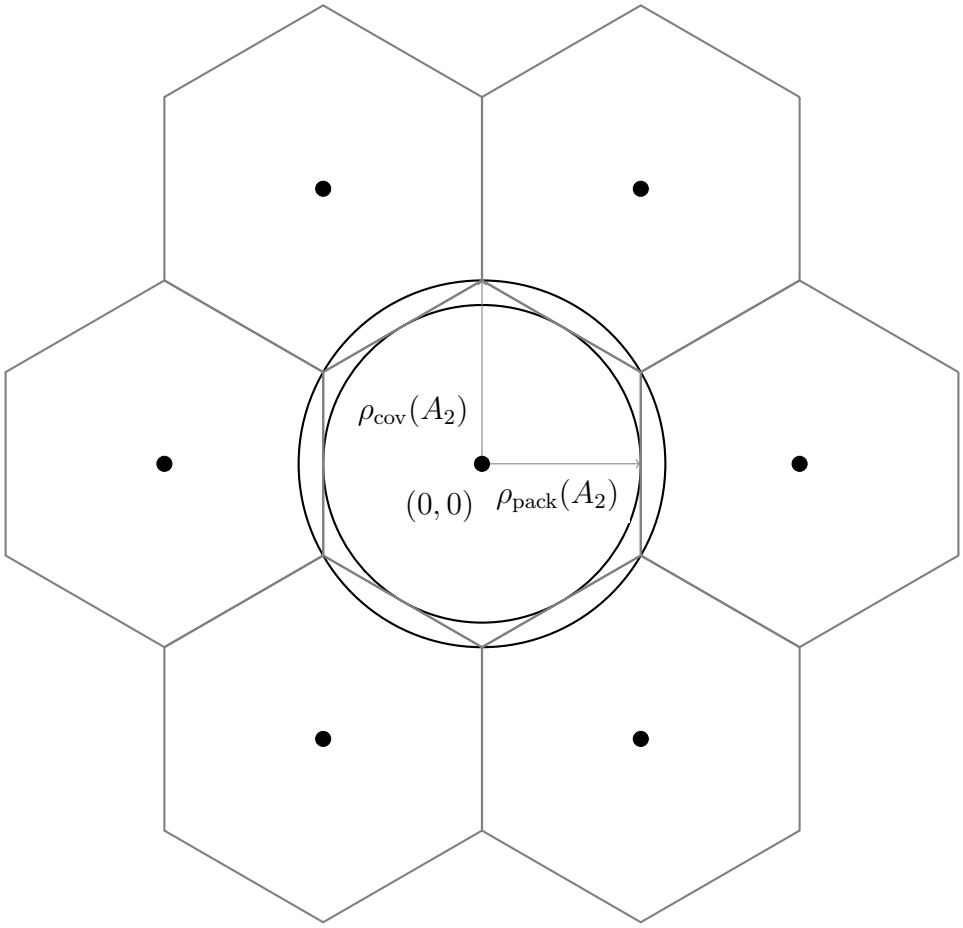

Figure 2.5: The packing and covering radius of the hexagonal lattice A_2 in $\mathbb{R}^2$. It is clearly visible how the Voronoi region of the lattice is included between the packing and the covering spheres.

2.4.1 The sphere packing problem

Let $B_{\mathbf{c},n}(\rho) \subseteq \mathbb{R}^n$ be the ball centred at $\mathbf{c}$ of radius ρ. Given a lattice $\Lambda \subseteq \mathbb{R}^n$, let

$$B_{\Lambda,n}(\rho) = \bigcup_{\mathbf{x}\in\Lambda} B_{\mathbf{x},n}(\rho).$$

This set is called a *packing* if

$$B_{\mathbf{x},n}(\rho) \cap B_{\mathbf{y},n}(\rho) = \emptyset, \forall \mathbf{x}, \mathbf{y} \in \Lambda \text{ and } \mathbf{x} \neq \mathbf{y}.$$

Definition 2.20 (Packing radius)**.** *The* packing radius ρ_{pack} *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is defined as*

$$\rho_{\text{pack}}(\Lambda) = \sup\{\rho \in \mathbb{R} : B_{\Lambda,n}(\rho) \textit{ is a packing}\}. \tag{2.10}$$

For a fixed lattice volume, the *sphere packing problem* in dimension n consists in looking for the Λ with the given volume such that its packing radius is the greatest

possible. Observe that

$$B_{\mathbf{x},n}\left(\rho_{\text{pack}}(\Lambda)\right) \subseteq \mathcal{V}(\mathbf{x}) \tag{2.11}$$

for every $\mathbf{x} \in \Lambda$ (see also Figure 2.5). This leads to say that a lattice with a big packing radius is a lattice whose Voronoi region is "very spherical". In other terms, we look for lattices whose Voronoi regions are not too "narrow" or asymmetric, so that they can contain a big packing sphere. More formally, let us express this concept starting with another definition:

Definition 2.21 (Effective radius)**.** *The* effective radius *of a lattice $\Lambda \subseteq \mathbb{R}^n$ is the radius $\rho_{\text{eff}}(\Lambda)$ such that*

$$\text{Vol}\left(B_{\mathbf{0},n}\left(\rho_{\text{eff}}(\Lambda)\right)\right) = \text{Vol}\left(\mathcal{V}(\Lambda)\right).$$

The inclusion in (2.11) implies that the quantity $\delta_{\text{pack}}(\Lambda) = \rho_{\text{pack}}(\Lambda)/\rho_{\text{eff}}(\Lambda)$, often called *packing efficiency* of Λ, is smaller than 1 (in particular, it is equal to 1 only in dimension $n = 1$; otherwise, the inequality is strict). We would like to find lattices whose packing efficiency is the closest possible to 1.

Mathematically speaking, this problem is probably the central one in lattice theory and it has been (and still is) the object of profound investigation. Nevertheless, it is far from being easy and it has not come to a complete solution. We briefly recall here two asymptotic results (as provided in [ELZ05]): if Λ_n denotes a generical n-dimensional lattice, let

$$\delta^*_{\text{pack}} = \limsup_{n\to\infty} \sup_{\Lambda_n} \delta_{\text{pack}}(\Lambda_n)$$

be the optimal asymptotic packing efficiency. The Minkowski-Hlawka theorem [Rog64] gives the best known lower bound for δ^*_{pack}, namely $\delta^*_{\text{pack}} \geq 1/2$. On the other hand, it is known that δ^*_{pack} is strictly smaller than 1 and in particular $\delta^*_{\text{pack}} \leq 0.660211...$ (cf. [KL78, CS99]).

Definition 2.22 (Goodness for packing)**.** *We say that a sequence of lattices Λ_n is (asymptotically)* good for packing *if*

$$\limsup_{n\to\infty} \sup_{\Lambda_n} \delta_{\text{pack}}(\Lambda_n) \geq \frac{1}{2},$$

that is, if it achieves Minkowski's bound.

As we have previously mentioned, Construction A lattices are shown to be good for packing [ELZ05, Rog64].

2.4.2 The sphere covering problem

The *sphere covering problem* is somehow dual to the sphere packing problem. Keeping the same notation as before, we say that the set $B_{\Lambda,n}(\rho)$ is a *covering* if $\mathbb{R}^n = B_{\Lambda,n}(\rho)$. This means that the ensemble of all the balls of radius ρ centred at the points of Λ cover the whole space. Dually to the notion of packing radius, we can give the following definition:

Definition 2.23 (Covering radius)**.** *The* covering radius *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is defined as*

$$\rho_{\text{cov}}(\Lambda) = \min\{\rho \in \mathbb{R} : B_{\Lambda,n}(\rho) \text{ is a covering}\}.$$

For a fixed lattice volume, the *sphere covering problem* consists in finding lattices with the smallest possible covering radius. As it is represented in Figure 2.5, for every lattice point $\mathbf{x} \in \Lambda$,

$$\mathcal{V}(\mathbf{x}) \subseteq B_{\mathbf{x},n}\left(\rho_{\text{cov}}(\Lambda)\right).$$

So, once again, "spherical" Voronoi regions imply that a lattice has a small covering radius. We define the *covering efficiency* of a lattice Λ as the quotient $\delta_{\text{cov}}(\Lambda) = \rho_{\text{cov}}(\Lambda)/\rho_{\text{eff}}(\Lambda)$. Of course, the smallest possible covering efficiency equals 1. Let δ^*_{cov} be the optimal asymptotic covering efficiency:

$$\delta^*_{\text{cov}} = \liminf_{n\to\infty} \inf_{\Lambda_n} \delta_{\text{cov}}(\Lambda_n);$$

Rogers [Rog59] showed that $\delta^*_{\text{cov}} = 1$ and we give the following definition:

Definition 2.24 (Goodness for covering)**.** *A sequence of lattices* Λ_n *is (asymptotically)* good for covering *if*

$$\liminf_{n\to\infty} \inf_{\Lambda_n} \delta_{\text{cov}}(\Lambda_n) = 1,$$

that is, if it achieves Rogers' equality.

Construction A lattices are good for covering too [ELZ05].

2.4.3 The quantisation problem

We defined in Section 2.1 what a *lattice quantiser* is (cf. Definition 2.4). In *Mean Squared Error (MSE) quantisation*, the quantiser associated with the Voronoi region of a lattice Λ is considered (we denoted it by $Q_\Lambda(\cdot)$). A result of Gersho [Ger79] puts in relation the distortion for high-resolution lattice quantisation of a source and the following quantity:

Definition 2.25 (Normalised second moment)**.** *The* normalised second moment *of a lattice* $\Lambda \subseteq \mathbb{R}^n$ *is defined as*

$$G(\Lambda) = \frac{1}{n \operatorname{Vol}(\mathcal{V}(\Lambda))^{1+2/n}} \int_{\mathcal{V}(\Lambda)} ||\mathbf{y}||^2 d\mathbf{y}.$$

Namely, the normalised second moment and the distortion are proportional. The *quantisation problem* translates into finding lattices whose normalised second moment is the smallest possible.

Let Λ_n be the generic lattice of $\mathbb{R}^n$ and let $G_n = \min_{\Lambda_n} G(\Lambda_n)$. For geometric reasons, G_n is greater than the normalised second moment of a sphere. Moreover, it is known that the latter tends to $1/2\pi e$ when n goes to infinity. Rate-distortion theory and the Shannon lower bound state that the random-code formula for the distortion is the same as the distortion formula for a lattice family for which $G(\Lambda_n)$ tends to $1/2\pi e$. This justifies the following definition:

Definition 2.26 (Goodness for quantisation)**.** *We say that a lattice family Λ_n is (asymptotically)* good for quantisation *if*

$$\lim_{n\to\infty} G(\Lambda_n) = \frac{1}{2\pi e}.$$

Once again, "goodness" of lattices deals with their Voronoi region and lattices for which it is "spherical" (this time, in the sense of the normalised second moment) are good. It is known that there exist good lattice quantisers (cf. [ZF96]) and Construction A lattices are among them (cf. [ELZ05]).

2.4.4 The channel coding problem

Finally, we mention the *channel coding problem*. It is simply the problem of coding over the unconstrained AWGN channel, as it was formulated by Poltyrev in [Pol94]). We have already introduced it in Section 2.3.2 and we have dedicated an entire paragraph to the related concept of Poltyrev capacity. Now, we just want to complete the formulation of this problem adding a definition of "goodness" for AWGN coding:

Definition 2.27 (Goodness for coding)**.** *We say that a lattice family Λ_n is* good for coding *over the unconstrained AWGN channel with noise variance σ^2, if:*

1. *It is Poltyrev-capacity-achieving (cf. Definition 2.19).*

2. *When the dimension n tends to infinity, its error decoding probability decreases to 0 like $e^{-nE_U(\Lambda_n,\sigma^2)+o(1)}$, where $E_U(\Lambda_n,\sigma^2)$ is the error exponent derived by Poltyrev in [Pol94] for random lattices.*

Construction A lattices are good for coding, too [ELZ05].

For the sake of completeness, we report below here the values of $E_U(\Lambda_n,\sigma^2)$ and we refer the reader to [Pol94] to learn how they are computed. Let $\nu = \mathrm{Vol}(\Lambda_n)^{2/n}/2\pi e\sigma^2$. Recall that Theorem 2.1 states that reliable decoding is possible if $\nu > 1$; vice versa, it is not possible if $\nu < 1$ and $\nu = 1$ corresponds to Poltyrev capacity $\sigma^2_{\max}$ (cf. (2.9)). Then

$$E_U(\Lambda_n,\sigma^2) = E_U(\nu) = \begin{cases} \frac{1}{2}\left((\nu-1)-\ln\nu\right) & \text{if } 1\le\nu\le 2 \\ \frac{1}{2}\ln(e\nu/4), & \text{if } 2\le\nu\le 4 \\ \nu/8, & \text{if } \nu\ge 4. \end{cases}.$$

2.5 Some useful lemmas

The following lemmas deal with probability theory, combinatorics and geometry and we list them one by one here below, even if for now they may appear unrelated.

2.5.1 Chebyshev's inequality

Chebyshev's inequality is well-known and very useful. For the sake of completeness, we give a short proof of it, following [Bal98].

Lemma 2.1 (Chebyshev's inequality)**.** *Let X be a random variable and let $\tau > 0$ be any positive quantity.*

$$\mathcal{P}\left\{|X - \mathbb{E}[X]| > \tau\right\} \leq \frac{\mathrm{Var}(X)}{\tau^2}.$$

Proof. Consider the auxiliary random variable

$$Y = \begin{cases} \tau^2, & \text{if } |X - \mathbb{E}[X]| > \tau \\ 0, & \text{otherwise} \end{cases}.$$

If $|X - \mathbb{E}[X]| > \tau$, then $(X - \mathbb{E}[X])^2 > \tau^2 = Y$ and similarly, if $|X - \mathbb{E}[X]| \leq \tau$, we have $(X - \mathbb{E}[X])^2 \geq 0 = Y$. Then $(X - \mathbb{E}[X])^2 \geq Y$ in both cases and we conclude:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] \geq \mathbb{E}[Y] = \tau^2 \mathcal{P}\{|X - \mathbb{E}[X]| > \tau\}.$$

□

2.5.2 The typical norm of a random noise vector

In the previous section we have mentioned additive white Gaussian noise. The next classical lemma describes the "typical" norm of a random noise vector in very high dimension. More formally:

Lemma 2.2 (Typical norm of the AWG noise)**.** *Consider n i.i.d. random variables $X_1, \ldots, X_n$, each of them following a Gaussian distribution of mean 0 and variance σ^2. Let $\rho = \sqrt{\sum_{i=1}^n X_i^2}$. Then, for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathcal{P}\left\{\sigma\sqrt{n}\,(1-\varepsilon) \leq \rho \leq \sigma\sqrt{n}\,(1+\varepsilon)\right\} = 1.$$

Proof. It is known that, since $X_i \sim \mathcal{N}(0, \sigma^2)$, $i = 1, \ldots, n$, then X_i^2 has a gamma distribution and $\mathbb{E}[X_i^2] = \sigma^2$, $\mathrm{Var}(X_i^2) = 2\sigma^4$. Consequently, by the independence of the X_i,

$$\mathbb{E}[\rho^2] = n\sigma^2, \ \mathrm{Var}(\rho^2) = 2n\sigma^4.$$

Lemma 2.1 with $Y = \rho^2$ and $\tau = \kappa\sqrt{2n}\sigma^2$ for some $\kappa > 0$ gives

$$\mathcal{P}\left\{|\rho^2 - n\sigma^2| > \kappa\sqrt{2n}\sigma^2\right\} \leq \frac{1}{\kappa^2}.$$

If we choose $\kappa = \kappa(n)$ such that $\lim_{n\to\infty} \kappa = +\infty$, then

$$\lim_{n\to\infty} \mathcal{P}\left\{|\rho^2 - n\sigma^2| \leq \kappa\sqrt{2n}\sigma^2\right\} = 1. \tag{2.12}$$

As a consequence,

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho^2 \leq \sigma^2 n\left(1+\kappa\sqrt{\frac{2}{n}}\right)\right\} = 1.$$

Taking for example $\kappa = \ln n$, we have that $\lim_{n\to\infty} \kappa\sqrt{2/n} = 0$. This implies that for n big enough and for every $\varepsilon > 0$

$$\sqrt{1+\kappa\sqrt{\frac{2}{n}}} < 1+\varepsilon$$

and

$$\mathcal{P}\left\{\rho \leq \sigma\sqrt{n}\left(\sqrt{1+\kappa\sqrt{\frac{2}{n}}}\right)\right\} \leq \mathcal{P}\left\{\rho \leq \sigma\sqrt{n}\,(1+\varepsilon)\right\}.$$

This is enough to conclude that

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho \leq \sigma\sqrt{n}\,(1+\varepsilon)\right\} = 1,$$

too, which proves the statement restricted to the second inequality. But notice that (2.12) also implies that

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho^2 \geq \sigma^2 n\left(1-\kappa\sqrt{\frac{2}{n}}\right)\right\} = 1.$$

This leads to the conclusion that

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho \geq \sigma\sqrt{n}\,(1-\varepsilon)\right\} = 1,$$

too, and the lemma is proved. □

2.5.3 Integer points in a sphere

In the next chapters, we will often need to count the number of integer points inside a sphere of a given radius. For this purpose, we will use the following lemma:

Lemma 2.3 (Integer points inside a sphere)**.** *Let* $B_{\mathbf{c},n}(\rho) = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}-\mathbf{c}||^2 \leq \rho^2\}$ *be the ball centred at* $\mathbf{c}$ *of radius* ρ*. Let* $N = |\mathbb{Z}^n \cap B_{\mathbf{c},n}(\rho)|$*. Then*

$$\mathrm{Vol}\left(B_{\mathbf{c},n}(\rho)\right)\left(\max\left\{1-\frac{\sqrt{n}}{2\rho}, 0\right\}\right)^n \leq N \leq \mathrm{Vol}\left(B_{\mathbf{c},n}(\rho)\right)\left(1+\frac{\sqrt{n}}{2\rho}\right)^n.$$

Proof. Consider, for every $\mathbf{z} \in \mathbb{Z}^n$, the cube $\mathcal{C}_{\mathbf{z}}$ centred at $\mathbf{z}$ of edge (and volume) equal to 1. Let

$$U = \bigcup_{\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{c},n}(\rho)} \mathcal{C}_{\mathbf{z}}$$

and notice that $|\mathbb{Z}^n \cap B_{\mathbf{c},n}(\rho)| = \mathrm{Vol}(U)$.

Now, let S_1 be the sphere inscribed in U and S_2 the one circumscribed to U. The definition of U and the fact that the length of the diagonal of any $\mathcal{C}_{\mathbf{z}}$ is $\sqrt{n}$ imply that the radius of S_1 is at least $\rho - \sqrt{n}/2$, while the one of S_2 is at most $\rho + \sqrt{n}/2$. Therefore,

$$\mathrm{Vol}\left(B_{\mathbf{c},n}(\rho)\right)\left(1 - \frac{\sqrt{n}}{2\rho}\right)^n = \mathrm{Vol}\left(B_{\mathbf{c},n}\left(\rho - \frac{\sqrt{n}}{2}\right)\right) \leq \mathrm{Vol}(S_1) \leq \mathrm{Vol}(U)$$

and

$$\mathrm{Vol}(U) \leq \mathrm{Vol}(S_2) \leq \mathrm{Vol}\left(B_{\mathbf{c},n}\left(\rho + \frac{\sqrt{n}}{2}\right)\right) = \mathrm{Vol}\left(B_{\mathbf{c},n}(\rho)\right)\left(1 + \frac{\sqrt{n}}{2\rho}\right)^n.$$

Since $|\mathbb{Z}^n \cap B_{\mathbf{c},n}(\rho)| = \mathrm{Vol}(U)$, these two inequalities give us the wanted result. □

2.5.4 Approximations of the binomial coefficient

We recall the following result:

Lemma 2.4 (Stirling's approximation of the factorial function)**.** *Let n be a natural number. Then*

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{1}{12n} - \frac{1}{360n^3}\right) < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{1}{12n}\right) \tag{2.13}$$

and, asymptotically,

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

We omit the proof, which can be found in some classical handbooks of analysis or number theory. Of course, the symbol $\sim$ indicates the "asymptotic equality" relation: we say that a function $f(n)$ is *asymptotic to* a function $g(n)$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$$

and we denote this relation by $f(n) \sim g(n)$. We move on recalling the following definition:

Definition 2.28 (Binary entropy function)**.** *The* binary entropy *is the function:*

$$h(x) = -x\log_2 x - (1-x)\log_2(1-x), \text{ for all } 0 < x < 1.$$

By continuity, we extend the definition to $h(1) = h(0) = 0$.

This function is of help to establish an accurate approximation of the binomial coefficient:

Lemma 2.5. *Let n be a natural number and let $0 \leq \theta \leq 1$ be any rational number such that θn is natural, too. Then:*

$$\frac{1}{\sqrt{8n\theta(1-\theta)}} 2^{nh(\theta)} \leq \binom{n}{\theta n} \leq \frac{1}{\sqrt{2\pi n\theta(1-\theta)}} 2^{nh(\theta)}.$$

The proof of the lemma, as it is proposed by [MS77], is nothing more than a direct computation that employs inequality (2.13) to approximate the factorial functions in the binomial coefficient.

2.5.5 The volume of a sphere

Lemma 2.6. *Let $B = B_{\mathbf{c},n}(\rho) \subseteq \mathbb{R}^n$ be the ball centred at $\mathbf{c}$ of radius ρ. Then*

$$\mathrm{Vol}(B) = \frac{(\sqrt{\pi}\rho)^n}{\Gamma\left(\frac{n}{2}+1\right)} \sim \frac{1}{\sqrt{\pi n}} \left(\frac{\sqrt{2\pi e}\rho}{\sqrt{n}}\right)^n,$$

where $\Gamma(\cdot)$ is Euler's Gamma function.

We do not provide the proof of this result, which is classical; the asymptotic approximation comes from Stirling's formula.

Chapter 3

Infinite LDA lattice constellations

This chapter is entirely dedicated to a particular family of lattices, called Low-Density Construction A (LDA) lattices. In Section 2.3, we have already explained what we mean by achieving Poltyrev capacity with an infinite constellation. We have also recalled that this result is theoretically affordable when we consider Construction A families of lattices (see the proof of [Loe97]). Besides this non-constructive achievement some lattice coding schemes have been put forward with the intention of finding families which have simultaneously practically manageable encoding and decoding algorithms and close-to-capacity performance. Some of these families are inspired by LDPC and turbo codes [SBP06, BC08, SFS08, SSP12, SS13]. The most recent of these works concerns polar lattices [YLW13, YL12], for which a theoretical result about capacity achievement is available, too. However, to the best of our knowledge no other analysis has been provided till now that gives both a strong theoretical result and satisfactory numerical performance. This is actually the purpose of this dissertation and this chapter concerns the abstract analysis of the capacity-achieving properties of LDA lattices:

Definition 3.1 (LDA lattice)**.** *A lattice $\Lambda \subseteq \mathbb{R}^n$ is called a* Low-Density Construction A *(or briefly* LDA*) lattice if it is built with Construction A (cf. Definition 2.13) starting from an LDPC code.*

LDA lattices were first envisaged in [Ere02] and we reintroduced them together with an efficient iterative algorithm in [dPBZB12]. We also carried out a theoretical analysis of their capacity-achieving qualities in [dPBZB13, dPBZ13]. Some of those results are proposed anew and completed in this dissertation. LDA lattices put together the strength of Construction A and LDPC codes (over a non-binary prime field) and will be described and investigated in depth in the sequel. Their main feature is that their corresponding parity-check matrix is sparse. As one can guess, this is the key idea to reconduct their decoding to well-performing, implementable LDPC decoding algorithms. At the same time, a theoretical approach is still manageable to mathematically prove the good properties of the families, even if some sharper and more technical examination is required in contrast with the more general Construction A results [Loe97, ELZ05, GZ07]. This is due to the "less

random" nature of the ensemble we consider, that complicates but does not impede the analysis.

In particular, we propose two different constructions of LDA lattices over a prime field $\mathbb{F}_p$ and show that *they both achieve Poltyrev capacity* of the unconstrained AWGN channel under lattice (ML) decoding. The difference between them concerns the number of non-zero coefficients h_i in a parity-check equation $\sum_{i=1}^{n} h_i x_i \equiv 0 \bmod p$. We call this number the *degree* of the parity-check equation. By definition of an LDA lattice, it has to be small with respect to n. How small? We will consider the two following settings:

1. The case in which the degree grows logarithmically with n. The corresponding random ensemble is described in Section 3.2.1 and the capacity-achieving result is given in Theorem 3.1. This setting is the direct generalisation to lattices of the LDPC code theory. In fact, it was already shown by Gallager that binary LDPC codes need logarithmically growing parity-check equation degrees to achieve the capacity of the binary symmetric channel (see [Gal63, Mac99]). Previous studies of non-binary LDPC codes for modulo additive channels [EM05] similarly require parity-check equations with weight tending to infinity to achieve capacity. We take inspiration from these results and adapt it to our LDA lattices.

2. The case in which the degree is (a well-defined) constant with respect to n. We present this second LDA family in Section 3.3.3 and the capacity-achieving result will be Theorem 3.2. In this case, the probabilistic model is different and somewhat less natural; we also need a particular expansion result concerning the Tanner graphs associated with the LDA random ensemble. This is presented in Section 3.3.2. Moreover, we will also have a further condition on the size of the prime p with respect to the lattice dimension n. At first sight, it appears quite unexpected that LDA lattices achieve capacity even with constant parity-check equation degrees. This definitely stands in sharp contrast with the analogue for binary LDPC codes.

A final remark: in all the proofs of Chapter 3 and Chapter 4, we let the lattice dimension n tend to infinity, as it is usual in this context. The prime number p also needs to tend to infinity (see also [Loe97, EZ04, OE12]) and we define $p = n^{\lambda}$ for some positive constant λ. It is clear that if n changes and λ is fixed, then it is not true that n^{λ} is always a prime number. It would be more precise to say that $p(\lambda)$ is the closest prime number to n^{λ}, or that $p = n^{\lambda(n)}$ for some $\lambda(n)$ assuming values in an interval properly centred at our fixed value λ. Nevertheless, it is possible to show that this variation of $\lambda(n)$ concerns a range which is narrow enough not to impact any of the asymptotic estimations that we compute letting n tend to infinity. In other words, there always exists a prime number p close enough to n^{λ} to make accurate the approximation $p = n^{\lambda}$. Despite the slight abuse of notation, we prefer to keep it that way from now on, in order to write the proofs in the clearest way possible and avoid the overabundance of symbols.

3.1 Poltyrev capacity for Construction A lattices

Before moving on to our main results, we just recall that when the channel random noise is AWG, lattice decoding is equivalent to ML decoding in the case of infinite constellations. Poltyrev capacity (cf. Definition 2.19) corresponds to the threshold noise variance per dimension value that separates reliably decodable and undecodable noise (cf. Section 2.3.2). For n-dimensional lattices with fixed fundamental volume V_n, this maximum "tolerable" noise variance per dimension is equal to (cf. (2.9))

$$\sigma^2_{\max} = \frac{V_n^{2/n}}{2\pi e}.$$

Now, consider Construction A lattices over $\mathbb{Z}$ for some prime p (see Definition 2.13). We have already computed (cf. (2.3)) that the volume of these lattices is equal to $p^{n(1-R)}$, if R is the rate of the underlying linear code. This implies that for a fixed R, Poltyrev capacity for Construction A lattices is equal to

$$\sigma^2_{\max} = \frac{p^{2(1-R)}}{2\pi e}, \tag{3.1}$$

independently of the dimension n. Moreover, this value is also independent of the type of linear code that we use to construct the lattices.

3.2 Random LDA lattices achieve Poltyrev capacity

In this section we prove that there exists a Poltyrev-capacity-achieving random family of LDA lattices with logarithmic-degree parity-check equations. This result is inspired by what is already known about LDPC codes [Gal63, Mac99] and adapted to non-binary Construction A over $\mathbb{Z}$.

We will see very soon what is in detail the LDA lattice family that we consider and the formal proof of our result. But before that, we would like to summarise our strategy as an introduction that should help the reader not to get lost in the mathematical details. Thus, the proof of Theorem 3.1 will develop in this way:

1. Given a random lattice Λ of the family presented in Section 3.2.1, we suppose that the channel input is the point $\mathbf{0} \in \Lambda$.

2. Given the AWGN channel output $\mathbf{y}$, our aim is to show that it is closer to $\mathbf{0}$ than to any other lattice point.

3. Lemma 2.2 guarantees that for n tending to infinity the AWG noise typically lies inside a sphere of a well-determined radius (function of the noise variance). Equivalently, we can consider a sphere (called the *decoding sphere*) centred at $\mathbf{y}$ with the same radius as before. This sphere will typically contain $\mathbf{0}$ and we restrict the argument of the previous item only to lattice points inside this sphere.

4. We first get rid of all the points of $p\mathbb{Z}^n$, showing that they will be typically further from $\mathbf{y}$ than $\mathbf{0}$ (cf Lemma 3.1 and recall that $p\mathbb{Z}^n$ is always contained in Λ).

5. All the other integer points of the decoding sphere have a well-determined probability of belonging to the random lattice. We conclude the proof by an average argument that shows that typically there is no other lattice point than $\mathbf{0}$ inside the decoding sphere.

Finally, we specify that we will let the lattice dimension n tend to infinity. The prime number p also needs to tend to infinity and we define $p = n^\lambda$ for some positive constant λ (see also the end of the introduction of this chapter for some further comment about this choice). The proof of Theorem 3.1 does not need any condition on λ, except that it is constant and positive.

3.2.1 The LDA random ensemble: logarithmic degree of the parity-check equations

It is now the moment of describing our LDA random ensemble. As we have already anticipated, we will do it starting from the associated parity-check matrices. Hence, let p be a prime number and let H be a matrix of size $n(1-R) \times n$, for some $0 < R < 1$ with entries in $\{0, 1, \ldots, p-1\}$. More precisely, let each row of the matrix be a random vector, built independently from each other as follows. Let $0 \leq \Delta \leq n$ be an integer. For a given row of H, let us choose, following a uniform random distribution, exactly Δ coordinates among all the n ones. We assign to these coordinates a value, chosen uniformly at random in $\{0, 1, \ldots, p-1\}$; furthermore, we impose that all the other $n - \Delta$ coordinates are deterministically equal to 0. What we obtain is a matrix in which every row contains exactly $n - \Delta$ zeros and Δ random entries, placed in random positions.

This matrix can be viewed as the parity-check matrix of a k-dimensional random code $C = C[n, k]_p \subseteq {\mathbb{F}_p}^n$, for which all parity-check equations have at most Δ non-zero coefficients. This implies that the rate of C it at least R (but it may also be bigger, due to random choices of the entries).

Of course, if Δ is small with respect to n, C is an LDPC code. We will take into account the set of all LDA lattices $\Lambda = C + p\mathbb{Z}^n \subseteq \mathbb{Z}^n$ such that C is built at random as we have just explained and $\Delta = \beta \ln n$, for some constant β. Theorem 3.1 will need some (mild) conditions on this β.

3.2.2 A lemma on the points of $p\mathbb{Z}^n$

We state and prove a lemma which is used in the proof of Theorem 3.1. It concerns a particular subset of points of our random lattices: the points of $p\mathbb{Z}^n$. They have the characteristic property of always belonging to a lattice $\Lambda = C + p\mathbb{Z}^n$, independently of the random choice of the linear code C. For this reason, we treat them separately with respect to the other (random) lattice points.

We show here that if $\mathbf{0}$ is the sent point, the random noise produces a channel output which is typically closer to $\mathbf{0}$ itself than to any other point of $p\mathbb{Z}^n$. From the point of view of the lattice decoder, this means that the points of $p\mathbb{Z}^n$ do not typically induce a decoding error.

Lemma 3.1. *Let $\Lambda \subseteq \mathbb{R}^n$ be a Construction A lattice, let $\mathbf{0} \in \Lambda$ be the lattice point to be sent over the AWGN channel and let $\mathbf{w}$ be the random noise vector. Furthermore, suppose that the noise variance per dimension is equal to $\sigma^2 = \sigma^2_{\max}(1-\delta)^2$, for some constant $0 < \delta < 1$ and $\sigma^2_{\max} = p^{2(1-R)}/2\pi e$ (see (3.1)). Then, for every $\mathbf{z} \in p\mathbb{Z}^n \setminus \{\mathbf{0}\}$,*

$$\lim_{n\to\infty} \mathcal{P}\{||\mathbf{w}||^2 \geq ||\mathbf{w}-\mathbf{z}||^2\} = 0.$$

Proof. Since $\mathbf{z}$ belongs to $p\mathbb{Z}^n$, a necessary condition when $||\mathbf{w}||^2 \geq ||\mathbf{w}-\mathbf{z}||^2$ is that at least one of the coordinates of $\mathbf{w}$ is bigger than $p/2$ in absolute value. Hence

$$\begin{aligned}\mathcal{P}\{||\mathbf{w}||^2 \geq ||\mathbf{w}-\mathbf{z}||^2\} &\leq \mathcal{P}\{|w_i| \geq p/2, \exists i \in \{1,2,\ldots,n\}\} \\ &\leq \sum_{i=1}^{n} \mathcal{P}\{|w_i| \geq p/2\}. \end{aligned} \tag{3.2}$$

Now, $w_i \sim \mathcal{N}(0,\sigma^2)$ for every $i \in \{1,2,\ldots,n\}$ and the probabilities in the previous sum are all identical and independent from i.

Consider the function $Q(\cdot)$, the tail probability of the standard normal distribution:

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u.$$

For positive y, the Chernoff bound states that

$$Q(y) \leq \frac{1}{2} e^{-\frac{y^2}{2}}.$$

Hence, we can go back to (3.2) and write

$$\begin{aligned}\sum_{i=1}^{n} \mathcal{P}\{|w_i| \geq p/2\} &\leq n\mathcal{P}\{|w_1| \geq p/2\} \\ &= 2nQ\left(\frac{p}{2\sigma}\right) \\ &\leq n\exp\left(-\frac{p^2}{8\sigma^2}\right) \\ &= n\exp\left(-\frac{\pi e p^{2R}}{4(1-\delta)^2}\right),\end{aligned}$$

which decreases to 0 because $p = n^\lambda$. □

3.2.3 The capacity-achieving theorem

We have introduced all the useful elements for stating and proving the main result:

Theorem 3.1. *Let n be a positive integer and let R be a real number such that $0 < R < 1$. Let $p = n^\lambda$ be a prime number for some $\lambda > 0$. Let $\Delta = \beta \ln n$ be an integer number, for some $\beta \in \mathbb{R}$. If*

$$\beta > \lambda + \frac{3}{2(1-R)}, \tag{3.3}$$

then the n-dimensional LDA random ensemble of Section 3.2.1 achieves Poltyrev capacity. The row degree in the parity-check matrix of the underlying LDPC codes is at most Δ and their rate at least R.

Proof. First of all, observe that $\beta \ln n$ is not always an integer, when n changes and β is fixed. For this reason, we should properly write $\Delta = \lfloor \beta \ln n \rfloor$ in all the following computations. Anyway, we prefer to drop the integer part symbols and slightly abuse in notation, pretending that $\beta \ln n$ is integer. This does not change the substance of the proof and allows to lighten some formal analytical aspects.

Let Λ be a lattice of this family. Since we let n change all along the proof, we may probably call it Λ_n. Anyway, we omit the index n for the sake of simplicity and we simply keep implicit this dependence on the dimension. This will not lead to any misunderstanding.

Because of the lattice symmetry and the independence between random noise and channel input, we can suppose that the point of Λ transmitted over the channel is the point $\mathbf{0}$. The AWG noise vector is $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ and the channel output is $\mathbf{y} = \mathbf{w}$. We suppose that the channel noise variance is $\sigma^2 = \sigma^2_{\max}(1-\delta)^2$ for some $0 < \delta < 1$ that can be as small as wanted. Of course, $\sigma^2_{\max}$ is the noise variance value that corresponds to Poltyrev capacity, as made explicit in (3.1).

Lemma 2.2 states that, when n is very large, the vector $\mathbf{y}$ tends to lie within a sphere of radius a bit greater than $\sigma\sqrt{n}$ and centred at $\mathbf{0}$.

Let us consider the *decoding sphere* $\mathcal{B} = B_{\mathbf{y},n}(\sigma\sqrt{n}(1+\varepsilon))$ centred at $\mathbf{y}$, with $\varepsilon > 0$ chosen such that

$$\varepsilon < \frac{\delta}{1-\delta}; \tag{3.4}$$

this last condition will be explained in the sequel.

When n goes to infinity, the point $\mathbf{0}$ is inside the decoding sphere with probability tending to 1; if this occurs, the probability of making a decoding error with a lattice decoder is smaller than the probability that one or more lattice points different from $\mathbf{0}$ lie inside the sphere: if $\mathbf{0}$ is the only lattice point in $\mathcal{B}$, then lattice decoding gives the correct answer; otherwise, an error will possibly come out. Furthermore, Lemma 3.1 guarantees that the possible presence of points of $p\mathbb{Z}^n$ inside the decoding sphere does not actually impede good decoding.

Summarising, we are left to show that, if $\mathcal{N}$ is the random variable that counts the number of lattice points inside $\mathcal{B}$ that do not belong to $p\mathbb{Z}^n$, then

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{N} = 0\} = 1. \tag{3.5}$$

In order to do this, for every integer point $\mathbf{x} \in \mathcal{B} \cap \mathbb{Z}^n$ let $X_{\mathbf{x}}$ be the random variable defined by

$$X_{\mathbf{x}} = \begin{cases} 1, & \text{if } \mathbf{x} \in \Lambda \\ 0, & \text{if } \mathbf{x} \notin \Lambda \end{cases}.$$

By definition,

$$\mathcal{N} = \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n} X_{\mathbf{x}}$$

and to prove (3.5) it is sufficient to show that

$$\lim_{n \to \infty} \mathbb{E}[\mathcal{N}] = 0. \tag{3.6}$$

Observe that

$$\mathbb{E}[\mathcal{N}] = \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n} \mathbb{E}[X_{\mathbf{x}}] = \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{x} \in \Lambda\}. \tag{3.7}$$

If H is the parity-check matrix of C,

$$\begin{aligned} \mathcal{P}\{\mathbf{x} \in \Lambda\} &= \mathcal{P}\{H\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\} \\ &= \left(\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\}\right)^{n(1-R)}, \end{aligned}$$

where $\mathbf{h}$ represents any randomly built row of H (all rows of H are i.i.d., see Section 3.2.1). Then,

$$\mathbb{E}[\mathcal{N}] = \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n} \left(\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\}\right)^{n(1-R)}. \tag{3.8}$$

Given an integer point $\mathbf{x} \in \mathbb{Z}^n \smallsetminus p\mathbb{Z}^n$, its *support* is the set

$$\operatorname{Supp}(\mathbf{x}) = \{i \in \{1, 2, \ldots, n\} : x_i \neq 0\}.$$

Suppose that $|\operatorname{Supp}(\mathbf{x})| = s > 0$, for some integer $1 \leq s \leq n$. We define *support* of the random vector $\mathbf{h} = (h_1, h_2, \ldots, h_n)$ the set of indices of the Δ coordinates of $\mathbf{h}$ that are not deterministically equal to 0. If $I = \{i \in \operatorname{Supp}(\mathbf{x}) \cap \operatorname{Supp}(\mathbf{h})\}$, we have

$$\begin{aligned} \mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\} &= \mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p \mid |I| = 0\}\mathcal{P}\{|I| = 0\} \\ &\quad + \mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p \mid |I| \neq 0\}\mathcal{P}\{|I| \neq 0\} \\ &= 1 \cdot \mathcal{P}\{|I| = 0\} + \frac{1}{p} \cdot \mathcal{P}\{|I| \neq 0\} \\ &\leq \mathcal{P}\{|I| = 0\} + \frac{1}{p}. \end{aligned} \tag{3.9}$$

There are two different situations:

- If $1 \leq s \leq n - \Delta$,

$$\begin{aligned}
\mathcal{P}\{|I| = 0\} &= \frac{\binom{n-s}{\Delta}}{\binom{n}{\Delta}} \\
&= \frac{n-\Delta}{n} \cdot \frac{n-1-\Delta}{n-1} \cdots \frac{n-s+1-\Delta}{n-s+1} \\
&= \left(1 - \frac{\Delta}{n}\right) \cdot \left(1 - \frac{\Delta}{n-1}\right) \cdots \left(1 - \frac{\Delta}{n-s+1}\right) \\
&\leq \left(1 - \frac{\Delta}{n}\right)^s \\
&= \left(1 - \frac{\beta \ln n}{n}\right)^s \qquad (3.10) \\
&\leq \frac{1}{n^{\beta s/n}}; \qquad (3.11)
\end{aligned}$$

(3.10) comes from the hypothesis on Δ, while (3.11) comes from the fact that

$$\begin{aligned}
&\left(1 - \frac{\beta \ln n}{n}\right)^s \leq \frac{1}{n^{\beta s/n}} \\
\Leftrightarrow &\ln\left(1 - \frac{\beta \ln n}{n}\right) \leq \ln \frac{1}{n^{\beta/n}} = -\frac{\beta \ln n}{n},
\end{aligned}$$

which is true because

$$\ln(1 - x) \leq -x \text{ for all } x < 1.$$

Note that for n big enough $\ln(1 - \beta \ln n/n)$ is well-defined and $0 < \beta \ln n/n < 1$.

- If instead $n - \Delta < s \leq n$, $\mathcal{P}\{|I| = 0\} = 0$.

Therefore, if n is big enough, putting together (3.8), (3.9) and what we have just shown, recalling that $p = n^\lambda$, we get

$$\begin{aligned}
\mathbb{E}[\mathcal{N}] &= \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n} \left(\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\}\right)^{n(1-R)} \\
&\leq \sum_{s=1}^{n-\Delta} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})| = s}} \left(\frac{1}{n^{\beta s/n}} + \frac{1}{n^\lambda}\right)^{n(1-R)} \qquad (3.12) \\
&\quad + \sum_{s=n-\Delta+1}^{n} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})| = s}} \left(\frac{1}{n^\lambda}\right)^{n(1-R)}. \qquad (3.13)
\end{aligned}$$

First of all, let us show that (3.13) goes to 0 when n tends to infinity.

$$\sum_{s=n-\Delta+1}^{n} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{B}\setminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})|=s}} \left(\frac{1}{n^\lambda}\right)^{n(1-R)}$$

$$\leq \sum_{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{B}} \left(\frac{1}{n^\lambda}\right)^{n(1-R)} \tag{3.14}$$

$$= |\mathbb{Z}^n\cap\mathcal{B}| \left(\frac{1}{n^\lambda}\right)^{n(1-R)} \tag{3.15}$$

$$\leq \operatorname{Vol}(\mathcal{B}) \left(1+\frac{1}{2(1+\varepsilon)\sigma}\right)^n \left(\frac{1}{n^\lambda}\right)^{n(1-R)} \tag{3.16}$$

$$\sim \frac{(\sigma\sqrt{2\pi e}(1+\varepsilon))^n}{\sqrt{\pi n}} \left(1+\frac{1}{2(1+\varepsilon)\sigma}\right)^n \left(\frac{1}{n^\lambda}\right)^{n(1-R)} \tag{3.17}$$

$$= \frac{((1-\delta)(1+\varepsilon))^n}{\sqrt{\pi n}} \left(1+\frac{\sqrt{2\pi e}}{2(1+\varepsilon)(1-\delta)n^{\lambda(1-R)}}\right)^n. \tag{3.18}$$

Note that in (3.16) we have used Lemma 2.3, (3.17) follows by Lemma 2.6 and (3.18) from the fact that $\sigma = (1-\delta)\sigma_{\max}$. Now, one can show that the term in the big parenthesis is either asymptotic to a constant or, at worst, subexponential (i.e. asymptotic to $\exp(\mu n^\nu)$, for some constants μ and $0<\nu<1$); hence the dominating term in (3.18) is $((1-\delta)(1+\varepsilon))^n$. But $(1-\delta)(1+\varepsilon)<1$, thanks to condition (3.4):

$$(1-\delta)(1+\varepsilon)<1 \Leftrightarrow \varepsilon < \frac{1}{1-\delta}-1 = \frac{\delta}{1-\delta}.$$

This implies that (3.18) tends to 0 as n tends to infinity and the same holds for (3.13).

At this point, we only have to study the behavior of (3.12). In order to do it, we will separate the analysis into three subcases: let

$$1 < a < 1+\frac{2}{3+2\lambda(1-R)}$$

be a constant such that $a\lambda/\beta<1$. We will consider separately:

1. $1 \leq s \leq \lambda n/\beta$;
2. $\lambda n/\beta < s < a\lambda n/\beta$;
3. $a\lambda n/\beta \leq s \leq n-\Delta$.

Note that $a\lambda n/\beta$ is really less than $n-\Delta$ if n is big enough.

First case. $1 \leq s \leq \lambda n/\beta$ (that is, $\lambda \geq \beta s/n$). First of all, recall that $B_{\mathbf{c},n}(\rho)$ is the n-dimensional sphere of radius ρ, centred at $\mathbf{c}$. Observe that

$$\begin{aligned}
&|\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \setminus p\mathbb{Z}^n : |\operatorname{Supp}(\mathbf{x})| = s\}| \\
&\leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : |\operatorname{Supp}(\mathbf{x})| = s\}| \\
&\leq \binom{n}{s} |\mathbb{Z}^s \cap B_{\mathbf{y},s}(\sigma\sqrt{n}(1+\varepsilon))| \\
&\leq n^s |\mathbb{Z}^s \cap [-\sigma\sqrt{n}(1+\varepsilon), \ldots, \sigma\sqrt{n}(1+\varepsilon)]^s| \\
&\leq n^s (2\sigma\sqrt{n}(1+\varepsilon)+1)^s.
\end{aligned} \tag{3.19}$$

The restriction on s implies that

$$\frac{1}{n^{\beta s/n}} \geq \frac{1}{n^{\lambda}},$$

thus,

$$\begin{aligned}
&\sum_{s=1}^{\lfloor \lambda n/\beta \rfloor} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \setminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})| = s}} \left(\frac{1}{n^{\beta s/n}} + \frac{1}{n^{\lambda}} \right)^{n(1-R)} \\
&\leq \sum_{s=1}^{\lfloor \lambda n/\beta \rfloor} n^s (2\sigma\sqrt{n}(1+\varepsilon)+1)^s \left(\frac{2}{n^{\beta s/n}} \right)^{n(1-R)} \\
&\leq \sum_{s=1}^{\lfloor \lambda n/\beta \rfloor} \left(C_1 n^{\lambda(1-R)+3/2-\beta(1-R)} \right)^s,
\end{aligned}$$

where C_1 is a positive constant. The last inequality is obtained recalling that the noise variance per coordinate is fixed to be $\sigma = \sigma_{\max}(1-\delta) = n^{\lambda(1-R)}(1-\delta)/\sqrt{2\pi e}$. We conclude by pointing out that the previous sum is a geometric series and it tends to 0 because the exponent of n is negative, thanks to condition (3.3).

Second case. $\lambda n/\beta < s < a\lambda n/\beta$ (and $\beta s/n < a\lambda$). First of all, notice that, if we bound $\binom{n}{s}$ with 2^n instead of n^s in (3.19), we have

$$\begin{aligned}
&|\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \setminus p\mathbb{Z}^n : |\operatorname{Supp}(\mathbf{x})| = s\}| \\
&\leq 2^n (2\sigma\sqrt{n}(1+\varepsilon)+1)^s \\
&\leq 2^n (C_2 n^{1/2+\lambda(1-R)})^s,
\end{aligned}$$

where C_2 is a positive constant. This implies that

$$\begin{aligned}
&\sum_{s=\lfloor \lambda n/\beta \rfloor+1}^{\lfloor a\lambda n/\beta \rfloor} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{B}\smallsetminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})|=s}} \left(\frac{1}{n^{\beta s/n}}+\frac{1}{n^{\lambda}}\right)^{n(1-R)} \qquad (3.20)\\
&= \sum_{s=\lfloor \lambda n/\beta \rfloor+1}^{\lfloor a\lambda n/\beta \rfloor} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{B}\smallsetminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})|=s}} \left(\frac{1}{n^{\beta s/n}}\right)^{n(1-R)} \left(1+n^{\frac{\beta s}{n}-\lambda}\right)^{n(1-R)}\\
&= \sum_{s=\lfloor \lambda n/\beta \rfloor+1}^{\lfloor a\lambda n/\beta \rfloor} |\{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{B}\smallsetminus p\mathbb{Z}^n : |\operatorname{Supp}(\mathbf{x})|=s\}| \left(\frac{1}{n^{\beta(1-R)}}\right)^{s} \left(1+n^{\frac{\beta s}{n}-\lambda}\right)^{n(1-R)}\\
&\leq 2^n \left(1+n^{\lambda(a-1)}\right)^{n(1-R)} \cdot {C_2}^n \sum_{s=\lfloor \lambda n/\beta \rfloor+1}^{\lfloor a\lambda n/\beta \rfloor} \left(n^{1/2+\lambda(1-R)-\beta(1-R)}\right)^{s}.
\end{aligned}$$

Let $\gamma = 1/2+\lambda(1-R)-\beta(1-R)$. The summation is a (partial) geometric series and it is equal to:

$$\begin{aligned}
\frac{1-n^{\gamma(\lfloor a\lambda n/\beta \rfloor+1)}}{1-n^{\gamma}} - \frac{1-n^{\gamma(\lfloor \lambda n/\beta \rfloor+1)}}{1-n^{\gamma}} &= \frac{n^{\gamma(\lfloor \lambda n/\beta \rfloor+1)}-n^{\gamma(\lfloor a\lambda n/\beta \rfloor+1)}}{1-n^{\gamma}}\\
&\sim n^{\gamma(\lfloor \lambda n/\beta \rfloor+1)},
\end{aligned}$$

since γ is negative by hypothesis (3.3) and $a>1$.

This implies that (3.20) is bounded by a function which is asymptotic to

$$(2C_2)^n n^{n(\lambda(a-1)(1-R)+\gamma/n(\lfloor \lambda n/\beta \rfloor+1))} \leq (2C_2)^n n^{n(\lambda(a-1)(1-R)+\gamma\lambda/\beta)}, \qquad (3.21)$$

which goes to 0 if $(\lambda(a-1)(1-R)+\gamma\lambda/\beta)$ is negative. Let us check that this is true:

$$\begin{aligned}
\lambda(a-1)(1-R)+\gamma\lambda/\beta = \lambda(a-1)(1-R)+\left(\frac{1}{2}+\lambda(1-R)-\beta(1-R)\right)\frac{\lambda}{\beta} &< 0\\
\Leftrightarrow \beta > \frac{\lambda}{2-a}+\frac{1}{2(2-a)(1-R)}&.
\end{aligned}$$

This is true thanks to hypothesis (3.3) and the condition $a<1+2/(3+2\lambda(1-R))$, which imply

$$\lambda+\frac{3}{2(1-R)} > \frac{\lambda}{2-a}+\frac{1}{2(2-a)(1-R)}.$$

All of this allows us to conclude that (3.21) and (3.20) tend to 0 when n tends to infinity.

Third case. $a\lambda n/\beta \leq s \leq n - \Delta$. We have

$$\begin{aligned} &\sum_{s=\lfloor a\lambda n/\beta \rfloor + 1}^{n-\Delta} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} \smallsetminus p\mathbb{Z}^n \\ |\operatorname{Supp}(\mathbf{x})| = s}} \left(\frac{1}{n^{\beta s/n}} + \frac{1}{n^{\lambda}} \right)^{n(1-R)} \\ &\leq \sum_{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B}} \left(\frac{1}{n^{\lambda}} \right)^{n(1-R)} \left(1 + \frac{1}{n^{\lambda(a-1)}} \right)^{n(1-R)} \\ &= \left(|\mathbb{Z}^n \cap \mathcal{B}| \left(\frac{1}{n^{\lambda}} \right)^{n(1-R)} \right) \left(1 + \frac{1}{n^{\lambda(a-1)}} \right)^{n(1-R)} . \end{aligned} \tag{3.22}$$

We know that the left term is (asymptotically) bounded by (3.18) and goes to 0 exponentially. The right term is at most subexponential in n:

$$\left(1 + \frac{1}{n^{\lambda(a-1)}} \right)^{n(1-R)} \sim \exp\left(n^{1-\lambda(a-1)}(1-R) \right),$$

with $a > 1$ by hypothesis. Then, (3.22) is bounded by a quantity in which the dominating term is $((1+\varepsilon)(1-\delta))^n$, which goes to 0 as n tends to infinity.

This ends the proof of (3.6), which is enough to conclude that the theorem is true. □

3.3 A stronger result with constant degrees

The result of the previous section is already sufficient to affirm that random LDA lattices achieve Poltyrev capacity of the unconstrained AWGN channel under lattice decoding. Nevertheless, an interesting question arises: is the condition on the (logarithmically) growing row degree of the parity-check matrix strictly necessary? For binary LDPC codes to achieve capacity of the binary symmetric channel, the answer is known to be yes (see [Gal63, Mac99]), without possibility of change. Surprisingly, for LDA lattices that hypothesis can be bypassed and relaxed. Namely, in this section we show that Poltyrev capacity can be achieved by a random LDA ensemble whose parity-check equations have degree bounded from above by constants (cf. Theorem 3.2), even if the dimension is still taken to progressively tend to infinity.

Construction A is again applied to LDPC codes over $\mathbb{F}_p$ for some prime $p = n^{\lambda}$. The value λ is constant as before, but now some further conditions will be necessary. In other terms, we lose some flexibility in the choice of the prime p, but it still will remain appreciably small and in particular smaller than n for typical values of the LDPC code rate (see condition (3.56)).

The newest element in the proof of this result concerns the Tanner graphs of the LDPC codes at the root of our LDA ensemble. Namely, we are interested in some *expansion properties* of our family of random graphs. This will be properly treated in Section 3.3.2, but we would like to give now a simplified explanation of our argument.

3.3.1 Overview of the proof for constant degrees

Geometrically, we follow the same path of the proof of Theorem 3.1 (see also the beginning of Section 3.2). That is, after having defined a *decoding sphere* $\mathcal{B}$, centred at the channel output and containing (very probably) the channel input, we look for lattice points inside this sphere. If the sent point is there and is the only one, no decoding error occurs. More formally, we employ an average argument to show that the probability that this does not happen tends to 0 when n tends to infinity. We will end up to estimate the same summation as (3.7), depending on the probabilities that the integer points inside the decoding sphere also belong to the lattice.

Now, it was already clear in the proof of Theorem 3.1 that in the case of LDA lattices, because of the sparseness of the parity-check matrix, this probability depends on the *weight* of the integer point we deal with (intended as the number of its non-zero coordinates). A low-weight $\mathbf{x} \in \mathbb{Z}^n$ has a bigger probability of satisfying all the parity-check equations than a high-weight one. For now, we have computed this probability looking at every parity-check equation one by one and independently and we have derived the expressions in (3.12) and (3.13). The new strategy will be of examining all the parity-check equations at the same time. We will understand what that means thanks to the following example: consider the parity-check matrix

$$H = \begin{bmatrix} h_{1,1} & 0 & 0 & h_{1,4} & h_{1,5} & 0 \\ 0 & h_{2,2} & 0 & h_{2,4} & h_{2,5} & 0 \\ 0 & h_{3,2} & h_{3,3} & 0 & 0 & h_{3,6} \\ h_{4,1} & 0 & h_{4,3} & 0 & 0 & h_{4,6} \end{bmatrix},$$

in which the zero entries are fixed and the $h_{i,j}$'s are i.i.d. uniform random variables which take value in $\mathbb{F}_p$. Consider then a low-weight point $\mathbf{x} = (0, 0, 0, x_4, x_5, 0)$, with $x_4, x_5 \neq 0$. When we compute the product $H\mathbf{x}^T$, there are two possible situations:

1. First two parity-check equations: the intersection of their supports with the support of $\mathbf{x}$ is non-empty and the probability that each one of them is satisfied is equal to $1/p$, depending on the random choice of the $h_{i,j}$'s.

2. Second two parity-check equations: their support does not intersect the support of $\mathbf{x}$ and the probability that they are satisfied is 1, independently of the random values taken by the $h_{i,j}$'s.

Summarising,

$$\mathcal{P}\{H\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\} = \left(\frac{1}{p}\right)^2.$$

In particular, this probability is bigger than $(1/p)^4$, the probability corresponding to a full-weight $\mathbf{x} = (x_1, x_2, \ldots, x_6)$ with non-zero coordinates.

Generalising this argument, if $T_{\mathbf{x}}$ is the set of parity-check equations that are not trivially satisfied by $\mathbf{x}$, our average estimation will lead to show that the following

sum tends to 0 when n tends to infinity (cf. equation (3.60)):

$$\sum_{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n}\frac{1}{p^{|T_{\mathbf{x}}|}}=\sum_{t=1}^{n(1-R)}\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n\\|T_{\mathbf{x}}|=t}}\frac{1}{p^t}.$$

The key argument concerns the estimation of $|\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}|$ for every single $t\in\{1,2,\ldots,n(1-R)\}$. When $t=n(1-R)$, then everything substantially works like in (3.14). When t is smaller, one has to hope that the cardinality of $\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}$ is small enough to compensate the fact that $(1/p)^t>(1/p)^{n(1-R)}$, in order to still have a vanishing summation.

How to ensure that? We exploit some expansion properties of the Tanner graph associated with the parity-check matrices. Namely, we remark that the random graphs in our family are typically such that any "small" subset of variable nodes is linked to a "big enough" subset of check nodes (cf. Definition 3.3 and Lemma 3.3). This translates into the condition that, for any given value of the parameter t that we defined before, the support of an $\mathbf{x}$ such that $T_{\mathbf{x}}=t$ cannot be too big (cf. Lemma 3.2). Consequently, $|\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}|$ is small, too.

We have given the intuitive guideline that lies behind our proof. The rest of the section will present in much more detail and mathematical formalism:

1. The expansion properties of random graphs that we are interested in.
2. The random LDA ensemble that we use for the proof.
3. The proof itself, with all the similarities with the proof of Theorem 3.1 and the novelties due to the "expansion approach", as well.

3.3.2 Graph-theoretical tools

As we have already anticipated, the family of LDPC codes that we are interested in is characterised by Tanner graphs with somewhat non-standard expansion properties. It is now the time to formally state these properties and proving that random, big enough bipartite graphs satisfy them with very high probability.

Let $\mathcal{G}=(V,P,E)$ be an undirected bipartite graph; $V\cup P$ is its set of vertices and E its set of edges. Later, V and P will stand for the sets of variable and check nodes of the Tanner graph respectively. Let $|V|=n$ and $|P|=n(1-R)$, for some $0<R<1$. Parallel edges are accepted, that is, there might be two or more edges connecting the same two vertices.

Definition 3.2 (Neighbourhood)**.** *If S is a subset of vertices of a graph $\mathcal{G}$, its* neighbourhood $N(S)$ *is defined as the set of vertices of the graph that are incident to a vertex of S.*

In a bipartite graph $\mathcal{G}=(V,P,E)$, the neighbourhood $N(S)\subseteq P$ for every $S\subseteq V$ and, vice versa, $N(T)\subseteq S$ for every $T\subseteq P$. See Figure 3.1 for a simple example.

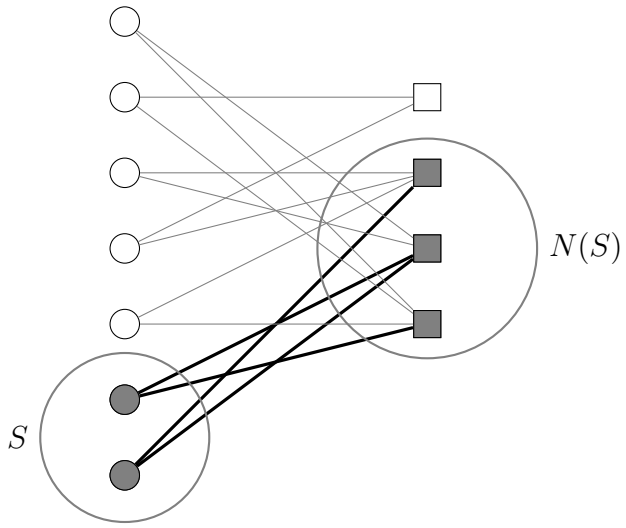

Figure 3.1: A bipartite graph with an example of neighbourhood of a subset of vertices. V is the set of round vertices, P is the set of square vertices. Observe that $S \subseteq N(N(S))$ and the inclusion is generally strict.

From now on, we will consider only graphs with the following variation of the *biregularity* property: the number of edges incident to any single vertex of V (resp. P) has constant cardinality Δ_V (resp. Δ_P). Consequently, the neighbourhood of any single vertex of V (resp. P) has cardinality at most Δ_V (resp. Δ_P). If the graph has no parallel edges, these cardinalities are exactly Δ_V and Δ_P and the graph is biregular, according to the standard definition. Denote by $\mathcal{F}(n, R, \Delta_V, \Delta_P)$ the family of graphs just defined. Note that biregularity implies the relations:

$$n \times \Delta_V = n(1-R) \times \Delta_P \quad \text{and} \quad \Delta_P = \frac{\Delta_V}{(1-R)}. \tag{3.23}$$

We are interested in some particular expansion properties of this kind of graphs. Thus we give the following definition:

Definition 3.3 ((α, A, β, B)-good graphs)**.** *Let α, A, β and B be four natural numbers such that*

$$A > \alpha \geq 1 \quad \textit{and} \quad \frac{1}{(1-R)} < \beta < \min\left\{\frac{2}{(1-R)}, B\right\}. \tag{3.24}$$

Let ε and ϑ be two small fixed positive constants ($0 < \vartheta, \varepsilon < 1$). We say that a graph

of $\mathcal{F}(n, R, \Delta_V, \Delta_P)$ *is* (α, A, β, B)-good *if*

1. *For every* $S \subseteq V$ *such that* $|S| \leq \lceil \varepsilon n \rceil$, *then* $|N(S)| \geq A|S|$. (3.25)
2. *For every* $S \subseteq V$ *such that* $|S| \leq \left\lceil \frac{n(1-R)}{2\alpha} \right\rceil$, *then* $|N(S)| \geq \alpha|S|$. (3.26)
3. *For every* $T \subseteq P$ *such that* $|T| \leq \frac{n(1-R)}{2}$, *then* $|N(T)| \geq \beta|T|$. (3.27)
4. *For every* $T \subseteq P$ *such that* $|T| \leq \vartheta n(1-R)$, *then* $|N(T)| \geq B|T|$. (3.28)

The four conditions above mean in quantitatively different ways that all "small" subsets of V or P have "big enough" sets of neighbours. Observe that the definition of an (α, A, β, B)-good graph implicitly depends on the choice of ε and ϑ. This will not lead to any ambiguity, since these constants will always be clearly and explicitly fixed any time we will deal with these graphs.

Before going on, we would like to put in evidence a direct consequence of the previous conditions:

Lemma 3.2. *Let* $\mathcal{G} \in \mathcal{F}(n, R, \Delta_V, \Delta_P)$ *be an* (α, A, β, B)*-good graph. The following statements hold true for every* $S \subseteq V$*:*

1. *If* $|N(S)| < A\lceil \varepsilon n \rceil$, *then* $|S| \leq |N(S)|/A$. (3.29)
2. *If* $|N(S)| < n(1-R)/2$, *then* $|S| \leq |N(S)|/\alpha$. (3.30)
3. *If* $|N(S)| \geq n(1-R)/2$, *then* $|S| \leq \beta|N(S)| - n(\beta(1-R) - 1)$. (3.31)
4. *If* $|N(S)| \geq (1-\vartheta)n(1-R)$, *then* $|S| \leq B|N(S)| - n(B(1-R) - 1)$. (3.32)

Proof. We will only prove the second and third statement and we leave to the reader the task of deducing (3.29) from (3.30) and (3.32) from (3.31). This will not require a big effort, since the remaining proofs are substantially identical to the given ones, after a minimal change of the parameters.

Let us start proving (3.30); in order to do this, we suppose that $|S| > |N(S)|/\alpha$ and we will argue that this implies $|N(S)| \geq n(1-R)/2$. Our hypothesis is equivalent to $|N(S)| < \alpha|S|$. Then (3.26) implies that $|S| > \lceil n(1-R)/2\alpha \rceil$. In particular, this means that there exists some $S' \subseteq S$ such that $|S'| = \lceil n(1-R)/2\alpha \rceil$. We can apply (3.26) and obtain $|N(S')| \geq \alpha|S'| = \alpha\lceil n(1-R)/2\alpha \rceil$. But $N(S') \subseteq N(S)$ by definition and therefore

$$|N(S)| \geq |N(S')| \geq \alpha|S'| = \alpha \left\lceil \frac{n(1-R)}{2\alpha} \right\rceil \geq \frac{n(1-R)}{2}.$$

For the proof of (3.31), let $T = N(S)$ and let $T^c = P \setminus T$. The hypothesis says that $|T^c| \leq n(1-R)/2$ and (3.27) implies that $|N(T^c)| \geq \beta|T^c|$. Note that $S \subseteq (V \setminus N(T^c))$, hence

$$|S| \leq n - |N(T^c)| \leq n - \beta|T^c| = n - \beta(n(1-R) - |T|) = n(\beta(1-R) - 1) + \beta|T|.$$

□

Let $h(\cdot)$ be the *binary entropy function*, already introduced in Definition 2.28. We have set the necessary notation to state the following lemma:

Lemma 3.3. *Let $n, \Delta_V \in \mathbb{N}$, with $\Delta_V \geq 3$. Let $0 < R < 1$ and let $\mathcal{G}$ be a graph in $\mathcal{F}(n, R, \Delta_V, \Delta_P)$, chosen uniformly at random in the family. Fix $\alpha, A, \beta, B \in \mathbb{N}$ satisfying* (3.24) *and let ε and ϑ be two positive constants such that*

$$0 < \varepsilon < \frac{(1-R)(\Delta_V - A - 1)}{A(\Delta_V - 2 + R)}, \tag{3.33}$$

$$0 < \vartheta < \frac{\Delta_V - (B+1)(1-R)}{B(1-R)(\Delta_V - 2 + R)}. \tag{3.34}$$

If

$$\begin{aligned} \Delta_V > \max \Bigg\{ & \frac{h\left(\frac{1-R}{2\alpha}\right) + 1 - R}{h\left(\frac{1-R}{2\alpha}\right) - \frac{1}{2}h\left(\frac{1-R}{\alpha}\right)}, R + 2\alpha, A + 1, \frac{h(\varepsilon) + (1-R)h\left(\frac{A\varepsilon}{1-R}\right)}{h(\varepsilon) - \frac{A\varepsilon}{1-R}h\left(\frac{1-R}{A}\right)}, \\ & \frac{1 - R + h\left(\frac{\beta(1-R)}{2}\right)}{1 - \frac{\beta(1-R)}{2}h\left(\frac{1}{\beta(1-R)}\right)}, \frac{(2+\beta R)(1-R)}{2 - \beta(1-R)}, (1-R)(B+1), \\ & \frac{(1-R)h(\vartheta) + h(B\vartheta(1-R))}{h(\vartheta) - B\vartheta(1-R)h\left(\frac{1}{B(1-R)}\right)} \Bigg\}, \end{aligned} \tag{3.35}$$

then

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ is not } (\alpha, A, \beta, B)\text{-good}\} = 0.$$

Remark: the proof of the previous lemma is inspired by the similar argument that can be found in [Bas81]. Nevertheless, we have derived it from scratch and adapted to our setting.

Proof. First of all, let us order the set V (putting it in bijection with $\{1, 2, \ldots, n\}$) and the set P (in bijection with $\{1, 2, \ldots, n(1-R)\}$); let us also order the set E of edges and call $e_1, e_2, \ldots, e_{\Delta_V}$ the edges linked to the first element of V, $e_{\Delta_V+1}, e_{\Delta_V+2}, \ldots, e_{2\Delta_V}$ the edges linked to the second element of V and so on. At the same time, call $f_1, f_2, \ldots, f_{\Delta_P}$ the edges linked to the first element of P, $f_{\Delta_P+1}, f_{\Delta_P+2}, \ldots, f_{2\Delta_P}$ the edges linked to the second element of P and so on. Then, a graph is determined by a permutation of $\{1, 2, \ldots, n\Delta_V\}$ that assigns to every e_m one of the f_l.

By the union bound,

$$\begin{aligned} \mathcal{P}\{\mathcal{G} \text{ is not } (\alpha, A, \beta, B)\text{-good}\} \leq & \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.26)}\} && (3.36) \\ & + \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.25)}\} && (3.37) \\ & + \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.27)}\} && (3.38) \\ & + \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.28)}\}. && (3.39) \end{aligned}$$

We will separately evaluate these four probabilities, which corresponds to counting the number of permutations of $\{1, 2, \ldots, n\Delta_V\}$ that do not guarantee the expansion properties.

Estimation of (3.36)**.**

$$
\begin{aligned}
&\mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.26)}\} \\
&\quad = \mathcal{P}\{\exists S \subseteq V : |S| \leq \lceil n(1-R)/2\alpha \rceil \text{ and } |N(S)| < \alpha|S|\} \\
&\quad \leq \sum_{\substack{S \subseteq V \\ 1 \leq |S| \leq \lceil n(1-R)/2\alpha \rceil}} \mathcal{P}\{|N(S)| < \alpha|S|\} \\
&\quad \leq \sum_{\substack{S \subseteq V \\ 1 \leq |S| \leq \lceil n(1-R)/2\alpha \rceil}} \sum_{\substack{T \subseteq P \\ |T| = \alpha|S|}} \mathcal{P}\{N(S) \subseteq T\} \\
&\quad = \sum_{\substack{S \subseteq V \\ 1 \leq |S| \leq \lceil n(1-R)/2\alpha \rceil}} \binom{n(1-R)}{\alpha|S|} \binom{\alpha|S|\Delta_P}{|S|\Delta_V} \Big/ \binom{n\Delta_V}{|S|\Delta_V} \\
&\quad = \sum_{s=1}^{\lceil n(1-R)/2\alpha \rceil} \binom{n}{s} \binom{n(1-R)}{\alpha s} \binom{\alpha s \Delta_P}{s\Delta_V} \Big/ \binom{n\Delta_V}{s\Delta_V} \\
&\quad = n \binom{n(1-R)}{\alpha} \binom{\alpha\Delta_P}{\Delta_V} \Big/ \binom{n\Delta_V}{\Delta_V} + \\
&\qquad + \sum_{s=2}^{\lceil n(1-R)/2\alpha \rceil} \binom{n}{s} \binom{n(1-R)}{\alpha s} \binom{\alpha s \Delta_P}{s\Delta_V} \Big/ \binom{n\Delta_V}{s\Delta_V} \\
&\quad \leq C_1 n^{(1+\alpha-\Delta_V)} + \sum_{s=2}^{\lceil n(1-R)/2\alpha \rceil} \binom{n}{s} \binom{n(1-R)}{\alpha s} \binom{\alpha s \Delta_P}{s\Delta_V} \Big/ \binom{n\Delta_V}{s\Delta_V}, \qquad (3.40)
\end{aligned}
$$

where C_1 is a constant that depends on α, Δ_V and R. Now, let $s = \zeta n$; we have

$$
\begin{aligned}
\binom{n(1-R)}{\alpha s} &= \binom{n(1-R)}{\frac{\zeta\alpha}{(1-R)} n(1-R)}, \\
\binom{\alpha s \Delta_P}{s\Delta_V} &= \binom{\alpha\zeta n \Delta_P}{\frac{(1-R)}{\alpha} \alpha\zeta n \Delta_P}
\end{aligned}
$$

and, by Lemma 2.5,

$$
\begin{aligned}
(3.40) &\leq C_1 n^{(1+\alpha-\Delta_V)} \qquad (3.41) \\
&\quad + C_2 \sum_{s=2}^{\lceil n(1-R)/2\alpha \rceil} 2^{n(-(\Delta_V-1)h(\zeta) + (1-R)h(\alpha\zeta/(1-R)) + (\Delta_V \alpha\zeta/(1-R))h((1-R)/\alpha))}, \qquad (3.42)
\end{aligned}
$$

for some new constant C_2, also recalling that $\Delta_P = \Delta_V/(1-R)$ (cf. (3.23)). Let

$$
\gamma(\zeta) = -(\Delta_V - 1)h(\zeta) + (1-R)h\left(\frac{\alpha\zeta}{1-R}\right) + \frac{\Delta_V \alpha\zeta}{1-R} h\left(\frac{1-R}{\alpha}\right)
$$

be the coefficient of n in the exponential function in (3.42). A simple computation shows that the derivative of γ is

$$\gamma'(\zeta) = -(\Delta_V - 1)\log_2\left(\frac{1}{\zeta} - 1\right) + \alpha\log_2\left(\frac{1-R}{\alpha\zeta} - 1\right) + \frac{\alpha\Delta_V}{1-R}h\left(\frac{1-R}{\alpha}\right),$$

while the second derivative is

$$\gamma''(\zeta) = \frac{\Delta_V - 1}{\zeta(1-\zeta)} - \frac{\alpha(1-R)}{\zeta(1-R-\alpha\zeta)}.$$

Recalling that ζ is limited to the range $2/n \leq \zeta \leq (1-R)/2\alpha$, one can easily check that $\gamma''(\zeta) > 0$ (the fact that $\Delta_V > R + 2\alpha$ by (3.35) is needed).

The positivity of $\gamma''(\zeta)$ implies that $\gamma'(\zeta)$ is increasing. A simple computation shows that

$$\lim_{n\to\infty} \gamma'\left(\frac{2}{n}\right) = -\infty.$$

Hence, for big enough n, $\gamma'(2/n) < 0$ and its monotonicity implies that either it is always negative or it has exactly one zero. In the first case, $\gamma(\zeta)$ is decreasing and

$$\max_{2/n\leq\zeta\leq(1-R)/2\alpha} \gamma(\zeta) = \gamma\left(\frac{2}{n}\right); \tag{3.43}$$

in the second case,

$$\max_{2/n\leq\zeta\leq(1-R)/2\alpha} \gamma(\zeta) = \max\left\{\gamma\left(\frac{2}{n}\right), \gamma\left(\frac{1-R}{2\alpha}\right)\right\};$$

but $\gamma((1-R)/2\alpha)$ is constant and negative (it can be easily verified, consequence of the fact that

$$\Delta_V > \frac{1 - R + h\left(\frac{1-R}{2\alpha}\right)}{h\left(\frac{1-R}{2\alpha}\right) - \frac{1}{2}h\left(\frac{1-R}{\alpha}\right)} \tag{3.44}$$

by condition (3.35)), while

$$\lim_{n\to\infty} \gamma\left(\frac{2}{n}\right) = 0.$$

This implies that (3.43) is always true when n is big enough. Then, we can deduce that

$$\begin{aligned}(3.41) + (3.42) \leq & C_1 n^{(1+\alpha-\Delta_V)} + C_2 n 2^{n\gamma(2/n)} \\ \leq & C_1 n^{(1+\alpha-\Delta_V)} + C_3 n\binom{n}{2}\binom{n(1-R)}{2\alpha}\binom{2\alpha\Delta_P}{2\Delta_V}\Big/\binom{n\Delta_V}{2\Delta_V}, \end{aligned} \tag{3.45}$$

where C_3 is a new constant and the last inequality is a consequence of Lemma 2.5.

Finally,

$$(3.45) \leq C_1 n^{(1+\alpha-\Delta_V)} + C_4 n^{1+2(1+\alpha-\Delta_V)},$$

where, again, C_4 is a constant, different from C_3.

Now, notice that the condition $\Delta_V > A+1$ (cf. (3.35)) implies that $\Delta_V > \alpha+2$, because $A \geq \alpha+1$ by (3.24). Hence

$$1 + 2(1 + \alpha - \Delta_V) \leq 1 + 2(1 + \alpha - \alpha - 2) = -1 < 0.$$

Thus,

$$\lim_{n\to\infty} C_1 n^{(1+\alpha-\Delta_V)} + C_4 n^{1+2(1+\alpha-\Delta_V)} = 0$$

and as a consequence

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.26)}\} = 0, \tag{3.46}$$

too.

Estimation of (3.37)**.** We want to show that

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.25)}\} = 0. \tag{3.47}$$

We will not treat all the details here, since this estimation is very similar to the previous one. Namely, the argument is the very same as before, while in the explicit computations one has just to substitute α with A and only consider sums for s going from 1 to $\lceil \varepsilon n \rceil$. One obtains:

$$\mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.25)}\} \leq \sum_{s=1}^{\lceil \varepsilon n \rceil} \binom{n}{s} \binom{n(1-R)}{As} \binom{As\Delta_P}{s\Delta_V} \Big/ \binom{n\Delta_V}{s\Delta_V}.$$

Again, Lemma 2.5 allows us to approximate binomial coefficients with exponential function and the same exponent coefficient $\gamma(\zeta)$ as before is found (again, with A instead of α). Now, the choice of ε small enough (cf. (3.33)) guarantees that the second derivative of $\gamma(\zeta)$ is positive, while the fact that

$$\Delta_V > \frac{h(\varepsilon) + (1-R)h\left(\frac{A\varepsilon}{1-R}\right)}{h(\varepsilon) - \frac{A\varepsilon}{1-R}h\left(\frac{1-R}{A}\right)}$$

(cf. condition (3.35)) is the analogue of (3.44) for determining the maximum of $\gamma(\zeta)$ in the range $1 \leq \zeta \leq \varepsilon$. Nothing more is required to conclude the analysis of (3.37) and its limit is 0 when n tends to infinity.

Estimation of (3.38)**.** We can now move our attention to the estimation of the

following probability:

$$\begin{aligned}
&\mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.27)}\} \\
&\quad = \mathcal{P}\{\exists T \subseteq P : |T| \leq n(1-R)/2 \text{ and } |N(T)| < \beta|T|\} \\
&\quad \leq \sum_{\substack{T\subseteq P \\ 1\leq|T|\leq n(1-R)/2}} \mathcal{P}\{|N(T)| < \beta|T|\} \\
&\quad \leq \sum_{\substack{T\subseteq P \\ 1\leq|T|\leq n(1-R)/2}} \sum_{\substack{S\subseteq V \\ |S|=\beta|T|}} \mathcal{P}\{N(T) \subseteq S\} \\
&\quad = \sum_{\substack{T\subseteq P \\ 1\leq|T|\leq n(1-R)/2}} \binom{n}{\beta|T|}\binom{\Delta_V\beta|T|}{|T|\Delta_P}\Big/\binom{n\Delta_V}{|T|\Delta_P} \\
&\quad = \sum_{t=1}^{\lfloor n(1-R)/2\rfloor} \binom{n(1-R)}{t}\binom{n}{\beta t}\binom{\Delta_V\beta t}{t\Delta_P}\Big/\binom{n\Delta_V}{t\Delta_P} \\
&\quad \leq C_5 n^{1+\beta-\Delta_V/(1-R)} + \sum_{t=2}^{\lfloor n(1-R)/2\rfloor} \binom{n(1-R)}{t}\binom{n}{\beta t}\binom{\Delta_V\beta t}{t\Delta_P}\Big/\binom{n\Delta_V}{t\Delta_P}, \qquad (3.48)
\end{aligned}$$

for some constant C_5.

The strategy that we will employ is the same as the one employed previously. So, define $t = \tau n(1-R)$. Thus, recalling that $\Delta_V = \Delta_P(1-R)$ and applying Lemma 2.5, we can write

$$(3.48) \leq C_5 n^{1+\beta-\Delta_V/(1-R)} + C_6 \sum_{t=2}^{\lfloor n(1-R)/2\rfloor} 2^{n\varphi(\tau)}, \qquad (3.49)$$

where C_6 is a new constant and the function $\varphi(\cdot)$ is equal to

$$\varphi(\tau) = -(\Delta_V - 1 + R)h(\tau) + h(\tau\beta(1-R)) + \tau\beta(1-R)\Delta_V h\left(\frac{1}{\beta(1-R)}\right).$$

What can we say about φ? We will adopt the same approach that we have used with γ. It is easy to compute that

$$\begin{aligned}
\varphi'(\tau) = &-(\Delta_V - 1 + R)\log_2\left(\frac{1}{\tau} - 1\right) + \beta(1-R)\log_2\left(\frac{1}{\beta(1-R)\tau} - 1\right) \\
&+ \beta(1-R)\Delta_V h\left(\frac{1}{\beta(1-R)}\right)
\end{aligned}$$

and

$$\varphi''(\tau) = \frac{\Delta_V - 1 + R}{\tau(1-\tau)} - \frac{\beta(1-R)}{\tau(1-\beta\tau(1-R))}.$$

This second derivative is positive if and only if

$$\tau < \frac{\Delta_V - 1 + R - \beta(1-R)}{\beta(1-R)(\Delta_V - 2 + R)}. \qquad (3.50)$$

This holds true because

$$\tau \leq \frac{1}{2} < \frac{\Delta_V - 1 + R - \beta(1-R)}{\beta(1-R)(\Delta_V - 2)},$$

thanks to the fact that $\Delta_V > (2+\beta R)(1-R)/(2-\beta(1-R))$ by condition (3.35). Therefore $\varphi'(\tau)$ is increasing and, since

$$\lim_{n\to\infty} \varphi'\left(\frac{2}{n(1-R)}\right) = -\infty,$$

we deduce that $\varphi'(\tau)$ has at most one zero (for big enough n). Similarly to the case of γ,

$$\max_{2/n(1-R)\leq\tau\leq 1/2} \varphi(\tau) = \max\left\{\varphi\left(\frac{2}{n(1-R)}\right), \varphi\left(\frac{1}{2}\right)\right\}.$$

On the one hand,

$$\lim_{n\to\infty} \varphi\left(\frac{2}{n(1-R)}\right) = 0,$$

on the other one, $\varphi(1/2)$ is negative and constant, because

$$\Delta_V > \frac{1 - R + h\left(\frac{\beta(1-R)}{2}\right)}{1 - \frac{\beta(1-R)}{2} h\left(\frac{1}{\beta(1-R)}\right)} \tag{3.51}$$

thanks to condition (3.35). This implies that for big enough n

$$\max_{2/n(1-R)\leq\tau\leq 1/2} \varphi(\tau) = \varphi\left(\frac{2}{n(1-R)}\right).$$

Applying this result to (3.49) and then using Lemma 2.5, we obtain

$$\begin{aligned}
(3.49) &\leq C_5 n^{1+\beta-\Delta_V/(1-R)} + C_6 \sum_{t=2}^{\lfloor n(1-R)/2\rfloor} 2^{n\varphi(2/n(1-R))} \\
&\leq C_5 n^{1+\beta-\Delta_V/(1-R)} + C_7 n \binom{n(1-R)}{2}\binom{n}{2\beta}\binom{2\beta\Delta_V}{\Delta_V/(1-R)}\Big/\binom{n\Delta_V}{2\Delta_V/(1-R)} \\
&\leq C_5 n^{1+\beta-\Delta_V/(1-R)} + C_8 n^{1+2(1+\beta-\Delta_V/(1-R))},
\end{aligned}$$

where, of course, C_7 and C_8 are some new constants.

The exponent $1+2(1+\beta-\Delta_V/(1-R))$ is negative because our parameters are such that $\Delta_V/(1-R) > B+1 \geq \beta+2$ (cf. (3.35) and (3.24)). Thus, we can argue that

$$1 + 2(1+\beta-\Delta_V/(1-R)) \leq 1 + 2(1+\beta-\beta-2) = -1.$$

This allows us to conclude that

$$\lim_{n\to\infty} C_5 n^{1+\beta-\Delta_V/(1-R)} + C_8 n^{1+2(1+\beta-\Delta_V/(1-R))} = 0$$

and

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.27)}\} = 0. \tag{3.52}$$

Estimation of (3.39)**.** As one can guess, we will take inspiration from the previous case to analyse this last probability, in the same way as the analysis of (3.36) has established the pattern to estimate (3.37). If we replace β by B and make the sum over t go to $\lfloor \vartheta n(1-R) \rfloor$ instead of $\lfloor n(1-R)/2 \rfloor$, the same argument as before tells that

$$\mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.28)}\} \leq C_9 n^{1+\beta-\Delta_V/(1-R)} + C_{10} \sum_{t=2}^{\lfloor n(1-R)/2 \rfloor} 2^{n\varphi(\tau)}, \tag{3.53}$$

for some appropriate constants C_9 and C_{10} and

$$\varphi(\tau) = -(\Delta_V - 1 + R)h(\tau) + h(\tau\beta(1-R)) + \tau\beta(1-R)\Delta_V h\left(\frac{1}{\beta(1-R)}\right)$$

is the same $\varphi(\cdot)$ as before with B instead of β.

We can conclude just with the same argument of the previous case, noticing the fact that

$$\vartheta < \frac{\Delta_V - (B+1)(1-R)}{B(1-R)(\Delta_V - 2 + R)}$$

(cf. (3.50) and (3.34)) guarantees that $\varphi''(\tau)$ is positive. Moreover, $\tau \leq \vartheta$ and $\varphi(\vartheta)$ can be shown to be a negative constant thanks to the fact that

$$\Delta_V > \frac{(1-R)h(\vartheta) + h(B\vartheta(1-R))}{h(\vartheta) - B\vartheta(1-R)h\left(\frac{1}{B(1-R)}\right)}$$

(see (3.35) and compare it to its analogue (3.51)). Hence

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ does not satisfy (3.28)}\} = 0. \tag{3.54}$$

Conclusion. Finally, we are only left with putting together what we have shown till now: (3.46), (3.47), (3.52) and (3.54) imply that

$$\lim_{n\to\infty} \mathcal{P}\{\mathcal{G} \text{ is not } (\alpha, A, \beta, B)\text{-good}\} = 0,$$

which is what we were looking for and the proof is concluded. □

3.3.3 The new random LDA lattice ensemble

Let $\mathcal{G}$ be any (α, A, β, B)-good graph, in the sense specified by Definition 3.3. A priori, it may contain parallel edges. Let us identify them and call again $\mathcal{G}$ the new graph, with at most one edge between any two vertices. It is still bipartite and (α, A, β, B)-good and represents also the Tanner graph of a binary LDPC code. Let H be the *binary* parity-check matrix with Tanner graph $\mathcal{G}$. Let p be a prime

number and let us associate a label to every edge of $\mathcal{G}$, independently of each other and chosen uniformly at random in the set $\{0, 1, \ldots, p-1\}$ of the representatives of classes modulo p. Equivalently, we are choosing a parity-check matrix $\mathbf{H}$ with entries in $\mathbb{F}_p$. Let $C = C[n,k]_p \subseteq \mathbb{F}_p$ be the k-dimensional linear code over the finite field $\mathbb{F}_p$ with parity-check matrix $\mathbf{H}$. The actual Tanner graph of $\mathbf{H}$ is a subgraph of $\mathcal{G}$ which may differ from the whole graph $\mathcal{G}$ if some random coordinates are chosen to be equal zero. Observe also that, for the same reason, the rate of C may be greater than $R = k/n$.

Definition 3.4 (Skeleton matrix)**.** *In this context, we call the binary matrix H the* skeleton *of the random matrix* $\mathbf{H}$.

Every $i \in P$ represents a parity-check equation of C and a row of $\mathbf{H}$, while a $j \in V$ is a coordinate of a codeword $\mathbf{c} \in C$. If Δ_V is small with respect to n, the code is an LDPC code and column (resp. row) weights (or degrees) are bounded from above by Δ_V (resp. Δ_P).

3.3.4 LDA lattices achieve Poltyrev capacity with constant parity-check matrix row degree

Theorem 3.2. *Let n be a positive integer number and let $0 < R < 1$. Let $p = n^\lambda$ be a prime number for some $\lambda > 0$ and let $\sigma^2 = p^{2(1-R)}(1-\delta)^2/2\pi e$ be the AWG noise variance per dimension, for some $0 < \delta < 1$. Let α, A, β, B be four natural numbers that obey* (3.24)*. Moreover, let ε and ϑ be two positive constants that satisfy conditions* (3.33) *and* (3.34) *and suppose also that*

$$\vartheta < \frac{\delta}{B(1-R)}. \tag{3.55}$$

Finally, let $\Delta_V \in \mathbb{N}$ be a constant, big enough to satisfy (3.35)*. If*

$$\lambda > \max\left\{\frac{1}{2(\alpha-1+R)}, \frac{3}{2(A-1+R)}, \frac{1}{B(1-R)-1}\right\}, \tag{3.56}$$

then there exists a Poltyrev-capacity-achieving family of LDA lattices $\Lambda_n = C_n + p\mathbb{Z}^n$ such that the rate of C_n is at least R and the row degree in the parity-check matrix of C_n is at most $\Delta_V/(1-R)$ (this means that $\sigma^2 = \sigma^2_{\max}(1-\delta)^2$, according to (3.1)*).*

Proof. In order to prove the theorem, we evaluate the probability of decoding error, averaged over all LDA lattices built at random following the model described in Section 3.3.3, which employs (α, A, β, B)-good graphs. Geometrically, the proof follows the same path as the one of Theorem 3.1 and some analogies can be found between them.

So, let $\mathcal{G}$ be a bipartite graph chosen at random in $\mathcal{F}(n, R, \Delta_V, \Delta_P)$; we know by Lemma 3.3 that, if n is big enough, $\mathcal{G}$ is an (α, A, β, B)-good graph with very high probability. Thus, we suppose that $\mathcal{G}$ is chosen to be (α, A, β, B)-good. If $\mathcal{G}$ was not

good, we would simply treat it as a "bad" choice and change it with another random graph of $\mathcal{F}(n, R, \Delta_V, \Delta_P)$. Let $\Lambda = C + p\mathbb{Z}^n$ be a random LDA lattice associated to $\mathcal{G}$, and suppose we use Λ for communication over the AWGN channel. Notice that the dimension n will change and be sent to infinity all along the proof. Nevertheless, we avoid the notation C_n and Λ_n, the indices being redundant and not necessary for a full comprehension.

First of all, because of the lattice symmetry and of the independence between random noise and channel input, we can suppose that the point of Λ sent over the channel is $\mathbf{0}$. The AWG noise vector is $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ and the channel output is $\mathbf{y} = \mathbf{w}$.

Let us consider the sphere $\mathcal{B} = B_{\mathbf{y},n}(\sigma\sqrt{n}(1+\xi)) \subseteq \mathbb{R}^n$ centred at $\mathbf{y}$, of radius $\sigma\sqrt{n}(1+\xi)$, with $\xi > 0$, for some small ξ chosen such that

$$\xi < \frac{\delta}{1-\delta}; \tag{3.57}$$

the latter condition has exactly the same utility in the proof as (3.4) for Theorem 3.1 (cf. (3.64)). Lemma 2.2 states that, when n tends to infinity, the point $\mathbf{0}$ is inside the sphere with probability tending to 1. Moreover, Lemma 3.1 states that, again with probability tending to 1, all non-zero vectors of $p\mathbb{Z}^n$ in $\mathcal{B}$ will be further away from the received vector $\mathbf{y}$ than the transmitted (zero) vector itself.

We are therefore only concerned with ensuring that the decoder does not return $\hat{\mathbf{x}} \not\equiv \mathbf{0} \bmod p$. To this end let us introduce the random variable $\mathcal{N}$ that counts the number of lattice points inside the sphere and not belonging to $p\mathbb{Z}^n$. Our goal is to show that

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{N}] = 0,$$

which is enough to prove the result.

For every integer point $\mathbf{x} \in \mathcal{B} \cap \mathbb{Z}^n$, let $X_{\mathbf{x}}$ be the random variable defined by

$$X_{\mathbf{x}} = \begin{cases} 1, & \text{if } \mathbf{x} \in \Lambda \\ 0, & \text{if } \mathbf{x} \notin \Lambda \end{cases}.$$

We have

$$\mathcal{N} = \sum_{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n} X_{\mathbf{x}}$$

and

$$\mathbb{E}[\mathcal{N}] = \sum_{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{x} \in \Lambda\}. \tag{3.58}$$

If $\mathbf{H}$ is the parity-check matrix of the p-ary code C associated with Λ, an integer point $\mathbf{x}$ belongs to Λ if and only if $\mathbf{H}\mathbf{x}^T \equiv \mathbf{0}^T \bmod p$. Remember that $\mathbf{H}$ is a sparse matrix so, if some of the coordinates of $\mathbf{x}$ are equal to 0 (in $\mathbb{F}_p$), some parity-check equations will be trivially satisfied, no matter what its random coefficients are.

More precisely, let $\mathbf{h}$ be any row of the binary skeleton matrix H of $\mathbf{H}$ (cf. Definition 3.4). We define the *support* of $\mathbf{x}$ (resp. $\mathbf{h}$) to be the set of indices that

correspond to non-zero coordinates of $\mathbf{x}$ (resp. $\mathbf{h}$). We have that if the supports of $\mathbf{x}$ and $\mathbf{h}$ have empty intersection then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\} = 1$. On the other hand, if the supports of $\mathbf{x}$ and $\mathbf{h}$ intersect in at least one coordinate, then we see that $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv 0 \bmod p\} = 1/p$ (recall that we do not consider $\mathbf{x} \in p\mathbb{Z}^n$).

For a fixed $\mathbf{x}$, if $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n(1-R)}$ are the rows of $\mathbf{H}$, let

$$T_{\mathbf{x}} = \{i \in \{1, 2, \ldots, n(1-R)\} : \mathrm{Supp}(\mathbf{h}_i) \cap \mathrm{Supp}(\mathbf{x}) \neq \emptyset\}. \tag{3.59}$$

Note that $T_{\mathbf{x}}$ is the neighbourhood in the Tanner graph $\mathcal{G}$ of $\mathrm{Supp}(\mathbf{x})$. Now let $t = |T_{\mathbf{x}}|$ be the number of parity-check equations that are not trivially satisfied by $\mathbf{x}$; then,

$$\mathcal{P}\{\mathbf{x} \in \Lambda\} = \left(\frac{1}{p}\right)^t,$$

because the coefficients that define the parity-check equations are chosen independently and therefore the events $\{\mathbf{x}$ satisfies the i-th parity-check$\}_{i=1,\ldots,n(1-R)}$ are independent. This means that, by (3.58)

$$\mathbb{E}[\mathcal{N}] = \sum_{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n} \frac{1}{p^{|T_{\mathbf{x}}|}} = \sum_{t=1}^{n(1-R)} \sum_{\substack{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}| = t}} \frac{1}{p^t} \tag{3.60}$$

In order to clarify our strategy, let us start considering only the $\mathbf{x}$'s such that $|T_{\mathbf{x}}| = n(1-R)$. The partial summation corresponding to them is:

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}| = n(1-R)}} \frac{1}{p^{n(1-R)}} \leq \frac{|\mathbb{Z}^n \cap \mathcal{B}|}{p^{n(1-R)}} \tag{3.61}$$

$$\leq \frac{\mathrm{Vol}(\mathcal{B})}{p^{n(1-R)}} \left(1 + \frac{1}{2(1+\xi)\sigma}\right)^n \tag{3.62}$$

$$\sim \frac{1}{\sqrt{\pi n}} \left((1+\xi)\frac{\sqrt{2\pi e}\sigma}{p^{(1-R)}}\right)^n \left(1 + \frac{1}{2(1+\xi)\sigma}\right)^n, \tag{3.63}$$

where we have used Lemma 2.3 in (3.62) and Lemma 2.6 for the asymptotic expression in (3.63).

One can show that the term in the right parentheses is at worst subexponential in n (i.e. asymptotic to $\exp(an^\mu)$, for some constants a and $0 < \mu < 1$); hence the dominating term in (3.63) is the central one. We aim to prove that $\lim_{n\to\infty} \mathbb{E}[\mathcal{N}] = 0$; at least for the special (but most frequent) case of the $\mathbf{x}$'s such that $|T_{\mathbf{x}}| = n(1-R)$, we are done, since (3.63) goes to 0 when n tends to infinity. Indeed, the base of the dominating exponential can be made smaller than 1 with a proper choice of ξ: recalling that $\sigma = \sigma_{\max}(1-\delta) = p^{(1-R)}(1-\delta)/\sqrt{2\pi e}$ for an LDA lattice, we have

$$\begin{aligned}(1+\xi)\frac{\sqrt{2\pi e}\sigma}{p^{(1-R)}} < 1 &\Leftrightarrow \sigma = \sigma_{\max}(1-\delta) < \frac{\sigma_{\max}}{1+\xi} \\ &\Leftrightarrow \xi < \frac{\delta}{1-\delta},\end{aligned} \tag{3.64}$$

which is the condition that we imposed in (3.57).

What happens in the more general case, when $|T_{\mathbf{x}}| < n(1-R)$? A priori, the power of p in (3.63) is not sufficient to guarantee the convergence to 0; this clearly happens, for example, when $|T_{\mathbf{x}}|$ is a constant with respect to n. But in that case, the inequality

$$\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=t<n(1-R)}} \frac{1}{p^t} \leq \frac{|\mathbb{Z}^n\cap\mathcal{B}|}{p^t},$$

i.e. the analogue of (3.61), is not precise enough and it does not take into account the fact that the integer points in $\mathcal{B}$ with such a small $|T_{\mathbf{x}}|$ are much less than $|\mathbb{Z}^n \cap \mathcal{B}|$. We need a more detailed analysis, based on an efficient estimation of $|\{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}| = t\}|$, which exploits the properties of (α, A, β, B)-good graphs.

We begin by cutting the summation in (3.60) into four different parts, depending on the constants ε and ϑ that we have already fixed in the statement of the theorem. Namely, we will consider the following four cases:

1. $t < A\lceil\varepsilon n\rceil$;

2. $A\lceil\varepsilon n\rceil \leq t < n(1-R)/2$;

3. $n(1-R)/2 \leq t < (1-\vartheta)n(1-R)$;

4. $(1-\vartheta)n(1-R) \leq t \leq n(1-R)$.

First case. Lemma 3.2 tells that, according to the choice of ε and since $\mathcal{G}$ is (α, A, β, B)-good, $t < A\lceil\varepsilon n\rceil$ implies that $|\operatorname{Supp}(\mathbf{x})| \leq t/A$. As a consequence, recalling that $B_{\mathbf{c},n}(\rho)$ is the n-dimensional sphere centred at $\mathbf{c}$ of radius ρ, we have that $\forall t < A\lceil\varepsilon n\rceil$

$$\begin{aligned}
&|\{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}| = t\}| && (3.65)\\
&\leq |\{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}) \smallsetminus p\mathbb{Z}^n : |\operatorname{Supp}(\mathbf{x})| \leq t/A\}| \\
&\leq \binom{n}{\lfloor t/A\rfloor} |\mathbb{Z}^{\lfloor t/A\rfloor} \cap B_{\mathbf{y},\lfloor t/A\rfloor}(\sigma\sqrt{n}(1+\xi))| \\
&\leq n^{t/A} |\mathbb{Z}^{\lfloor t/A\rfloor} \cap \mathcal{C}_{\lfloor t/A\rfloor}(2\sigma\sqrt{n}(1+\xi))| \\
&\leq n^{t/A}(2\sigma\sqrt{n}(1+\xi)+1)^{t/A}, && (3.66)
\end{aligned}$$

where $\mathcal{C}_{\lfloor t/A\rfloor}(2\sigma\sqrt{n}(1+\xi))$ is the $\lfloor t/A\rfloor$-dimensional cube of edge $2\sigma\sqrt{n}(1+\xi)$.

The first part of the summation (3.60) that we consider is:

$$\begin{aligned}
&\sum_{t=1}^{A\lceil \varepsilon n\rceil -1} \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=t}} \frac{1}{p^t} \\
&= \sum_{t=1}^{A\lceil \varepsilon n\rceil -1} \frac{|\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}|}{p^t} \\
&\leq \sum_{t=1}^{A\lceil \varepsilon n\rceil -1} \frac{n^{t/A}(2\sigma\sqrt{n}(1+\xi)+1)^{t/A}}{p^t} \\
&< \sum_{t=1}^{A\lceil \varepsilon n\rceil -1} \left(Dn^{(3/(2A)+\lambda(1-R)/A-\lambda)}\right)^t, \qquad (3.67)
\end{aligned}$$

where D is a constant term. The last inequality holds because $p = n^\lambda$ and $\sigma < \sigma_{\max} = n^{\lambda(1-R)}/\sqrt{2\pi e}$ (see (3.1)). Now, (3.67) is a geometric series and its limit for n going to infinity is 0 if the exponent of n is negative; it is, thanks to hypothesis (3.56).

Second case. Now $A\lceil \varepsilon n\rceil \leq t < n(1-R)/2$; this case is almost identical to the previous one. We will just change a little bit some estimations and adapt them to the fact that the t's we consider are at least linear in n.

We apply again Lemma 3.2 to say that for this range of t the corresponding $\mathbf{x}$'s are such that $|\operatorname{Supp}(\mathbf{x})| \leq t/\alpha$. Then, we do exactly the same estimations done from (3.65) to (3.66), with the only change that we replace A by α and bound $\binom{n}{\lfloor t/\alpha\rfloor}$ by 2^n instead of $n^{t/\alpha}$; we get:

$$|\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}| \leq 2^n(2\sigma\sqrt{n}(1+\xi)+1)^{t/\alpha}$$

and

$$\begin{aligned}
\sum_{t=A\lceil \varepsilon n\rceil}^{\lfloor n(1-R)/2\rfloor} \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=t}} \frac{1}{p^t} &= \sum_{t=A\lceil \varepsilon n\rceil}^{\lfloor n(1-R)/2\rfloor} \frac{|\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : |T_{\mathbf{x}}|=t\}|}{p^t} \\
&\leq \sum_{t=A\lceil \varepsilon n\rceil}^{\lfloor n(1-R)/2\rfloor} \frac{2^{nt/t}(2\sigma\sqrt{n}(1+\xi)+1)^{t/\alpha}}{p^t} \\
&\leq \sum_{t=A\lceil \varepsilon n\rceil}^{\lfloor n(1-R)/2\rfloor} \frac{2^{t/A\varepsilon}(2\sigma\sqrt{n}(1+\xi)+1)^{t/\alpha}}{p^t} \\
&< \sum_{t=A\lceil \varepsilon n\rceil}^{\lfloor n(1-R)/2\rfloor} \left(En^{(1/2\alpha+\lambda(1-R)/\alpha-\lambda)}\right)^t,
\end{aligned}$$

for some constant E. Similarly to the first case, it is condition (3.56) that ensures that the exponent of n is negative and the sum converges to 0.

Third case. Now we deal with the partial sum that corresponds to the range $n(1-R)/2 \le t < (1-\vartheta)n(1-R)$. Let us call $\gamma = \beta t - n(\beta(1-R)-1)$. The third statement of Lemma 3.2 implies that any $\mathbf{x}$ such that $|T_{\mathbf{x}}| = t$ also requires $|\operatorname{Supp}(\mathbf{x})| \le \gamma$. Then

$$\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=t}} \frac{1}{p^t} \tag{3.68}$$

$$= \sum_{\substack{T\subseteq P \\ |T|=t}} \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ T_{\mathbf{x}}=T}} \frac{1}{p^t}$$

$$= \sum_{\substack{T\subseteq P \\ |T|=t}} |\{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n : T_{\mathbf{x}}=T\}| \frac{1}{p^t}$$

$$\le \sum_{\substack{T\subseteq P \\ |T|=t}} \frac{|\mathbb{Z}^{\lfloor\gamma\rfloor}\cap B_{\mathbf{0},\lfloor\gamma\rfloor}(\sigma\sqrt{n}(1+\xi))|}{p^t}$$

$$\le \binom{n(1-R)}{t} \frac{\operatorname{Vol}(B_{\mathbf{0},\lfloor\gamma\rfloor}(\sigma\sqrt{n}(1+\xi))}{p^t} \left(1+\frac{\sqrt{\gamma}}{2\sigma\sqrt{n}(1+\xi)}\right)^{\gamma} \tag{3.69}$$

$$\sim \binom{n(1-R)}{t} \frac{1}{\sqrt{\pi\lfloor\gamma\rfloor}} \left(\frac{n}{\lfloor\gamma\rfloor}\right)^{\gamma/2} ((1-\delta)(1+\xi))^{\gamma} n^{\lambda((1-R)\gamma-t)} \left(1+\frac{\sqrt{\gamma}}{2\sigma\sqrt{n}(1+\xi)}\right)^{\gamma} \tag{3.70}$$

$$\le 2^{n(1-R)} \left(\frac{n}{\lfloor\gamma\rfloor}\right)^{\gamma/2} ((1-\delta)(1+\xi))^{\gamma} n^{\lambda((1-R)\gamma-t)} \left(1+\frac{1}{2\sigma(1+\xi)}\right)^{\gamma}, \tag{3.71}$$

where (3.69) is justified by Lemma 2.3, while the asymptotic expression in (3.70) comes from Lemma 2.6. Now, observe that $n(1-R)/2 \le t$ also implies

$$\frac{n}{\gamma} = \frac{1}{\beta t/n - (\beta(1-R)-1)} \le \frac{1}{1-\beta(1-R)/2}.$$

So, for some positive constant F,

$$(3.71) \lesssim F^{\gamma} n^{\lambda((1-R)\gamma-t)}.$$

The symbol $\lesssim$ is used for brevity and indicates the "asymptotic inequality" relation: we write $f(n) \lesssim g(n)$ is there exists $h(n)$ such that $f(n) \le h(n)$ and $h(n) \sim g(n)$.

Going on with our computation, we have

$$\begin{aligned}\sum_{t=\lceil n(1-R)/2\rceil}^{\lfloor(1-\vartheta)n(1-R)\rfloor}\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n\\ |T_{\mathbf{x}}|=t}}\frac{1}{p^t}&\lesssim\sum_{t=\lceil n(1-R)/2\rceil}^{\lfloor(1-\vartheta)n(1-R)\rfloor}F^\gamma n^{\lambda(1-R)\gamma-\lambda t}\\&=\sum_{t=\lceil n(1-R)/2\rceil}^{\lfloor(1-\vartheta)n(1-R)\rfloor}F^\gamma n^{\lambda(\beta(1-R)-1)(t-n(1-R))}\\&\le\sum_{t=\lceil n(1-R)/2\rceil}^{\lfloor(1-\vartheta)n(1-R)\rfloor}F^n n^{-n\lambda\vartheta(\beta(1-R)-1)(1-R)}.\end{aligned}$$

The whole summation tends to 0, because the exponent of the dominating term is negative (recall that $\beta(1-R)>1$ by condition (3.24)).

Fourth case. We are left to study the case of the biggest t's in our range: $(1-\vartheta)n(1-R)<t\le n(1-R)$. The same estimations done from (3.68) to (3.70) still hold true if we substitute β with B in the definition of γ:

$$\gamma=Bt-n(B(1-R)-1);$$

then, bounding the binomial coefficient by

$$\binom{n(1-R)}{t}=\binom{n(1-R)}{n(1-R)-t}\le(n(1-R))^{n(1-R)-t}\le n^{n(1-R)-t},$$

we obtain

$$\begin{aligned}&\sum_{t=\lceil(1-\vartheta)n(1-R)\rceil}^{n(1-R)}\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n\\ |T_{\mathbf{x}}|=t}}\frac{1}{p^t}\\&\qquad\lesssim\sum_{t=\lceil(1-\vartheta)n(1-R)\rceil}^{n(1-R)}n^{n(1-R)-t}\left(\frac{n}{\lfloor\gamma\rfloor}\right)^{\gamma/2}((1-\delta)(1+\xi))^\gamma\cdot\\&\qquad\qquad\cdot n^{\lambda((1-R)\gamma-t)}\left(1+\frac{1}{2\sigma(1+\xi)}\right)^\gamma.\end{aligned}$$

The addends of the summation are made of factors of three different kinds:

- the term $((1+\xi)(1-\delta)\sqrt{n/\lfloor\gamma\rfloor})^\gamma$ is exponential in n;
- the term $n^{n(1-R)-t+\lambda((1-R)\gamma-t)}=n^{(n(1-R)-t)(1-\lambda(B(1-R)-1))}$ is superexponential in n when t is close to $(1-\vartheta)n(1-R)$, while it is only polynomial in n when t is close to $n(1-R)$; in both cases, the fact that $1-\lambda(B(1-R)-1)<0$ by condition (3.56) implies $n^{(n(1-R)-t)(1-\lambda(B(1-R)-1))}\le 1$.
- the right term $(1+1/2\sigma(1+\xi))^\gamma$ is subexponential in n.

Now, observe that

$$\gamma = Bt - n(B(1-R)-1) \geq B(1-\vartheta)n(1-R) - n(B(1-R)-1) = n - nB(1-R)\vartheta.$$

Therefore, the closer to 0 is ϑ, the closer to n is γ. In other words, if ϑ is "small enough", then $\sqrt{n/\lfloor\gamma\rfloor}$ is very close to 1 and the product $(1+\xi)(1-\delta)\sqrt{n/\lfloor\gamma\rfloor}$ is smaller than 1 (in case, up to a proper choice of ξ, that can be taken as small as we need); condition (3.55) is precisely required to ensure that and qualitatively explicits the previous "small enough".

Finally, our summation becomes

$$\sum_{t=\lceil(1-\vartheta)n(1-R)/2\rceil}^{n(1-R)} \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B})\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=t}} \frac{1}{p^t} \lesssim \sum_{t=\lceil(1-\vartheta)n(1-R)\rceil}^{n(1-R)} \left(\sqrt{\frac{n}{\lfloor\gamma\rfloor}}(1-\delta)(1+\xi)\right)^{(1-\vartheta)n(1-R)} \cdot \cdot \left(1+\frac{1}{2\sigma(1+\xi)}\right)^n,$$

that tends to 0 thanks to the exponential behaviour of the main factor.

Conclusion. Putting together all the four cases and recalling (3.60), we get that $\lim_{n\to\infty}\mathbb{E}[\mathcal{N}] = 0$ and this is the end of the proof of Theorem 3.2.

□

As a concluding remark, let us specify something about the values of all the parameters involved in the proof of the previous theorem. What we would like to point out is that these constants are typically very reasonable and in particular that the required value of Δ_V is generally small, what makes it suitable for practical implementations. So, let us show this through an example; suppose that we target a value of λ which is smaller than 1, say $\lambda = 1/2$, and suppose that $R = 0.6$ (a value of the LDPC code rate that is experimented with in Chapter 5). The hypothesis of Theorem 3.2 are satisfied if

- $\alpha = 2$,
- $A = 4$,
- $\beta = 3$,
- $B = 8$

and the conditions of Lemma 3.3 only imply $\Delta_V \geq 9$.

Chapter 4

Finite lattice constellations

In this chapter we go a bit further with respect to the results presented in Chapter 3. We address the problem of attaining the capacity of the AWGN channel with lattice codes (or constellations, see Definition 2.16) and a lattice decoder and we abandon the concept of *Poltyrev capacity* that has lead to Theorem 3.1 and Theorem 3.2.

We already mentioned in Section 2.3.2 that Linder et al. showed in [LSZ93] that the capacity of the AWGN channel is achievable with spherical lattice codes under ML decoding and that subsequently the challenge has been of reaching the same goal with a *lattice decoder*, which is non-optimal for finite lattice constellations.

This goal was reached by Erez and Zamir [EZ04] and more recently Ordentlich and Erez [OE12] and Ling and Belfiore [LB13] have revisited this result, giving different proofs with respect to Erez and Zamir. [EZ04] and [OE12] are based on the so-called Voronoi constellations and the MLAN (Modulo-Lattice Additive Noise) channel; Ling and Belfiore, instead, propose a solution based on a "pseudo-Gaussian" distribution on the AWGN channel set of inputs (which is of course a lattice). We will briefly summarise their ideas in Section 4.1, putting in evidence their strategies, their strength and their disadvantages. This will also allow us to fix in detail the setting in which we will work.

In Section 4.2, we propose our demonstration that there exists a random Construction A ensemble that achieves the capacity of the AWGN channel (cf. Theorem 4.1). This lattice family is very similar to the one considered by Ordentlich and Erez and our proof presents some advantages with respect to the already existing ones. Moreover, we have also been motivated in this redemonstration work because our argument seems to be the most suitable to be adapted to LDA lattices. This is done in Section 4.3.

4.1 Previous work

4.1.1 Voronoi constellations and MLAN channel: Erez and Zamir's approach

We recall that a lattice decoder simply looks for the closest lattice point to the received channel output (cf. Section 2.3.2). This strategy is suboptimal with respect to MAP decoding of a finite constellation (or ML, when all codewords are a priori equiprobable), since the decision regions do not take into account the boundaries of the *shaping region*. This means that the decoder output is always a lattice point, but it may not lie in the constellation and thus not be a codeword. Until Erez and Zamir's paper [EZ04], previously existing lattice coding schemes based on a "pure" lattice decoder could not achieve the capacity of the AWGN channel [dB75, Pol94, Loe97] and were stuck to the value $1/2\log_2(\mathrm{SNR})$ for the maximum achievable rate. Then, Erez and Zamir proposed a new encoding and decoding scheme that solves this problem. To that end, they needed a certain number of ingredients, which we will briefly introduce and describe in this section:

1. Voronoi constellations.

2. Modulo-Lattice Additive Noise channel.

3. "Good" lattices for the covering and the channel coding problem (see also Section 2.4).

Voronoi constellations

Let $\Lambda \subseteq \Lambda_f \subseteq \mathbb{R}^n$ be two lattices. Λ and Λ_f are commonly called two *nested* lattices and give rise to the following definition, proposed for the first time by Conway and Sloane [CS82, For89]:

Definition 4.1 (Voronoi constellation)**.** *The* Voronoi lattice code *or* Voronoi constellation $\mathcal{C}$ *generated by* Λ_f *and* Λ *is simply a lattice code (cf. Definition 2.16) made of points of* Λ_f *for which the shaping region is* $\mathcal{V}(\Lambda)$*:*

$$\mathcal{C} = \Lambda_f \cap \mathcal{V}(\Lambda). \tag{4.1}$$

Equivalently, it is the set of representatives of the equivalence classes of Λ_f/Λ *that have minimum norm.*

Λ_f is called the *fine* lattice (what clarifies the choice of the index f) and Λ the *coarse* or *shaping* lattice. An example of Voronoi constellation in dimension 2 is given in Figure 4.1.

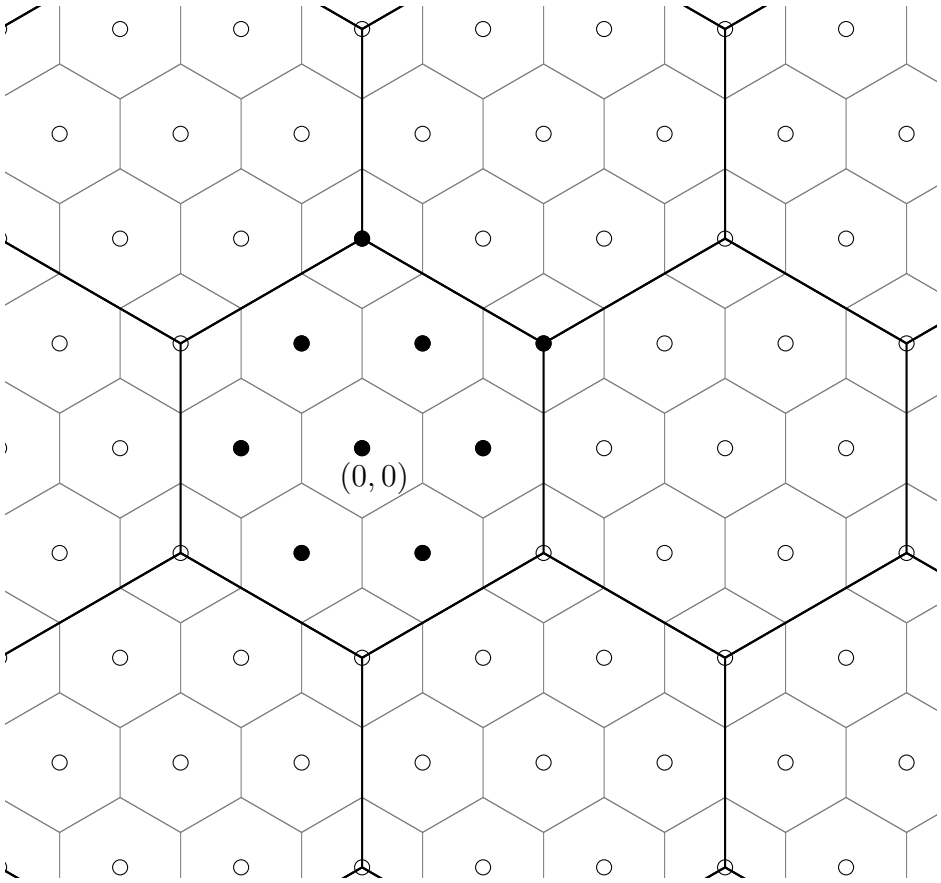

Figure 4.1: A Voronoi constellation of the hexagonal lattice in dimension 2. Filled and empty circles are the points of Λ_f. Small hexagons represent their Voronoi regions. Big hexagons are the Voronoi regions of $\Lambda = 3\Lambda_f$. The filled circles are the $3^2 = 9$ points of the constellation Λ_f/Λ, contained in $\mathcal{V}(\Lambda)$.

Encoding, decoding and the MLAN channel

As we have already pointed out, lattice decoding on its own seemed not to be sufficient to achieve the capacity of the Gaussian channel. That is the reason why Erez and Zamir proposed to base their transmission scheme on a transformation of the AWGN channel, firstly derived in [ESZ05]. This is called *Modulo-Lattice Additive Noise* (or *MLAN*) channel and involves the lattice Λ whose Voronoi region is the shaping region of the finite constellation. When this lattice is accurately chosen, one can show that the transition from AWGN to MLAN channel is information preserving at any SNR, asymptotically in the lattice dimension n. Erez and Zamir also proved in detail that reliable transmission of data is possible with this construction when the dimension n goes to infinity. Furthermore, they provided a precise analysis of the error probability exponents.

Let us introduce with some more precision the MLAN channel and the related transmission strategy. We will give many details, in order to properly understand

the construction, but we will skip all technical proofs (with the exception of Lemma 4.1) and refer the reader to [EZ04] for them.

The first key ingredient for this construction is a random variable $\mathbf{u}$ called a *dither*, which is uniformly distributed over the shaping region $\mathcal{V}(\Lambda)$. Every instance of the dither is known and shared by the sender and the receiver. Dithering is commonly used in lattice quantisation for source coding [For92].

The second ingredient is given by the following definition (see also [EZ04]):

Definition 4.2 (Wiener coefficient)**.** *Let* $\mathbf{x}$ *be the random variable that represents the AWGN channel input and let* $\mathbf{y} = \mathbf{x} + \mathbf{w}$ *be its random output, then the* Wiener coefficient *is*

$$\alpha = \arg\min_{\beta \in \mathbb{R}} \mathbb{E}[||\mathbf{x} - \beta\mathbf{y}||^2],$$

The minimum in the previous formula is usually called Minimum Mean Squared Error *and the Wiener coefficient is also called* MMSE coefficient*.*

Lemma 4.1. *Let* $\mathbb{E}[||\mathbf{x}||^2] = nP$ *and suppose that the AWG noise variance per dimension is* σ^2 *(i.e.* $w_i \sim \mathcal{N}(0, \sigma^2)$*). Then*

$$\alpha = \frac{P}{P + \sigma^2}.$$

Proof. Note that we have

$$\begin{aligned}\mathbb{E}[||\mathbf{x} - \beta\mathbf{y}||^2] &= \mathbb{E}[||(1-\beta)\mathbf{x} - \beta\mathbf{w}||^2] \\ &= (1-\beta)^2\mathbb{E}[||\mathbf{x}||^2] + \beta^2\mathbb{E}[||\mathbf{w}||^2] - 2(1-\beta)\beta\mathbb{E}[\mathbf{x}\mathbf{w}^T]. \qquad (4.2)\end{aligned}$$

Now, the coordinate-wise independent distribution of the Gaussian noise implies

$$\mathbb{E}[||\mathbf{w}||^2] = \mathbb{E}[\mathbf{w}\mathbf{w}^T] = \sum_{i=1}^{n} \mathbb{E}[w_i^2] = \sum_{i=1}^{n} \mathrm{Var}(w_i) = n\sigma^2.$$

Moreover, since x_i and w_i are independent for every i and $\mathbb{E}[w_i] = 0$,

$$\mathbb{E}[\mathbf{x}\mathbf{w}^T] = \sum_{i=1}^{n} \mathbb{E}[x_i w_i] = \sum_{i=1}^{n} \mathbb{E}[x_i w_i] = \sum_{i=1}^{n} \mathbb{E}[x_i]\mathbb{E}[w_i] = 0.$$

We go on from (4.2) and we obtain

$$\mathbb{E}[||\mathbf{x} - \beta\mathbf{y}||^2] = (1-\beta)^2 nP + \beta^2 n\sigma^2 = (nP + n\sigma^2)\beta^2 - 2nP\beta + nP = f(\beta).$$

Minimising $f(\beta)$ with a standard first derivative argument gives the wanted result. □

The third ingredient for Erez and Zamir's result is the "modulo-Λ" operation:

Definition 4.3 (Reduction modulo a lattice)**.** *Any point* $\mathbf{y} \in \mathbb{R}^n$ *can be written as* $\mathbf{y} = \mathbf{x} + \mathbf{v}$*, for some* $\mathbf{x} \in \Lambda$ *and* $\mathbf{v} \in \mathcal{V}(\Lambda)$*. This decomposition is unique if we consider the proper version of* $\mathcal{V}(\Lambda)$ *(see the comments after Definition 2.6). Hence, we define*

$$\mathbf{y} \bmod \Lambda = \mathbf{y} - Q_\Lambda(\mathbf{y}) = \mathbf{v};$$

recall that $Q_\Lambda(\cdot)$ *is the lattice quantiser associated with* $\mathcal{V}(\Lambda)$ *of Definition 2.4.*

We are now ready to describe the transmission scheme:

1. Let $\mathbf{t} \in \Lambda_f/\Lambda$ be the coded message that we want to send (that would correspond to the AWGN channel input).

2. Let $\mathbf{u} \sim \mathrm{Unif}(\mathcal{V}(\Lambda))$ be the dither; the encoder submits to the AWGN channel the quantity
$$\mathbf{x} = [\mathbf{t} - \mathbf{u}] \bmod \Lambda.$$

3. The receiver obtains the AWGN channel output $\mathbf{y} = \mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is the Gaussian noise. He multiplies it by the Wiener coefficient α and obtains:
$$\mathbf{y}'' = \alpha\mathbf{y} + \mathbf{u}.$$

4. Finally, the receiver passes to the lattice decoder (with respect to Λ_f) the point $\mathbf{y}''$; this computes
$$\hat{\mathbf{t}} = Q_{\Lambda_f}(\mathbf{y}'') \bmod \Lambda \in \Lambda_f/\Lambda$$
which is the decoding scheme output.

Erez and Zamir proved that this transmission scheme is equivalent to the channel whose input is $\mathbf{t}$ and whose output is $\mathbf{y}' = [\mathbf{y}'' \bmod \Lambda]$. It is properly the MLAN channel and we notice that

$$\begin{aligned}
\mathbf{y}' &= [\alpha\mathbf{y} + \mathbf{u}] \bmod \Lambda \\
&= [\alpha(\mathbf{x} + \mathbf{w}) + \mathbf{u}] \bmod \Lambda \\
&= [\alpha([\mathbf{t} - \mathbf{u}] \bmod \Lambda) + \mathbf{w}) + \mathbf{u}] \bmod \Lambda \\
&= [\mathbf{t} - (1 - \alpha)(\mathbf{t} - \mathbf{u}) + \alpha\mathbf{w}] \bmod \Lambda.
\end{aligned}$$

We can say that the effective noise of the MLAN channel is the sum of two components: the scaled Gaussian noise $\alpha\mathbf{w}$ and the *self-noise* $-(1-\alpha)(\mathbf{t}-\mathbf{u})$. The dither $\mathbf{u}$ in the self-noise has the role of decorrelating the self-noise from the channel input $\mathbf{t}$.

The dither plays an important role in [EZ04] and [OE12], but it will turn out that it is not strictly necessary to achieve the same goals. Indeed, our objective in proving Theorem 4.1 is precisely to show that Voronoi constellations can achieve capacity even without dithering.

On the other hand, the multiplication by α turns out to be the real key idea to gain the "missing 1" in the formula $1/2\log_2(\mathrm{SNR})$ and reach the capacity. MMSE

scaling can be intuitively thought as a method of bringing the AWGN channel output as close as possible to the channel input, before starting the lattice decoder (see also Section 4.2.3). The combination of scaling by α and lattice decoding is sometimes referred to as *MMSE lattice decoding*.

As a final remark, we put in evidence that in this setting the concept of *lattice encoding and decoding* is somehow redefined and now involves the dither (i.e. a source of shared randomness), the Wiener coefficient and, moreover, some reductions modulo Λ, the shaping lattice. Nevertheless, from a computational point of view, a reduction modulo Λ corresponds to an instance of the lattice decoder with respect to Λ, while addition and subtraction of $\mathbf{u}$ or the multiplication by α are negligible. Hence this scheme is concretely based on two lattice decoders (with respect to Λ and Λ_f) and we still can speak of lattice encoding and decoding.

Erez and Zamir's results with "good" nested lattices

Erez and Zamir proved under some hypotheses that the MLAN channel and the AWGN channel have (asymptotically) the same capacity and that it can be achieved by a particular family of random lattice constellations [EZ04]. Namely, they need:

- the shaping lattice Λ to be good for the quantisation problem;
- the fine lattice Λ_f to be good for coding over the AWGN channel with lattice decoding.

The reader can refer to Section 2.4 for the different definitions of "goodness" for lattices.

The explicit construction that Erez and Zamir provide is based on Construction A, whose properties of goodness are studied in [ELZ05]. It is a Construction A over $\mathbb{Z}$, with a very large p (exponential in n). Moreover, their shaping lattice has the property of being "Rogers-good", which concretely means good for the covering problem. This implies goodness for quantisation (cf. [ELZ05]) and this requirement is needed to accurately establish the error probability exponents that the modulo-lattice scheme achieves.

Ordentlich and Erez's simpler proof

Ordentlich and Erez gave in [OE12] a simpler proof of these capacity results, which avoids the analysis of the error exponents and needs a much less technical approach. It is worth the effort of presenting the family of nested lattices that they consider, since it is substantially very similar to the one we use in Theorem 4.1: let G_f be the generating matrix of a linear code C_f over $\mathbb{F}_p$; suppose that G_f has dimension $k_f \times n$. This code gives rise to the fine lattice Λ_f through (a scaled) Construction A:

$$\Lambda_f = \gamma p^{-1} C_f + \gamma \mathbb{Z}^n,$$

where $\gamma = 2\sqrt{n\,\mathrm{SNR}}$ is chosen for some computational reasons.

The coarse lattice comes from a subcode of C_f: let C be the linear code generated by the first k rows of G_f; then $C \subseteq C_f$ and

$$\Lambda = \gamma p^{-1} C + \gamma \mathbb{Z}^n \subseteq \Lambda_f.$$

Of course, the family is a random ensemble and all the entries of G_f are i.i.d. random variables uniformly distributed over $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$. In order to make the probability of having full rank matrices to be close to 1, the hypothesis $k_f < \beta n$ for some $0 < \beta < 1$ is made. Indeed, it can be shown that

$$\mathcal{P}\{\mathrm{rk}(G_f) < k_f\} < p^{k_f - n} \tag{4.3}$$

(see [ELZ05] to justify the previous bound).

The last point that we would like to discuss is that Ordentlich and Erez fix $1/2n^{3/2} \leq p \leq n^{3/2}$. In the perspective of a practical application of these results, this is a substantial improvement with respect to the construction of Erez and Zamir (for which p is exponential in n). Nevertheless, they do not specify if powers of n other than $3/2$ (and possibly smaller) can be chosen, keeping the proof valid. In Theorem 4.1, our range of admittable primes will be wider: the power $3/2$ can be replaced by any positive constant.

4.1.2 Lattice Gaussian coding: Ling and Belfiore's approach

Before moving on to our own work, we report another recent result by Ling and Belfiore [LB13]. Their strategy for achieving the AWGN channel capacity abandons Voronoi constellations, involves only one lattice and works under the hypothesis that SNR > 3. Their constellation could be defined as "probabilistically finite", because the codebook is a lattice good for coding and the a priori probability for the codewords follows a discrete Gaussian distribution (and in particular is not uniform over some set anymore). In such a way, even if the constellation is infinite, high-energy codewords have a small probability of being transmitted and an average power constraint is respected. So, there is no need of a shaping lattice and, moreover, of the dither. On the contrary, the Wiener coefficient is still employed for adapting the channel output and obtaining a better lattice decoding; MMSE lattice decoding is proved to be equivalent to MAP decoding in this setting. Finally, a key tool in Ling and Belfiore's construction is the so-called *flatness factor* [LLBS12], that will be introduced later.

Discrete Gaussian distribution over a lattice

We report here with our notation the definition of the distribution on the lattice points that Ling and Belfiore have chosen for the channel a priori probabilities.

Take $\sigma > 0$ and consider the n-dimensional Gaussian density function centred at $\mathbf{c}$:

$$f_{\sigma,\mathbf{c}}(\mathbf{y}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(\frac{-||\mathbf{y} - \mathbf{c}||^2}{2\sigma^2}\right).$$

For simplicity, we write $f_\sigma(\mathbf{y}) = f_{\sigma,\mathbf{0}}(\mathbf{y})$. From this distribution, we derive the function

$$f_{\sigma,\Lambda}(\mathbf{y}) = \sum_{\mathbf{x}\in\Lambda} f_{\sigma,\mathbf{x}}(\mathbf{y}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\mathbf{x}\in\Lambda} \exp\left(\frac{-||\mathbf{y}-\mathbf{x}||^2}{2\sigma^2}\right),$$

which is periodic on a lattice Λ, in the sense that $f_{\sigma,\Lambda}(\mathbf{y}) = f_{\sigma,\Lambda}(\mathbf{y}+\mathbf{x})$ for all $\mathbf{x} \in \Lambda$. One can easily see that this function, if restricted to $\mathbb{R}^n/\Lambda$, is a probability density.

Definition 4.4 (Discrete Gaussian distribution)**.** *Let $f_{\sigma,\mathbf{c}}(\Lambda) = \sum_{\mathbf{x}\in\Lambda} f_{\sigma,\mathbf{c}}(\mathbf{x})$. The* discrete Gaussian distribution *over the lattice $\Lambda \subseteq \mathbb{R}^n$ with center $\mathbf{c} \in \mathbb{R}^n$ is the following distribution, taking values in $\mathbf{x} \in \Lambda$:*

$$D_{\Lambda,\sigma,\mathbf{c}}(\mathbf{x}) = \frac{f_{\sigma,\mathbf{c}}(\mathbf{x})}{f_{\sigma,\mathbf{c}}(\Lambda)}.$$

In the case of a shifted lattice $\Lambda - \mathbf{c} = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \mathbf{x} - \mathbf{c}, \exists \mathbf{x} \in \Lambda\}$, we define for every $\mathbf{x} \in \Lambda$:

$$D_{\Lambda-\mathbf{c},\sigma}(\mathbf{x}-\mathbf{c}) = \frac{f_\sigma(\mathbf{x}-\mathbf{c})}{f_{\sigma,\mathbf{c}}(\Lambda)}$$

and we have $D_{\Lambda-\mathbf{c},\sigma}(\mathbf{x}-\mathbf{c}) = D_{\Lambda,\sigma,\mathbf{c}}(\mathbf{x})$.

The flatness factor

Definition 4.5 (Flatness factor)**.** *Let $\sigma > 0$, let $\Lambda \subseteq \mathbb{R}^n$ be a lattice and let $\mathcal{F}$ be any fundamental region of Λ. The* flatness factor *is*

$$\epsilon_\Lambda(\sigma) = \frac{\max_{\mathbf{y}\in\mathcal{F}} |f_{\sigma,\Lambda}(\mathbf{y}) - 1/\operatorname{Vol}(\Lambda)|}{1/\operatorname{Vol}(\Lambda)}.$$

The flatness factor measures the maximum variation of $f_{\sigma,\Lambda}(\mathbf{y})$ and, namely, the definition implies that

$$f_{\sigma,\mathbf{c}}(\Lambda) \in [1-\epsilon_\Lambda(\sigma), 1+\epsilon_\Lambda(\sigma)] \frac{1}{\operatorname{Vol}(\Lambda)}.$$

If the flatness factor is small (see the conditions below and [LLBS12] for more detail), the lattice Gaussian distribution can be proved to behave like a continuous Gaussian distribution from many respects. This is intuitively an advantage, because it is known that continuous Gaussian distributions achieve the capacity of the AWGN channel.

The encoding and decoding scheme

Let the SNR be given by P/σ^2, for some average power constraint P and a fixed AWGN variance per dimension σ^2. Let Λ be a lattice which is good for unconstrained AWGN channel coding. Furthermore, Ling and Belfiore supposed that the codebook is given by a shifted version of the lattice $\Lambda - \mathbf{c}$, since many practical applications

require it. This will not change the geometric properties of the lattice, of course. As already alluded, the encoder maps the information bits to the points $\mathbf{y} = \mathbf{x} - \mathbf{c} \in \Lambda - \mathbf{c}$ and the latter follow the lattice Gaussian distribution of parameter σ_0

$$D_{\Lambda-\mathbf{c},\sigma_0}(\mathbf{y}) = \frac{e^{-\frac{||\mathbf{y}||^2}{2\sigma_0^2}}}{(\sqrt{2\pi}\sigma_0)^n f_{\sigma_0,\mathbf{c}}(\Lambda)}.$$

After addition of the Gaussian noise $\mathbf{w}$, the decoder multiplies the channel output by the coefficient $\alpha' = \sigma_0^2/(\sigma_0^2+\sigma^2)$ and performs lattice decoding of the product $\alpha'(\mathbf{y}+\mathbf{w})$. In this setting, Ling and Belfiore showed explicitly that if the flatness factor is "small", α' is asymptotically equal to the Wiener coefficient and this decoding strategy is asymptotically equivalent to MAP decoding. The decoder output is then

$$\hat{\mathbf{y}} = Q_{\Lambda-\mathbf{c}}(\alpha'(\mathbf{y}+\mathbf{w})).$$

Ling and Belfiore's analysis puts in evidence that under some hypotheses the decoding error probability tends exponentially fast to 0 with the lattice dimension growing to infinity. In particular, they need

- Λ to be good for coding (according to Definition 2.27);
- $\lim_{n\to\infty} \epsilon_\Lambda\left(\frac{\sigma_0^2}{\sqrt{\sigma_0^2+\sigma^2}}\right) = 0$ and $\epsilon_\Lambda\left(\frac{\sigma_0}{2}\right) < 1$;
- $\mathrm{SNR} > e$.

So, actually, capacity is not achieved for every SNR, even if the previous requirement is quite mild.

Moreover, if $\mathbf{y} \sim D_{\Lambda-\mathbf{c},\sigma_0}$, it can be shown that (cf. [MR04, Ban93])

$$\mathcal{P}\{||\mathbf{y}|| > \sqrt{2\pi n}\sigma_0\} < \frac{1+\epsilon_\Lambda(\sigma_0)}{1-\epsilon_\Lambda(\sigma_0)} 2^{-n}.$$

Since $\epsilon_\Lambda(\sigma_0)$ is lower and upper bounded by positive constants, we are allowed to think that the points of the lattice which do not lie inside a sphere of radius $\sqrt{2\pi n}\sigma_0$ are not transmitted. This is how lattice Gaussian distribution resembles a finite constellation. On the other hand, the reader my be interested in [ZF96] to learn how a uniform distribution over a Voronoi constellation with a quantisation-good shaping lattice can be considered asymptotically close to a Gaussian distribution.

As a conclusion, we also recall that Ling and Belfiore's techniques are applied in another work by the same authors et al. [LLBS12], in which they address the problem of transmitting information over the wiretap channel.

4.2 A new approach to achieve capacity

In this section we provide a new proof that there exists a random ensemble of lattices that achieves capacity under (MMSE) lattice decoding when $\mathrm{SNR} > 1$. Our result

would like to put together some advantages of the results presented till now in the chapter. At the same time, we try to overcome some less attractive aspects of them. Our proof will present some novelties and its main features are listed below here.

- Our lattice family is a random Construction A ensemble. It is similar to the family proposed by Ordentlich and Erez (cf. Section 4.1.1), the difference being that it is defined by a set of parity-check matrices rather than generator matrices (cf. Section 4.2.1).

- We still adopt the technique of Voronoi constellations.

- We do not need dithering anymore. This meets the purpose of Belfiore and Ling of avoiding the unpractical sharing of common randomness between the sender and the receiver. However, they pay the price of a randomised encoder. Our proof instead does not need lattice Gaussian distribution and we still have an a priori uniform distribution over the lattice constellation. Moreover, an explicit bijection exists that maps uncoded messages to constellation points. This is desirable when we think of practical implementations of our encoding and decoding scheme.

- We still rely on the idea of scaling the AWGN channel output by the Wiener coefficient, before performing lattice decoding. This enhances the strength of the decoder.

- We restrict our construction to the case SNR > 1. The reasons of this choice will be explained in Section 4.2.3.

- We decrease the size of the prime number needed for Construction A as a function of n, with respect to Ordentlich and Erez' construction (recall that they have $p \approx n^{3/2}$), still attaining capacity. Again, this may have practical advantages.

Our encoding and decoding scheme is summarised in Figure 4.2 and treated in detail in the next sections.

4.2.1 The random ensemble of lattice codes

We still work with a random Construction A ensemble; it is simply given by a random parity-check matrix, whose entries are independent random variables uniformly distributed over $\{0, 1, \dots, p-1\} = \mathbb{F}_p$. In particular, let H be this matrix, of dimension $n(1-R) \times n$ for some $0 < R < 1$ and let H_f be its lower submatrix formed by the last $n(1-R_f)$ rows of H for some $R < R_f < 1$:

$$H = \left(\frac{H'}{H_f} \right) \tag{4.4}$$

1. **Generation of the random lattice.** Choose with uniform distribution over $\mathbb{F}_p$ a parity-check matrix H of dimension $(\ell + r) \times n$ (with $\ell = n(R_f - R)$ and $r = n(1 - R_f)$, see (4.4)).
2. **Encoding of a message $\mathbf{m} \in \mathbb{F}_p^\ell$.** Find a vector $\mathbf{x} \in \mathbb{Z}^n$ of smallest norm such that $H\mathbf{x}^T \equiv (\mathbf{m} \mid \mathbf{0})^T \in \mathbb{F}_p^\ell \times \mathbb{F}_p^r$. The uncoded messages are supposed to be uniformly chosen.
3. **Decoding of the received vector $\mathbf{y}$.** MMSE lattice decoding of $\mathbf{y}$: $\hat{\mathbf{x}} = Q_{\Lambda_f}(\alpha\mathbf{y})$, where α is the Wiener coefficient (Definition 4.2).

Figure 4.2: Our encoding and decoding scheme.

The submatrix H_f defines a linear code C_f over $\mathbb{F}_p$ and the whole matrix H defines a subcode C of C_f. If Λ is the lattice obtained by Construction A from C and Λ_f is the one coming from C_f, they are nested:

$$\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : H\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\} \subseteq \{\mathbf{x} \in \mathbb{Z}^n : H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\} = \Lambda_f.$$

The Voronoi constellation that we consider is then given by $\Lambda_f \cap \mathcal{V}(\Lambda)$. It should be clear that this approach is dual to the one adopted by Ordentlich and Erez, who described the random codes by the means of generator matrices.

Notice that we will let the dimension n change and tend to infinity all along our proofs, so it would be more proper to call the lattices of the random ensemble $\Lambda^{(n)}$ and $\Lambda_f^{(n)}$; nevertheless, since this will not cause any ambiguity, we will keep the lightened notation Λ and Λ_f.

If we suppose that all the rows of H are linearly independent, then R and R_f are the real rates of the codes C and C_f respectively. We have

$$\mathrm{Vol}(\Lambda) = p^{n(1-R)} \text{ and } \mathrm{Vol}(\Lambda_f) = p^{n(1-R_f)},$$

from which we deduce that the cardinality M of the lattice constellation is

$$M = |\Lambda_f/\Lambda| = \frac{\mathrm{Vol}(\Lambda)}{\mathrm{Vol}(\Lambda_f)} = p^{n(R_f-R)}. \tag{4.5}$$

Notice that the probability that the rank of H is strictly smaller than $n(1 - R)$ can be shown to decrease to 0 very fast when n tends to infinity (see also (4.3)); hence we will work as if H was always full-rank. In other terms, an H whose rows are not independent has to be considered a "bad" choice and discarded; the contribution of these "bad" matrices to the average arguments that we carry out asymptotically vanishes and is completely negligible.

4.2.2 Encoding and decoding

The points of the constellation (or equivalently the cosets of Λ_f/Λ) are indexed by the $p^{n(R_f-R)}$ different syndromes of the form $(s_1, s_2, \ldots, s_{n(R_f-R)}, 0, \ldots, 0)^T$ associated with the matrix H, where all the $s_i \in \mathbb{F}_p$. More explicitly, let $\ell = n(R_f - R)$ and let $\mathbb{F}_p^\ell$ be (in 1-1 correspondence with) the set of messages; the bijection

$$\begin{aligned} \varphi\colon \Lambda_f \cap \mathcal{V}(\Lambda) &\to \mathbb{F}_p^\ell \\ \mathbf{x} &\mapsto H'\mathbf{x}^T \bmod p \end{aligned} \tag{4.6}$$

makes possible a constructive encoding (recall that H' is the upper submatrix of H, see (4.4)). Our transmission scheme works as follows:

1. The sender pairs up a message and a syndrome and transmits $\mathbf{x}$, the corresponding constellation point (via φ^{-1}), over the AWGN channel (no dithering is required).
2. The receiver gets the channel output $\mathbf{y} = \mathbf{x}+\mathbf{w}$ and multiplies it by the Wiener coefficient α.
3. Then he performs lattice decoding of $\alpha\mathbf{y}$ and gets $\hat{\mathbf{x}} = Q_{\Lambda_f}(\alpha\mathbf{y})$.
4. The decoded message will be the one associated with $\varphi(\hat{\mathbf{x}})$.

A final remark on the bijection φ: for every $\mathbf{s}' \in \mathbb{F}_p^\ell$, let $\mathbf{x} \in \Lambda_f$ be any solution of the linear system $H'\mathbf{x}^T \equiv \mathbf{s}' \bmod p$. Then

$$\varphi^{-1}(\mathbf{s}') = \mathbf{x} - Q_\Lambda(\mathbf{x})$$

and encoding can be done substantially thanks to a lattice decoder, too.

4.2.3 How to achieve capacity - Overview and discussion on our proof

We will try now to give a general description of our proof, by the means of a heuristic argument that does not take into account all the probabilistic and asymptotic aspects of the rigorous demonstration.

Geometric description

Our result is based on the following facts:

- The points of the constellation typically have the same norm and lie very close to the surface of a sphere of a given radius (to be specified later, see Lemma 4.3).
- The AWGN noise is typically orthogonal to the sent vector, in the sense that, if $\mathbf{x}$ is our transmitted constellation point and $\mathbf{w}$ is the noise, then $\mathbf{x}\mathbf{w}^T \approx 0$ (cf. Lemma 4.4).

- The effective noise due to MMSE scaling and the sent point are not decorrelated. Consequently, it is not possible to show that (MMSE) lattice decoding works with very high probability independently of the sent point. Nevertheless, Theorem 4.1 is based on the fact that the number of points for which this does not happen is not big enough to perturb the average error probability of the family.

- As in the proofs of Theorem 3.1 and Theorem 3.2, for a certain channel output (MMSE-scaled, in this case), we look for lattice points inside a sphere centred at it and with a typical radius to be specified later. There will be no error decoding if the only point in this *decoding sphere* is the transmitted one.

Consider that when we use the adverb "typically", we mean "with probability tending to 1 when n tends to infinity". The accurate proof will be treated in all detail in the sequel, but let us try to understand the geometric sense of the elements that we have just listed. So, suppose that the channel input is a point $\mathbf{x}$ whose norm is fixed to be $||\mathbf{x}|| = \sqrt{nP}$, for some $P > 0$, which will turn out to be the average (and asymptotically maximum) power of the constellation. Suppose also that $\mathbf{x}\mathbf{w}^T = 0$; if $\mathbf{y} = \mathbf{x} + \mathbf{w}$ is the channel output, then $||\mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{w}||^2$. Now, let us multiply $\mathbf{y}$ by a scalar value β such that the distance between $\mathbf{x}$ and $\beta\mathbf{y}$ is minimised. Basic Euclidean geometry (see Figure 4.3) tells us that if $||\mathbf{w}||^2 = n\sigma^2$, then $\beta = P/(\sigma^2 + P) = \alpha$, the Wiener coefficient (cf. Lemma 4.1). This lets us guess that MMSE lattice decoding helps in bringing the decoder input closer to the sent point.

The receiver passes $\alpha\mathbf{y}$ to the lattice decoder and there will be no decoding error if there is no other lattice point closer to $\alpha\mathbf{y}$ than $\mathbf{x}$. We will show that this typically happens if:

1. SNR > 1.

2. $P \approx p^{2(1-R)}/2\pi e$;

3. $||\alpha\mathbf{y} - \mathbf{x}||^2$ asymptotically does not exceed $np^{2(1-R_f)}/2\pi e$.

Notice that the previous bound concretely means that our constellation tolerates an "effective" noise after MMSE scaling whose variance per dimension is at most as strong as the noise corresponding to Potyrev capacity. We intuitively understand that this can be the good condition, admitting that no issue comes from the fact that the "effective" noise and the sent point $\mathbf{x}$ are not decorrelated (here, we have no dither to guarantee that).

The condition on the signal-to-noise ratio can be simply understood with the following argument: let us call $\mathbf{h} = \alpha\mathbf{y} - \mathbf{x}$ and suppose that it takes the maximum value according to the second condition above here, $||\mathbf{h}||^2 = np^{2(1-R_f)}/2\pi e = n\sigma^2_{\text{dec}}$. We use the index "dec" to indicate that the quantity corresponds to the (upper bound of the) "decodable" effective noise and to the *decoding sphere* defined in the proof of Theorem 4.1. If we want good decoding, we need $\alpha\mathbf{y}$ to be closer to $\mathbf{x}$ than to $\mathbf{0}$, because the latter always belongs to the lattice constellation; in other terms,

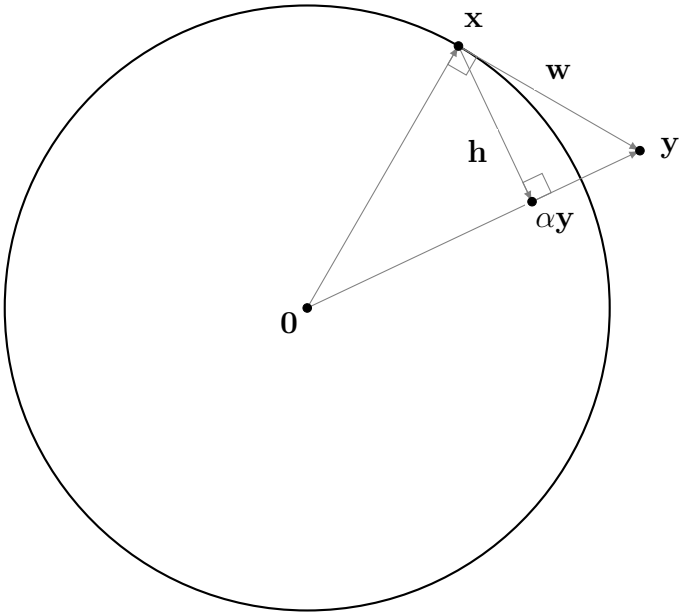

Figure 4.3: Geometric interpretation:

- $\mathbf{x}$: transmitted constellation point. $||\mathbf{x}||^2 = nP$.
- $\mathbf{w}$: AWG noise vector. $||\mathbf{w}||^2 = n\sigma^2$.
- $\mathbf{y}$: AWGN channel output. $\mathbf{y} = \mathbf{x} + \mathbf{w}$.
- α: Wiener coefficient. $\alpha = \frac{P}{P+\sigma^2}$.
- $\alpha\mathbf{y}$: lattice decoder input.
- $\mathbf{h}$: effective noise corresponding to MMSE scaling.

it is necessary that $||\alpha\mathbf{y}||^2 > ||\mathbf{h}||^2$. Again, a Euclidean geometry argument based on Figure 4.3 shows that (always supposing that $\mathbf{x}\mathbf{w}^T = 0$)

$$n\sigma^2_{\text{dec}} = ||\mathbf{h}||^2 = \frac{||\mathbf{x}||^2||\mathbf{w}||^2}{||\mathbf{y}||^2} = \frac{n^2P\sigma^2}{nP + n\sigma^2} = \frac{nP\sigma^2}{P + \sigma^2}, \tag{4.7}$$

while

$$||\alpha\mathbf{y}||^2 = \frac{P^2(nP + n\sigma^2)}{(P + \sigma^2)^2} = \frac{nP^2}{P + \sigma^2}.$$

Then, $||\alpha\mathbf{y}||^2 > ||\mathbf{h}||^2$ becomes

$$\frac{nP^2}{P + \sigma^2} > \frac{nP\sigma^2}{P + \sigma^2},$$

that is $P > \sigma^2$ or, equivalently, SNR > 1.

Taking $||\mathbf{h}||^2 = n\sigma^2_{\text{dec}}$ corresponds to a maximum rate for the constellation that equals capacity, as can be understood from the following calculation (that, again, is based on the approximations done till now and has only a demonstrative purpose): from (4.7) we can derive that

$$\sigma^2 = \frac{P\sigma^2_{\text{dec}}}{P - \sigma^2_{\text{dec}}}.$$

This implies that

$$\text{SNR} = \frac{P}{\sigma^2} = \frac{P}{\sigma^2_{\text{dec}}} - 1.$$

Observe that the previous formula shows how decoding $\alpha\mathbf{y}$ enhances the strength of the constellation, as if we had an "effective" SNR_{eff} equal to $P/\sigma^2_{\text{dec}} = \text{SNR} + 1$. This heuristically explains how we manage to gain the "plus 1" in the formula $1/2\log_2(\text{SNR})$. The same argument was pointed out in Erez and Zamir's work [EZ04]. To conclude, recall that we make the hypothesis that $P = p^{2(1-R)/2\pi e}$; this, together with (4.5) leads to

$$\begin{aligned} \frac{1}{2}\log_2(1+\text{SNR}) &= \frac{1}{2}\log_2\left(\frac{P}{\sigma^2_{\text{dec}}}\right) && (4.8)\\ &\approx \frac{1}{2}\log_2(p^{2(R_f-R)}) \\ &= \frac{1}{n}\log_2(p^{n(R_f-R)}), && (4.9) \end{aligned}$$

which is the rate of our constellation. A stronger rate would give a non-reliably decodable constellation, since the "effective" noise would make $||\mathbf{h}||^2$ exceed $n\sigma^2_{\text{dec}}$ and no reliable decoding could be guaranteed.

Originality of our proof and lattice decoding of $\alpha\mathbf{y}$

What we have explained till now gives an intuitive description of the typical geometry that characterises the AWG noise and the random Voronoi constellations of Construction A nested lattices. Nevertheless, it does not directly drop a hint on the original idea behind our proof that allows to avoid dithering. It is worth the effort of spending some words about that now, before moving on to the detailed proof.

The main argument is the following: if $\alpha\mathbf{y}$ is the real point that the receiver passes to the lattice decoder, we would like to ensure that the only lattice point inside the *decoding sphere* $B_{\alpha\mathbf{y},n}(\sqrt{n}\sigma_{\text{dec}})$ is the sent point. This is equivalent to saying that lattice decoding does not fail. Of course, this does not happen for every instance of the AWG noise and may not happen for every point of the constellation. Nevertheless, one hopes to be able to apply a similar technique as the one of the proofs of Theorem 3.1 and Theorem 3.2. In those settings, it was shown that the probability of a decoding failure is asymptotically negligible, independently on the sent point $\mathbf{x}$. This last feature makes the big difference with the case of MMSE lattice decoding. Indeed, the average argument that we apply will lead to the estimation

of (a more elaborated version of) the following sum:

$$\sum_{\mathbf{z}\in B_{\alpha\mathbf{y},n}(\sqrt{n}\sigma_{\mathrm{dec}})} \mathcal{P}\{\mathbf{z}\in\Lambda_f \mid \mathbf{x}\in\Lambda_f\}.$$

In Theorem 3.1 and Theorem 3.2, the two events $\{\mathbf{z}\in\Lambda_f\}$ and $\{\mathbf{x}\in\Lambda_f\}$ were independent, thanks to the action of the random noise. In our setting, instead, the multiplication by α adds some correlation between $\mathbf{x}$ and $\alpha\mathbf{y}$. One can interpret Erez and Zamir's dithering technique as a method of eliminating this correlation.

We do not use dither and consequently there will a priori be some points $\mathbf{x}$'s for which the previous sum takes a "big" value. Our analysis shows that the proportion of this kind of $\mathbf{x}$'s in the constellation is very small and the total error decoding probability still goes to 0 when n tends to infinity (see Lemma 4.5 and its application to (4.35) in the proof of Theorem 4.1).

4.2.4 The detailed proof

From now on, we will go into all the technical aspects of our proof that there exists a random capacity-achieving Construction A lattice family. This result will be formally stated and proved in Theorem 4.1. For the sake of clearness, we have taken out of the proof of the theorem a certain number of lemmas, that we present below here. The experienced reader may prefer to directly go to the proof of Theorem 4.1 and then to go back to the demonstration of all the involved lemmas.

The typical norm of a constellation point

We treat here the problem of the typical norm of a constellation point. Let $\rho_{\mathrm{eff}}^{(n)}$ be the effective radius of the n-dimensional shaping lattice Λ (see also Definition 2.21). It is the radius of the ball which has the same volume as $\mathcal{V}(\Lambda)$, the Voronoi region of the shaping lattice: $\mathrm{Vol}(\mathcal{V}(\Lambda)) = \mathrm{Vol}(B_{\mathbf{0},n}(\rho_{\mathrm{eff}}))$. Hence

$$\rho_{\mathrm{eff}}^{(n)} = p^{(1-R)}\,\mathrm{Vol}(B_{\mathbf{0},n}(1)^{-1/n}) \sim \frac{\sqrt{n}p^{(1-R)}}{\sqrt{2\pi e}},$$

by Lemma 2.6. We denote the asymptotic value

$$\rho_{\mathrm{eff}} = \frac{\sqrt{n}p^{(1-R)}}{\sqrt{2\pi e}}. \tag{4.10}$$

We claim that for n large enough almost all the points of the constellation lie very close to the surface of the ball $B_{\mathbf{0},n}(\rho_{\mathrm{eff}})$. For this reason, we call it the *shaping sphere*. Before formally proving this, we need the following lemma:

Lemma 4.2. *Let $\mathcal{B} = B_{\mathbf{c},n}(\rho)$ be the n-dimensional ball of radius ρ and centre $\mathbf{c}$. Let $\mathbf{x}$ be any point of $\mathcal{B}\cap\mathbb{Z}^n$, let p be a prime number and let $\mu\in\mathbb{F}_p$. Then,*

$$|\{\mathbf{z}\in\mathcal{B}\cap\mathbb{Z}^n : \mathbf{z}\equiv\mu\mathbf{x} \bmod p\}| \le 1 + \frac{4\rho^2}{p^2}\left(\frac{8n\rho^2}{p^2}\right)^{4\rho^2/p^2}.$$

Proof. Let us start with the case $\mu = 1$, that outlines the strategy for a more general μ. If $\mathbf{z} \equiv \mathbf{x} \bmod p$, then $\mathbf{x} - \mathbf{z} \in p\mathbb{Z}^n$. Hence, $x_i - z_i = k_i p$, for some $k_i \in \mathbb{Z}$. Let us call $N = \sum_{i=1}^{n} |k_i|$; we have

$$||\mathbf{x} - \mathbf{z}||^2 = \sum_{i=1}^{n} (x_i - z_i)^2 = \sum_{i=1}^{n} k_i^2 p^2 \geq p^2 \sum_{i=1}^{n} |k_i| = p^2 N.$$

This, together with the fact that both $\mathbf{x}$ and $\mathbf{z}$ lie in $\mathcal{B}$, gives

$$p^2 N \leq ||\mathbf{x} - \mathbf{z}||^2 \leq 4\rho^2$$

and

$$N \leq \frac{4\rho^2}{p^2}.$$

Then, the number of $\mathbf{z}$'s equivalent to $\mathbf{x}$ in $\mathcal{B}$ is bounded by the number L of different vectors $(k_1, k_2, \ldots, k_n) \in \mathbb{Z}^n$ such that $\sum_{i=1}^{n} |k_i| \leq \lfloor 4\rho^2/p^2 \rfloor$. One of this vectors is simply $\mathbf{0} \in \mathbb{Z}^n$. Hence, L itself is bounded by 1 plus the number of possible ways of:

1. fixing m coordinates among n (with $1 \leq m \leq \lfloor 4\rho^2/p^2 \rfloor$; $m = 0$ corresponds to $k_i = 0$ for every i and corresponds to the "1 plus");
2. for every one of the m fixed coordinates, deciding if k_i will be positive or negative (and, for now, fix $k_i = 0$);
3. choosing for $\lfloor 4\rho^2/p^2 \rfloor$ times to increment one of the m coordinates k_i's of ± 1 (according to the sign fixed at step 2).

As a consequence,

$$\begin{aligned}
|\{\mathbf{z} \in \mathcal{B} \cap \mathbb{Z}^n : \mathbf{z} \equiv \mathbf{x} \bmod p\}| &\leq L \\
&\leq 1 + \sum_{m=1}^{\lfloor 4\rho^2/p^2 \rfloor} \binom{n}{m} 2^m m^{\lfloor 4\rho^2/p^2 \rfloor} \\
&\leq 1 + \sum_{m=1}^{\lfloor 4\rho^2/p^2 \rfloor} n^m 2^m m^{\lfloor 4\rho^2/p^2 \rfloor} \\
&\leq 1 + \frac{4\rho^2}{p^2} n^{4\rho^2/p^2} 2^{4\rho^2/p^2} \left(\frac{4\rho^2}{p^2}\right)^{4\rho^2/p^2} \\
&= 1 + \frac{4\rho^2}{p^2} \left(\frac{8n\rho^2}{p^2}\right)^{4\rho^2/p^2},
\end{aligned}$$

The lemma is proved for $\mu = 1$. Now, let us consider the case in which μ takes another value and let $\mathbf{z}'$ be any point inside the sphere such that $\mathbf{z}' \equiv \mu\mathbf{x} \bmod p$. Then

$$|\{\mathbf{z} \in \mathcal{B} \cap \mathbb{Z}^n : \mathbf{z} \equiv \mu\mathbf{x} \bmod p\}| = |\{\mathbf{z} \in \mathcal{B} \cap \mathbb{Z}^n : \mathbf{z} \equiv \mathbf{z}' \bmod p\}|$$

and the proof works exactly in the same way as before, with $\mathbf{z}'$ instead of $\mathbf{x}$. □

We are ready to state and demonstrate the lemma about the typical norm of a constellation point. Of course, the constellation we consider is the one presented in Section 4.2.1:

Lemma 4.3. *Let* $\mathbf{s} = (s_1, s_2, \ldots, s_{n(R_f-R)}, 0, \ldots, 0) \in \mathbb{F}_p^{n(1-R)} \setminus \{\mathbf{0}\}$ *be any non-zero syndrome, keeping the previous notation. Suppose that* $p = n^\lambda$ *for some* $\lambda > 0$ *and let* ω *be a positive constant such that*

$$\omega < \min\{\lambda(1-R), 2\lambda R, 1\}. \tag{4.11}$$

If $\mathbf{x}$ *is the random (over the choice of the matrix* H*) constellation point associated with the syndrome* $\mathbf{s}$ *via* φ^{-1} *(cf.* (4.6)*), then*

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho_{\text{eff}}\left(1-\frac{1}{n^\omega}\right) \le ||\mathbf{x}|| \le \rho_{\text{eff}}\left(1+\frac{1}{n^\omega}\right)\right\} = 1. \tag{4.12}$$

Proof. Let X_ρ be the random variable that counts the number of points with syndrome $\mathbf{s}$ in the n-dimensional ball $B_{\mathbf{0},n}(\rho)$ centred at $\mathbf{0}$ with radius ρ. For any $\rho \ge 0$ and for any $\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$, we define the random variable

$$X_{\mathbf{x}} = \begin{cases} 1, & \text{if } H\mathbf{x}^T \equiv \mathbf{s}^T \bmod p \\ 0, & \text{otherwise} \end{cases}$$

that of course depends on the random choice of H. In particular,

$$\mathcal{P}\{X_{\mathbf{x}} = 1\} = \begin{cases} \left(\frac{1}{p}\right)^{n(1-R)}, & \text{if } \mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \setminus p\mathbb{Z}^n \\ 0, & \text{if } \mathbf{x} \in p\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \end{cases}$$

and clearly

$$X_\rho = \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} X_{\mathbf{x}}.$$

We will split the proof into two parts. First of all, we argue that

$$\lim_{n\to\infty} \mathcal{P}\{X_{\rho_{\text{eff}}\left(1-\frac{1}{n^\omega}\right)} > 0\} = 0. \tag{4.13}$$

Later, we show that

$$\lim_{n\to\infty} \mathcal{P}\{X_{\rho_{\text{eff}}\left(1+\frac{1}{n^\omega}\right)} = 0\} = 0. \tag{4.14}$$

These two results together imply (4.12).

Proof of (4.13)**.** When $\rho = \rho_{\text{eff}}\,(1 - 1/n^{\omega})$,

$$\begin{aligned}
\mathbb{E}[X_\rho] &= \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}} = 1\} \\
&\le |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| \left(\frac{1}{p}\right)^{n(1-R)} \qquad (4.15)\\
&\le \mathrm{Vol}\left(B_{\mathbf{0},n}\left(\rho + \frac{\sqrt{n}}{2}\right)\right)\left(\frac{1}{p}\right)^{n(1-R)} \\
&= \mathrm{Vol}\left(B_{\mathbf{0},n}(1)\right)\rho_{\text{eff}}^n \left(1 - \frac{1}{n^{\omega}}\right)^n \left(1 + \frac{\sqrt{n}}{2\rho}\right)^n \left(\frac{1}{p}\right)^{n(1-R)} \\
&\sim \frac{1}{\sqrt{\pi n}} \exp(-n^{1-\omega}) \exp\left(\frac{\sqrt{2\pi e}}{2} n^{-\lambda(1-R)} n \left(1 - \frac{1}{n^{\omega}}\right)^{-1}\right) \\
&= \frac{1}{\sqrt{\pi n}} \exp\left(\sqrt{\frac{\pi e}{2}}\left(1 - \frac{1}{n^{\omega}}\right)^{-1} n^{-\lambda(1-R)+1} - n^{1-\omega}\right), \qquad (4.16)
\end{aligned}$$

where we have used Lemma 2.6 for the asymptotic value of $\mathrm{Vol}\left(B_{\mathbf{0},n}(1)\right)$. The whole quantity tends to 0, since $1 - \omega > -\lambda(1-R) + 1$ by (4.11) and the argument of the exponential function goes to $-\infty$; considering the fact that $\mathcal{P}\{X_\rho > 0\} \le \mathbb{E}[X_\rho]$, we also have

$$\lim_{n\to\infty} \mathcal{P}\{X_{\rho_{\text{eff}}\left(1-\frac{1}{n^{\omega}}\right)} > 0\} = 0.$$

Proof of (4.14)**.** Now, let $\rho = \rho_{\text{eff}}\,(1 + 1/n^{\omega})$. Taking into account the fact that $|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n| \sim |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|$, we have

$$\begin{aligned}
\mathbb{E}[X_\rho] &= \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}} = 1\} \qquad (4.17)\\
&= |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n| \left(\frac{1}{p}\right)^{n(1-R)} \qquad (4.18)\\
&\sim |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| \left(\frac{1}{p}\right)^{n(1-R)} \\
&\ge \mathrm{Vol}\left(B_{\mathbf{0},n}\left(\rho - \frac{\sqrt{n}}{2}\right)\right)\left(\frac{1}{p}\right)^{n(1-R)} \\
&= \mathrm{Vol}\left(B_{\mathbf{0},n}(1)\right)\rho_{\text{eff}}^n \left(1 + \frac{1}{n^{\omega}}\right)^n \left(1 - \frac{\sqrt{n}}{2\rho}\right)^n \left(\frac{1}{p}\right)^{n(1-R)} \\
&\sim \frac{1}{\sqrt{\pi n}} \exp(n^{1-\omega}) \exp\left(-\frac{\sqrt{2\pi e}}{2} n^{-\lambda(1-R)} n \left(1 + \frac{1}{n^{\omega}}\right)^{-1}\right) \\
&= \frac{1}{\sqrt{\pi n}} \exp\left(n^{1-\omega} - \sqrt{\frac{\pi e}{2}}\left(1 + \frac{1}{n^{\omega}}\right)^{-1} \cdot n^{-\lambda(1-R)+1}\right), \qquad (4.19)
\end{aligned}$$

which tends to infinity, again thanks to (4.11). Hence

$$\lim_{n\to\infty} \mathbb{E}[X_\rho] = +\infty.$$

Suppose now for a moment that $\mathrm{Var}(X_\rho) \leq f(n)\mathbb{E}[X_\rho]$ for some $f(n) = o(\mathbb{E}[X_\rho])$; we would have

$$\begin{aligned} \mathcal{P}\{X_\rho = 0\} &\leq \mathcal{P}\{|X_\rho - \mathbb{E}[X_\rho]| \geq \mathbb{E}[X_\rho]\} \\ &\leq \frac{\mathrm{Var}(X_\rho)}{\mathbb{E}[X_\rho]^2} \\ &\leq \frac{f(n)}{\mathbb{E}[X_\rho]} \longrightarrow 0, \end{aligned} \tag{4.20}$$

where we have applied Chebyshev's inequality (Lemma 2.1) to obtain (4.20). This would be enough to prove (4.14) and conclude. For this reason, let us show that $\mathrm{Var}(X_\rho) \leq f(n)\mathbb{E}[X_\rho]$; to do this, we investigate the quantity

$$\mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) = \mathbb{E}[X_{\mathbf{x}}X_{\mathbf{z}}] - \mathbb{E}[X_{\mathbf{x}}]\mathbb{E}[X_{\mathbf{z}}],$$

for $\mathbf{x}, \mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$. Observe that, by the definition of the two random variables, if $\mathbf{h}_i$ is the i-th row of H,

$$\begin{aligned} \mathbb{E}[X_{\mathbf{x}}X_{\mathbf{z}}] &= \mathcal{P}\{X_{\mathbf{x}}X_{\mathbf{z}} = 1\} \\ &= \mathcal{P}\{X_{\mathbf{x}} = 1, X_{\mathbf{z}} = 1\} \\ &= \prod_{i=1}^{n(1-R)} \mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p, \mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p\}. \end{aligned}$$

There are three possibilities:

1. If $\mathbf{x} \not\equiv a\mathbf{z} \bmod p$ for all $a \in \mathbb{F}_p$, let i be an index such that $s_i \neq 0$ (there always exists, since $\mathbf{s} \neq \mathbf{0}$); then

$$\begin{aligned} &\mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p, \mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p\} \\ &= \mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p\}\mathcal{P}\{\mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p\} = \\ &= \begin{cases} 0, & \text{if } \mathbf{x} \text{ or } \mathbf{z} \text{ belong to } p\mathbb{Z}^n \\ \left(\frac{1}{p}\right)^2, & \text{otherwise} \end{cases}. \end{aligned}$$

   and $\mathbb{E}[X_{\mathbf{x}}X_{\mathbf{z}}] = (1/p)^{2n(1-R)} = \mathbb{E}[X_{\mathbf{x}}]\mathbb{E}[X_{\mathbf{z}}]$, that is $\mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) = 0$.

2. If $\mathbf{x} \equiv a\mathbf{z} \bmod p$ for some $a \in \mathbb{F}_p \setminus \{1\}$, again let i be an index such that $s_i \neq 0$. Hence either $a\mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p$ or $\mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p$, with no chance that the two events happen together. Then

$$\mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p, \mathbf{h}_i\mathbf{z}^T \equiv s_i \bmod p\} = 0,$$

   $\mathbb{E}[X_{\mathbf{x}}X_{\mathbf{z}}] = 0$ and $\mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) \leq 0$.

3. Finally, if $\mathbf{x} \equiv \mathbf{z} \bmod p$, then $X_{\mathbf{x}} X_{\mathbf{z}} = X_{\mathbf{x}}^2$ and $\mathbb{E}[X_{\mathbf{x}} X_{\mathbf{z}}] = \mathbb{E}[X_{\mathbf{x}}]$. That is, $\mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) \leq \mathbb{E}[X_{\mathbf{x}}]$.

Putting all of this together, we have

$$\begin{aligned}
\mathrm{Var}(X_\rho) &= \mathrm{Var}\left(\sum_{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)} X_{\mathbf{x}} \right) \\
&= \sum_{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)} \mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) \\
&= \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \\ \mathbf{x} \not\equiv a\mathbf{z}}} \mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) + \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \\ \mathbf{x} \equiv a\mathbf{z},\ a \neq 1}} \mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) \\
&\quad + \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \\ \mathbf{x} \equiv \mathbf{z}}} \mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}}) \\
&\leq \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)} \sum_{\substack{\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \\ \mathbf{x} \equiv \mathbf{z}}} \mathbb{E}[X_{\mathbf{x}}] \\
&\leq \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)} \left(1 + \frac{4\rho^2}{p^2} \left(\frac{8n\rho^2}{p^2} \right)^{4\rho^2/p^2} \right) \mathbb{E}[X_{\mathbf{x}}] \qquad (4.21) \\
&= \left(1 + \frac{4\rho^2}{p^2} \left(\frac{8n\rho^2}{p^2} \right)^{4\rho^2/p^2} \right) \mathbb{E}[X_\rho],
\end{aligned}$$

where (4.21) is a consequence of Lemma 4.2. The last thing we need to conclude is that

$$\lim_{n\to\infty} \frac{f(n)}{\mathbb{E}[X_\rho]} = \lim_{n\to\infty} \frac{1 + 4\rho^2/p^2 \, (8n\rho^2/p^2)^{4\rho^2/p^2}}{\mathbb{E}[X_\rho]} = 0.$$

Taking into account that $\rho = \sqrt{n} p^{(1-R)}(1 + 1/n^\omega)/2\pi e$ and $p = n^\lambda$, one can compute that the dominating term (up to some multiplicative constants in the exponent) of the numerator is $n^{n^{1-2\lambda R}} = \exp(n^{1-2\lambda R} \ln n)$. On the other hand, (4.19) and (4.11) tell that the dominating term in the asymptotic lower bound for the denominator is $\exp(n^{1-\omega})$. Hence, the limit is 0 if

$$1 - 2\lambda R < 1 - \omega,$$

which is true, again by (4.11). □

A property of the Gaussian noise

The following lemma formally explains in what (probabilistic, asymptotic) sense the typical AWG noise vector is orthogonal to constellation point vectors. Explicitly, we bound their scalar product by a quantity that in the proof of Theorem 4.1 turns out to be negligible with respect to their squared norms. Hence $||\mathbf{x}+\mathbf{w}||^2 \approx ||\mathbf{x}||^2 + ||\mathbf{w}||^2$.

Lemma 4.4 (Orthogonal noise)**.** *Let* $\mathbf{x} \in \mathbb{R}^n$ *and let* $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ *be a random AWG noise vector:* $w_i \sim \mathcal{N}(0, \sigma^2)$*. Then, for every function* $f(n)$ *such that* $\lim_{n\to\infty} f(n) = +\infty$*, we have*

$$\lim_{n\to\infty} \mathcal{P}\{|\mathbf{x}\mathbf{w}^T| \leq f(n)\sigma||\mathbf{x}||\} = 1.$$

Proof. Of course, if $\mathbf{x} = \mathbf{0}$, the statement is trivially true. So, suppose from now on that $\mathbf{x} \neq \mathbf{0}$. The scalar product $\mathbf{x}\mathbf{w}^T = \sum_{i=1}^n x_i w_i$ is a sum of i.i.d. Gaussian random variables, weighted by the x_i's, then it is well-known that it follows a Gaussian distribution, too. More precisely, $\mathbb{E}[\mathbf{x}\mathbf{w}^T] = 0$ and

$$\operatorname{Var}\left(\sum_{i=1}^n x_i w_i\right) = \sum_{i=1}^n x_i^2 \operatorname{Var}(w_i) = \sigma^2||\mathbf{x}||^2.$$

Consider $Q(\cdot)$, the tail probability of the standard normal distribution:

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u.$$

For positive y, the Chernoff bound states that

$$Q(y) \leq \frac{1}{2} e^{-\frac{y^2}{2}}.$$

We apply this bound to our probability and we have

$$\begin{aligned}\mathcal{P}\{|\mathbf{x}\mathbf{w}^T| > f(n)\sigma||\mathbf{x}||\} &= 2Q\left(\frac{f(n)\sigma||\mathbf{x}||}{\sigma||\mathbf{x}||}\right) \\ &\leq \exp\left(-\frac{f(n)^2}{2}\right),\end{aligned}$$

which tends to 0 because of the choice of $f(n)$ by hypothesis.

Hence,

$$\lim_{n\to\infty} \mathcal{P}\{|\mathbf{x}\mathbf{w}^T| \leq f(n)\sigma||\mathbf{x}||\} = 1.$$

□

Multiple points modulo p in the decoding sphere

The following lemma may appear independent from the context and the discussion that we have done till now. Nevertheless, it will be useful for the analysis of the decoding error probability in the proof of Theorem 4.1.

If $z \in \mathbb{Z}$ we denote by $\overline{z} \in \mathbb{Z}$ the element of the class of z modulo p with the smallest absolute value; that is, $z \equiv \overline{z} \bmod p$ and $\overline{z}$ is the class representant lying in $\{-(p-1)/2, -(p-3)/2, \ldots, (p-1)/2\}$.

Lemma 4.5. *Consider $\mathcal{B} = B_{\mathbf{0},n}(\rho_{\text{eff}}(1+1/n^{\omega}))$ and let $\rho = p^{1-R_f}\sqrt{n}(1+\varepsilon)/\sqrt{2\pi e}$ (where $\rho_{\text{eff}} = \sqrt{n}p^{(1-R)}/\sqrt{2\pi e}$, $p = n^{\lambda}$ for some constant λ, ω is chosen as in Lemma 4.3, and R and R_f are the same that we have considered till now in this section). Moreover, suppose that*

$$R > 1/2.$$

Let $\mu \in \mathbb{F}_p = \{-(p-1)/2, -(p-3)/2, \ldots, (p-1)/2\} \setminus \{0,1,2\}$. We define

$$N(\mu) = |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : \exists \mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{x},n}(2\rho) \text{ for which } \mathbf{z} \equiv \mu\mathbf{x} \bmod p\}|.$$

Then, if n is big enough,

$$N = \sum_{\mu \in \mathbb{F}_p \setminus \{0,1,2\}} N(\mu) = o\left(\frac{p^{n(1-R)}}{h(n)}\right),$$

for every function $h(n)$ subexponential in n.

Proof. First of all, notice that $\mathbf{z} \equiv \mu\mathbf{x} \bmod p$ means $\mathbf{z} = \mu\mathbf{x} + p\mathbf{k}$, for some $\mathbf{k} \in \mathbb{Z}^n$. Hence, if we call $\nu = 1 - \mu$,

$$\begin{aligned} ||\mathbf{x} - \mathbf{z}||^2 &= ||\mathbf{x} - \mu\mathbf{x} - p\mathbf{k}||^2 \\ &\geq ||\overline{(1-\mu)\mathbf{x} - p\mathbf{k}}||^2 \\ &= ||\overline{(1-\mu)\mathbf{x}}||^2 \\ &= ||\overline{\nu\mathbf{x}}||^2. \end{aligned}$$

If $||\overline{\nu\mathbf{x}}||^2 > 4\rho^2$, then $||\mathbf{x} - \mathbf{z}||^2 > 4\rho^2$, too. In other words, $\mathbf{z}$ lies outside $B_{\mathbf{x},n}(2\rho)$ and $\mathbf{x}$ does not have to be counted among the ones contributing to $N(\mu)$. That is,

$$N(\mu) \leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : ||\overline{\nu\mathbf{x}}||^2 \leq 4\rho^2\}| = N'(\mu).$$

Now, let $C \geq 3$ be a constant to be fixed later. Given $\mathbf{x}$, let

$$J = \left\{ i \in \{1, 2, \ldots, n\} : |x_i| < \frac{\sqrt{p}}{C^2} \right\};$$

Let us prove that if n is large enough, then $|J| \geq \gamma n$ for every constant $0 < \gamma < 1$. Indeed, suppose by contradiction that $|J| < \gamma n$, then we would have at least $(1-\gamma)n$ coordinates x_i of $\mathbf{x}$ such that $|x_i| \geq \sqrt{p}/C^2$. We employ the hypothesis $R > 1/2$ to get:

$$\begin{aligned} ||\mathbf{x}||^2 &\geq (1-\gamma)n\frac{p}{C^4} \\ &> n\frac{p^{2(1-R)}}{2\pi e}\left(1 + \frac{1}{n^{\omega}}\right)^2 \\ &= \rho_{\text{eff}}^2\left(1 + \frac{1}{n^{\omega}}\right)^2 \\ &\geq ||\mathbf{x}||^2, \end{aligned}$$

which is a nonsense (notice that the second - strict - inequality is true for n large enough).

Before going on, for a given $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and for a subset of indices $I \subseteq \{1, 2, \dots, n\}$ we define $\mathbf{x}(I)$ to be the vector $(x(I)_1, x(I)_2, \dots, x(I)_n)$ such that

$$\mathbf{x}(I)_i = \begin{cases} 0, & \text{if } i \in I \\ x_i, & \text{otherwise} \end{cases}.$$

First estimate: $|\nu| \leq \sqrt{p}$. When ν is "small", denoting $J^c = \{1, 2, \dots, n\} \setminus J$, we have

$$||\overline{\nu\mathbf{x}}||^2 \geq ||\overline{\nu\mathbf{x}}(J^c)||^2 = ||\nu\mathbf{x}(J^c)||^2 \geq 4||\mathbf{x}(J^c)||^2;$$

the equality holds by definition of J and because $|\nu| \leq \sqrt{p}$, while the second inequality comes from the hypothesis on the range of μ, that implies $|\nu| \geq 2$. Now, if $||\mathbf{x}(J^c)||^2 > \rho^2$, then the previous chain of inequalities gives $||\overline{\nu\mathbf{x}}||^2 > 4\rho^2$. Hence

$$N'(\mu) \leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : ||\mathbf{x}(J^c)||^2 \leq \rho^2\}| \leq |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|,$$

noticing that $B_{\mathbf{0},n}(\rho) \subseteq \mathcal{B}$ because $\rho_{\text{eff}}(1 + 1/n^\omega) > \rho$ (provided that ε is small enough).

Second estimate: $|\nu| > \sqrt{p}$. Let η be a constant to be fixed later such that $0 < \eta < 1$. We say that $\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B}$ is *heavy* if for all $I \subseteq \{1, 2, \dots, n\}$ such that $|I| \leq (1-\eta)n$, we have $||\mathbf{x}(I)||^2 > 4\rho^2/C^2$. This definition qualitatively means that a heavy $\mathbf{x}$ is such that every "quite small" subset of coordinates still gives a "big enough" contribution to the total norm of $\mathbf{x}$ itself.

Now, let

$$I = \{i \in \{1, 2, \dots, n\} : |\overline{\nu x_i}| < C|x_i|\}.$$

Suppose that $\mathbf{x}$ is heavy, then, if $|I| \leq (1-\eta)n$,

$$||\overline{\nu\mathbf{x}}||^2 \geq ||\overline{\nu\mathbf{x}}(I)||^2 \geq C^2||\mathbf{x}(I)||^2 > 4\rho^2,$$

where the second inequality is a direct consequence of the definition of I. This means that in this case

$$N'(\mu) \leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : \mathbf{x} \text{ is not heavy}\}| + |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : \mathbf{x} \text{ is heavy, } |I| > (1-\eta)n\}|.$$

Let us call $N_1(\mu)$ the first addend and $N_2(\mu)$ the second one and estimate them.

Estimation of $N_1(\mu)$. If $\mathbf{x}$ is not heavy, there exists $I \subseteq \{1, 2, \dots, n\}$ such that $|I| \leq (1-\eta)n$ and $||\mathbf{x}(I)||^2 \leq 4\rho^2/C^2$. Notice that if this is true for $I = \emptyset$, then the same property holds a fortiori for a bigger I. Then

$$\begin{aligned} N_1(\mu) &\leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : \exists I \subseteq \{1, 2, \dots, n\} \text{ with } |I| \leq (1-\eta)n, ||\mathbf{x}(I)||^2 \leq 4\rho^2/C^2\}| \\ &\leq |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : ||\mathbf{x}(I)||^2 \leq 4\rho^2/C^2 \text{ when } I = \emptyset\}| \\ &= |\{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{B} : ||\mathbf{x}||^2 \leq 4\rho^2/C^2\}| \\ &= |\mathbb{Z}^n \cap B_{\mathbf{0},n}(2\rho/C)| \\ &\leq |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|, \end{aligned}$$

since $C \geq 3$ by definition.

Estimation of $N_2(\mu)$. Let $\mathbf{x}$ be heavy and suppose that $|I| > (1-\eta)n$. Now, let us fix $\gamma > \eta$ (say γ close to 1 and η close to 0) and let $I' = I \cap J$. Then, $|I'| \neq 0$ and

$$|I'| \geq |J| - |I^c| \geq (\gamma - \eta)n.$$

Let $S \subseteq \mathcal{B}$ be the set of integer points whose cardinality is $N_2(\mu)$ and that we have to estimate. We will create a relation $\phi : S \to \mathcal{B}$ (a "function" with more than one image per point), as follows: if $\mathbf{x} \in S$, fix $|I'|/2$ coordinates of I' and add to each of them 1 or -1, in such a way that the new point is still in $\mathcal{B}$. The set of images of $\mathbf{x}$ is made of all the $\binom{|I'|}{|I'|/2}$ new points that we obtain with the $\binom{|I'|}{|I'|/2}$ different choices of coordinates to modify. We denote it by $\phi(\mathbf{x}) \subseteq \mathcal{B}$. We have implicitly supposed that $|I'|/2$ is integer, but nothing would substantially change if $|I'|$ were odd. Observe that the number of images of each single $\mathbf{x}$ is bounded from below by

$$\binom{|I'|}{|I'|/2} \geq \frac{1}{\sqrt{2|I'|}} 2^{|I'|} \geq \frac{1}{\sqrt{2n}} 2^{(\gamma-\eta)n},$$

independently from I'. Of course, we have used Lemma 2.5 to approximate the binomial coefficient.

Now, let

$$S' = \{\mathbf{x}' \in \mathcal{B} : \mathbf{x}' \in \phi(\mathbf{x}), \exists \mathbf{x} \in \mathcal{B}\} \subseteq \mathcal{B}.$$

It is possible that a certain $\mathbf{x}' \in S'$ has more than one counterimage in S. We would like to estimate how many they can be. In order to count them, pay attention to the following facts: given an $\mathbf{x}$ in the set that we are considering, for all $i \in I' = I \cap J$ we have that

- $|x_i| < \sqrt{p}/C^2$ (by definition of J),
- $|\overline{\nu x_i}| < C|x_i|$ (by definition of I).

The two conditions together say that $|\overline{\nu x_i}| < \sqrt{p}/C$, while $|\nu| > \sqrt{p}$ by hypothesis. Then

$$|\overline{\nu(x_i \pm 1)}| = |\overline{\nu x_i \pm \nu}| > \sqrt{p}\left(1 - \frac{2}{C}\right) \geq \frac{\sqrt{p}}{C} \geq C|x_i \pm 1|.$$

Now, consider $\mathbf{x}' \in \phi(\mathbf{x})$ for some $\mathbf{x} \in S$; what we have just shown implies that all the coordinates x'_i of $\mathbf{x}'$ that are equal to a coordinate of $\mathbf{x}$ plus or minus 1 (i.e., all the "modified" coordinates of $\mathbf{x}$), are such that $|\overline{\nu x'_i}| \geq C|x'_i|$. As a consequence and by definition of I, every $\mathbf{x}' \in S'$ has between $|I'|/2$ and $|I'|/2 + \lfloor \eta n \rfloor$ coordinates such that $|\overline{\nu x'_i}| \geq C|x'_i|$. On the other hand, every $\mathbf{x} \in S$ has between 0 and $\lfloor \eta n \rfloor$ of them. This means that an upper bound for the number K of counterimages of $\mathbf{x}' \in S'$ is given by the number of possible modifications of plus or minus 1 (only towards the surface of $\mathcal{B}$, since ϕ always "pushes" a point towards the inner region) of $|I'|/2$ coordinates chosen among the at most $|I'|/2 + \eta n$ such that $|\overline{\nu x'_i}| \geq C|x'_i|$;

in formulae,

$$\begin{aligned} K &\leq \sum_{k=|I'|/2}^{|I'|/2+\lfloor \eta n \rfloor} \binom{|I'|/2+\lfloor \eta n \rfloor}{k} \\ &\leq \eta n 2^{|I'|/2+\lfloor \eta n \rfloor} \\ &\leq \eta n 2^{\frac{n}{2}+\eta n}. \end{aligned}$$

Summarising, we have created a relation ϕ that associates every point in S with at least $2^{(\gamma-\eta)n}/\sqrt{2n}$ points in S' and every point in S' with at most $\eta n 2^{(1/2+\eta)n}$ counterimages in S. In other terms,

$$N_2(\mu)\frac{1}{\sqrt{2n}}2^{(\gamma-\eta)n} \leq |S'|\eta n 2^{\frac{n}{2}+\eta n} \leq |\mathcal{B}\cap\mathbb{Z}^n|\eta n 2^{(\frac{1}{2}+\eta)n}$$

and

$$N_2(\mu) \leq \sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n}|\mathcal{B}\cap\mathbb{Z}^n|.$$

Putting together the estimation of $N_1(\mu)$ and $N_2(\mu)$, we get

$$N'(\mu) \leq |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| + \sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n}|\mathcal{B}\cap\mathbb{Z}^n|.$$

Conclusion. We have shown that for every value of ν (hence of μ), it is true that

$$N(\mu) \leq N'(\mu) \leq |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| + \sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n}|\mathcal{B}\cap\mathbb{Z}^n|.$$

Since the number of total different μ's is bounded by p, we can multiply by p the previous bound and get

$$N = \sum_{\mu\in\mathbb{F}_p\setminus\{0,1,2\}} N(\mu) \leq p|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| + p\sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n}|\mathcal{B}\cap\mathbb{Z}^n|.$$

Recall that the goal of this lemma is to prove that $N = o(p^{n(1-R)}/h(n))$ for every subexponential function $h(n)$. Let us split the previous sum and analyse the two addends separately. First of all, we claim that some standard computations, similar to the ones already carried out in this thesis (see the proofs of Theorem 3.1 and Lemma 4.3), tell that for some $f(n)$ subexponential in n,

$$\begin{aligned} \frac{p|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|h(n)}{p^{n(1-R)}} &\lesssim \frac{pf(n)h(n)(1+\varepsilon)^n p^{n(1-R_f)}}{p^{n(1-R)}} \\ &= \frac{pf(n)h(n)(1+\varepsilon)^n}{p^{n(R_f-R)}} \\ &= \frac{pf(n)h(n)(1+\varepsilon)^n}{\sqrt{1+\mathrm{SNR}}^n}; \end{aligned}$$

this last equality comes from (4.8) and 4.9. The ratio tends to 0 when n tends to infinity because ε can be taken as small as wanted.

With the same techniques as before, let $g(n)$ be the subexponential function in n coming from the estimation of $|\mathcal{B} \cap \mathbb{Z}^n|$. The second summand is treated as follows:

$$\frac{p\sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n} |\mathcal{B} \cap \mathbb{Z}^n| h(n)}{p^{n(1-R)}} \lesssim \frac{p\sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n} g(n)h(n)p^{n(1-R)}}{p^{n(1-R)}}$$
$$= p\sqrt{2}\eta n^{\frac{3}{2}} 2^{(2\eta+\frac{1}{2}-\gamma)n} g(n)h(n),$$

which tends to 0 because the dominating term is the exponential $2^{(2\eta+\frac{1}{2}-\gamma)n}$ and η and γ can be chosen in such a way that the exponent is negative. This ends the proof. □

The proof that capacity is achieved

We are now ready to state and prove the main result of this section:

Theorem 4.1. *The random ensemble of nested Construction A lattices introduced in Section 4.2.1 achieves the capacity of the AWGN channel under MMSE lattice decoding, when* $\mathrm{SNR} > 1$, $R > 1/2$ *and* $p = n^\lambda$ *for some constant* $\lambda > (1+R)^{-1}$.

Proof. The AWGN channel is defined by the $\mathrm{SNR} = P/\sigma^2 > 1$, for some AWG noise variance per dimension σ^2 and some power constraint P (cf. (2.6)). The capacity is then known to be

$$C = \frac{1}{2}\log_2(1 + \mathrm{SNR}).$$

We would like to show that for every fixed rate smaller than capacity, the random ensemble of Section 4.2.1 corresponding to that rate can be reliably decoded.

Dually, let us fix $1/2 < R < R_f < 1$ the rates of the $\mathbb{F}_p$-linear codes generating the nested lattice ensemble and $p = n^\lambda$ for some $\lambda > (1+R)^{-1}$; then the constellation $\mathcal{C} = \Lambda_f \cap \mathcal{V}(\Lambda)$ has rate

$$R_{\mathcal{C}} = \frac{\log_2 p^{n(R_f - R)}}{n}$$

and asymptotically the power constraint gives (see Lemma 4.3 and (4.10))

$$P = \frac{\rho_{\mathrm{eff}}^2}{n} = \frac{p^{2(1-R)}}{2\pi e}.$$

The inequality $R_{\mathcal{C}} < C$ is equivalent to

$$\sigma^2 < \frac{P}{|\mathcal{C}|^{2/n} - 1} = \frac{p^{2(1-R)}}{2\pi e(p^{2(R_f - R)} - 1)} = \sigma_{\max}^2. \tag{4.22}$$

We have called $\sigma_{\max}^2$ this upper bound, which is of course different from the $\sigma_{\max}^2$ of Chapter 3 (cf. (3.1)). Nevertheless, we keep the same notation because achieving capacity in this setting is still equivalent to prove that a random lattice in our ensemble can be reliably decoded (in big enough dimension) for every AWG noise variance value $\sigma^2 = \sigma_{\max}^2(1-\delta)^2$ with $0 < \delta < 1$, just like in the proofs of Theorem

3.1 and Theorem 3.2. The rest of the proof will be consecrated to demonstrate the previous statement.

The transmission scheme is of course the same that we have presented all along this section and schematized in Figure 4.2. Hence, let us consider any syndrome $\mathbf{s} = (s_1, s_2, \ldots, s_{n(R_f-R)}, 0, \ldots, 0)^T \in \mathbb{F}_p^n$ representing an uncoded message. We recall that the messages are supposed to be a priori equiprobable. Let $\mathbf{x}$ be the associated coded point for some random constellation in the family. If $\mathbf{w}$ is the channel noise (with coordinate-wise variance σ^2) and $\alpha = P/(P+\sigma^2)$ is the Wiener coefficient, we claim that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathcal{P}\{||\alpha\mathbf{y} - \mathbf{x}||^2 \leq \alpha n\sigma^2(1+\varepsilon)^2\} = 1. \tag{4.23}$$

If $\mathbf{s} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$ and $\mathbf{y} = \mathbf{w}$. The claim is a straightforward consequence of Lemma 2.2 (the fact that $\alpha < 1$ is also used). If instead $\mathbf{s} \neq \mathbf{0}$, let $\varepsilon' < \varepsilon$ be a positive constant, let $f(n)$ be a function such that $\lim_{n\to\infty} f(n) = +\infty$ (to be specified later) and let $\mathcal{E}_1$ be the event

$$\mathcal{E}_1 = \{||\mathbf{x}||^2 \leq nP(1+\varepsilon')^2\} \cap \{||\mathbf{w}||^2 \leq n\sigma^2(1+\varepsilon')^2\} \cap \{|\mathbf{x}\mathbf{w}^T| \leq f(n)\sigma||\mathbf{x}||\}.$$

Note that, provided that ε' is small enough, the event $\mathcal{E}_1$ is (asymptotically) contained in the event $\{||\alpha\mathbf{y} - \mathbf{x}||^2 \leq \alpha n\sigma^2(1+\varepsilon)^2\}$: indeed, $\mathcal{E}_1$ implies

$$\begin{aligned}
||\alpha\mathbf{y} - \mathbf{x}||^2 &= (\alpha-1)^2||\mathbf{x}||^2 + \alpha^2||\mathbf{w}||^2 + 2\alpha(\alpha-1)\mathbf{x}\mathbf{w}^T \\
&\leq \frac{\sigma^4}{(P+\sigma^2)^2}||\mathbf{x}||^2 + \frac{P^2}{(P+\sigma^2)^2}||\mathbf{w}||^2 + \frac{2\sigma^2 P}{(P+\sigma^2)^2}|\mathbf{x}\mathbf{w}^T| \\
&\leq \frac{\sigma^4 nP(1+\varepsilon')^2}{(P+\sigma^2)^2} + \frac{P^2 n\sigma^2(1+\varepsilon')^2}{(P+\sigma^2)^2} + \frac{2\sigma^2 P f(n)\sigma||\mathbf{x}||}{(P+\sigma^2)^2} \\
&\leq \frac{nP\sigma^2}{P+\sigma^2}\left((1+\varepsilon')^2 + \frac{2f(n)\sigma\sqrt{P}(1+\varepsilon')}{\sqrt{n}(P+\sigma^2)}\right)
\end{aligned} \tag{4.24}$$

and

$$\begin{aligned}
\lim_{n\to\infty} \frac{2f(n)\sigma\sqrt{P}(1+\varepsilon')}{\sqrt{n}(P+\sigma^2)} &\leq \lim_{n\to\infty} \frac{2f(n)\max\{\sigma^2, P\}(1+\varepsilon')}{\sqrt{n}(P+\sigma^2)} \\
&\leq \lim_{n\to\infty} \frac{2f(n)(1+\varepsilon')}{\sqrt{n}} = 0,
\end{aligned}$$

taking $f(n) = o(\sqrt{n})$. Thus, we can go back to (4.24) and obtain (for n big enough and ε' small enough with respect to ε) that

$$(4.24) \leq \frac{nP\sigma^2}{P+\sigma^2}(1+\varepsilon)^2 = \alpha n\sigma^2(1+\varepsilon)^2.$$

We are done, because

$$\mathcal{P}\{||\alpha\mathbf{y} - \mathbf{x}||^2 \leq \alpha n\sigma^2(1+\varepsilon)^2\} \geq \mathcal{P}\{\mathcal{E}_1\} \to 1, \tag{4.25}$$

by Lemma 4.3, Lemma 2.2 and Lemma 4.4.

We have just shown that with very high probability when n is big enough, the sent point $\mathbf{x}$ lies inside a sphere of radius $\rho_{\text{dec}} = \sqrt{\alpha n}\sigma(1+\varepsilon)$ centred at $\alpha\mathbf{y}$. We call this sphere the *decoding sphere* $\mathcal{B}$ and no decoding error occurs if the only point of $\Lambda_f \cap \mathcal{B}$ is $\mathbf{x}$. Let us call the "good decoding" event $\mathcal{E}_2 = \{\Lambda_f \cap \mathcal{B} = \{\mathbf{x}\}\}$ and $\mathcal{E}_2^c$ its complement.

To conclude the proof, we will show that for every syndrome $\mathbf{s}$, the probability that $\alpha\mathbf{y}$ is not well decoded tends to 0 for a randomly chosen lattice constellation in the ensemble. Let us call $P_e(\mathbf{s})$ this probability and let $X_\mathbf{s}$ be the random variable that represents the constellation point associated to $\mathbf{x}$; $X_\mathbf{s}$ takes a priori a different (n-dimensional) value for every different choice of a random constellation.

Let us start with $\mathbf{s} = \mathbf{0}$. In this case, $\mathcal{P}\{X_\mathbf{s} = \mathbf{0}\} = 1$. A very simple computation implies $\alpha\sigma^2 = p^{2(1-R_f)}(1-\delta)^2/2\pi e$ (one can also deduce it from (4.7)); Lemma 3.1 states that with probability tending to 1 no point of $p\mathbb{Z}^n$ different from $\mathbf{0}$ inside $\mathcal{B}$ can lead to bad decoding (notice that the multiplication by the Wiener coefficient here does not change the proof of the lemma). Hence we will restrict our error probability analysis only to points not belonging to $p\mathbb{Z}^n$ and

$$\begin{aligned} P_e(\mathbf{0}) &\sim \mathcal{P}\{\exists \mathbf{z} \in \Lambda_f \cap \mathcal{B} \setminus p\mathbb{Z}^n \text{ and } \mathbf{z} \neq \mathbf{0}\} \\ &\leq \sum_{\mathbf{z} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{z} \in \Lambda_f\} \\ &= \sum_{\mathbf{z} \in (\mathbb{Z}^n \cap \mathcal{B}) \setminus p\mathbb{Z}^n} \left(\frac{1}{p}\right)^{n(1-R_f)} \\ &\leq |\mathbb{Z}^n \cap \mathcal{B}| \left(\frac{1}{p}\right)^{n(1-R_f)}. \end{aligned} \tag{4.26}$$

The estimation of the previous sum is exactly the same done from (3.15) to (3.18) in the proof of Theorem 3.1, with R_f instead of R. In other words, we already know that its limit is 0 when n tends to infinity, if we simply choose ε to satisfy (3.4).

Passing to the case $\mathbf{s} \neq \mathbf{0}$ and observing that no point of $p\mathbb{Z}^n$ can be the codeword associated to $\mathbf{s}$, we have

$$\begin{aligned} P_e(\mathbf{s}) &= \sum_{\mathbf{x} \in \mathbb{Z}^n} \mathcal{P}\{X_\mathbf{s} = \mathbf{x}, \mathcal{E}_2^c\} \\ &= \sum_{\mathbf{x} \in \mathbb{Z}^n \setminus p\mathbb{Z}^n} \mathcal{P}\{X_\mathbf{s} = \mathbf{x}, \mathcal{E}_2^c\} \\ &\leq \sum_{\mathbf{x} \in \mathbb{Z}^n \setminus p\mathbb{Z}^n} \mathcal{P}\{X_\mathbf{s} = \mathbf{x}, \mathbf{x} \notin \mathcal{B}\} + \sum_{\mathbf{x} \in \mathbb{Z}^n \setminus p\mathbb{Z}^n} \mathcal{P}\{X_\mathbf{s} = \mathbf{x}, \exists \mathbf{z} \in \Lambda_f \cap \mathcal{B} \text{ and } \mathbf{z} \neq \mathbf{x}\}. \end{aligned}$$

Let us call P_1 the left summation and P_2 the right one. We will separately show that they tend to 0 when n tends to infinity, which is enough to conclude.

Estimation of P_1. By the definition of conditional probability,

$$P_1 = \sum_{\mathbf{x} \in \mathbb{Z}^n \setminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{x} \notin \mathcal{B} \mid X_\mathbf{s} = \mathbf{x}\}\mathcal{P}\{X_\mathbf{s} = \mathbf{x}\}.$$

(4.23) tells us that the term $\mathcal{P}\{\mathbf{x} \notin \mathcal{B} \mid X_{\mathbf{s}} = \mathbf{x}\}$ is a vanishing term $T_1(n)$, independently of $\mathbf{x}$. Hence,

$$\begin{aligned} P_1 &\leq T_1(n) \sum_{\mathbf{x} \in \mathbb{Z}^n \setminus p\mathbb{Z}^n} \mathcal{P}\{X_{\mathbf{s}} = \mathbf{x}\} \\ &\leq T_1(n) \to 0. \end{aligned}$$

Estimation of P_2. First of all, notice that, choosing ω as in (4.11), Lemma 4.3 implies that the P_2 restricted to the $\mathbf{x}$'s lying outside the sphere

$$\mathcal{B}_{\text{eff}} = B_{\mathbf{0},n}(\rho_{\text{eff}}(1 + 1/n^{\omega}))$$

tends to 0 when n tends to infinity. For this reason, we will only need to show that

$$\lim_{n \to \infty} \sum_{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}) \setminus p\mathbb{Z}^n} \mathcal{P}\{X_{\mathbf{s}} = \mathbf{x}, \exists \mathbf{z} \in \Lambda_f \cap \mathcal{B} \text{ and } \mathbf{z} \neq \mathbf{x}\} = 0. \tag{4.27}$$

Before going on, let us start by making some considerations about the error probability:

1. First of all, does the point $\mathbf{z} = \mathbf{0} \in \Lambda_f$ typically induce a decoding error? Actually not, since we claim that

   $$\lim_{n \to \infty} \mathcal{P}\{||\alpha \mathbf{y}||^2 > \alpha n \sigma^2 (1 + \varepsilon)^2\} = 1.$$

   This and (4.25) mean that

   $$\lim_{n \to \infty} \mathcal{P}\{||\alpha \mathbf{y} - \mathbf{x}||^2 \leq ||\alpha \mathbf{y}||^2\} = 1$$

   and $\mathbf{x}$ is asymptotically closer to $\alpha \mathbf{y}$ than $\mathbf{0}$. Let us prove the claim: the condition $||\alpha \mathbf{y}||^2 > \alpha n \sigma^2 (1 + \varepsilon)^2$ is equivalent to

   $$||\mathbf{y}||^2 > \frac{n\sigma^2(1+\varepsilon)^2}{\alpha} = \frac{n\sigma^2(P+\sigma^2)(1+\varepsilon)^2}{P} = \frac{n(P+\sigma^2)(1+\varepsilon)^2}{\text{SNR}}.$$

   At the same time, Lemma 4.3, Lemma 2.2 and Lemma 4.4 imply that with probability tending to 1 as n tends to infinity, the event

   $$\mathcal{E}_1' = \{||\mathbf{x}||^2 \geq nP(1-\varepsilon')^2\} \cap \{||\mathbf{w}||^2 \geq n\sigma^2(1-\varepsilon')^2\} \cap \{|\mathbf{x}\mathbf{w}^T| \leq f(n)\sigma||\mathbf{x}||\}$$

   occurs and

   $$\begin{aligned} ||\mathbf{y}||^2 &= ||\mathbf{x}||^2 + ||\mathbf{w}||^2 + 2\mathbf{x}\mathbf{w}^T \\ &\geq nP(1-\varepsilon')^2 + n\sigma^2(1-\varepsilon')^2 - 2f(n)\sigma||\mathbf{x}|| \\ &\geq n(P+\sigma^2)(1-\varepsilon')^2 - 2f(n)\sigma\sqrt{nP}(1+\varepsilon') \\ &= n(P+\sigma^2)\left((1-\varepsilon')^2 - \frac{2f(n)\sigma\sqrt{P}(1+\varepsilon')}{\sqrt{n}(P+\sigma^2)}\right) \\ &\sim n(P+\sigma^2)(1-\varepsilon')^2, \end{aligned}$$

where the last asymptotical equality can be derived with the same observations done for (4.24). Thus, it is sufficient to say that

$$\frac{n(P+\sigma^2)(1+\varepsilon)^2}{\mathrm{SNR}} < n(P+\sigma^2)(1-\varepsilon')^2,$$

which is true because SNR is constant and bigger than 1 by hypothesis and $(1+\varepsilon)^2/(1-\varepsilon')^2$ can be taken to be as close to 1 as wanted, then a fortiori smaller than SNR.

2. The previous argument states that $\mathbf{0}$ does not asymptotically cause any decoding error (with probability tending to 1). We would like now to treat the case of all the other points $\mathbf{z} \in p\mathbb{Z}^n$. Notice that one of these points can be the lattice decoder output only if it is closer to $\alpha\mathbf{y}$ than $\mathbf{0}$ itself. That is, dangerous points $\mathbf{z} \in p\mathbb{Z}^n \setminus \{\mathbf{0}\}$ are such that $||\alpha\mathbf{y} - \mathbf{z}|| \leq ||\alpha\mathbf{y}||$. This implies that there exists $i \in \{1, 2, \ldots, n\}$ such that $|\alpha y_i - z_i| \leq |\alpha y_i|$ and $z_i \neq 0$; then, the fact that $\mathbf{z} \in p\mathbb{Z}^n$ means that $|z_i| \geq p$. Consequently, $|\alpha y_i|$ has to be bigger than $p/2$ and, a fortiori, $|y_i| > p/2$, too, because $\alpha < 1$. Now, $y_i = x_i + w_i$ and a necessary condition for having $|x_i + w_i| > p/2$ is that at least one between $|x_i|$ and $|w_i|$ is bigger than $p/4$. The probability that $|w_i| > p/4$ can be shown to decrease to 0 when n tends to infinity with the same argument used in Lemma 3.1 to treat (3.2). Hence, asymptotically speaking, there can be a decoding error due to points $\mathbf{z} \in p\mathbb{Z}^n \setminus \{\mathbf{0}\}$ only for the $\mathbf{x}$'s such that $|x_i| > p/4$ for some i. Let us show that also this case does not does not represent a real problem: recall that H is the parity-check matrix of Λ and consider the sum

$$\begin{aligned}
&\sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B}_{\mathrm{eff}})\setminus p\mathbb{Z}^n \\ |x_i|>p/4,\ \exists i}} \mathcal{P}\{X_{\mathbf{s}} = \mathbf{x}, \exists \mathbf{z} \in \Lambda_f \cap \mathcal{B} \text{ and } \mathbf{z} \neq \mathbf{x}\} \\
&\qquad \leq \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B}_{\mathrm{eff}})\setminus p\mathbb{Z}^n \\ |x_i|>p/4,\ \exists i}} \mathcal{P}\{X_{\mathbf{s}} = \mathbf{x}\} \\
&\qquad \leq \sum_{\substack{\mathbf{x}\in(\mathbb{Z}^n\cap\mathcal{B}_{\mathrm{eff}})\setminus p\mathbb{Z}^n \\ |x_i|>p/4,\ \exists i}} \mathcal{P}\{H\mathbf{x}^T \equiv \mathbf{s} \bmod p\} \\
&\qquad = |\{\mathbf{x} \in (\mathbb{Z}^n \cap \mathcal{B}_{\mathrm{eff}}) \setminus p\mathbb{Z}^n : |x_i| > p/4, \exists i\}| \left(\frac{1}{p}\right)^{n(1-R)} \\
&\qquad \leq n \left|\mathbb{Z}^{n-1} \cap B_{\mathbf{0},n-1}\left(\sqrt{\rho_{\mathrm{eff}}^2(1+1/n^\omega)^2 - p^2/16}\right)\right| \left(\frac{1}{p}\right)^{n(1-R)} \\
&\qquad \leq n\frac{|\mathbb{Z}^n \cap \mathcal{B}_{\mathrm{eff}}|}{p^{n(1-R)}} \frac{\left(\sqrt{\rho_{\mathrm{eff}}^2(1+1/n^\omega)^2 - p^2/16}\right)^n}{\rho_{\mathrm{eff}}^n\,(1+1/n^\omega)^n} \\
&\qquad = n\frac{|\mathbb{Z}^n \cap \mathcal{B}_{\mathrm{eff}}|}{p^{n(1-R)}} \left(\sqrt{1 - \frac{p^2}{16\rho_{\mathrm{eff}}^2(1+1/n^\omega)^2}}\right)^n.
\end{aligned}$$

Now, the ratio $|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}|/p^{n(1-R)}$ has substantially already been studied from (4.15) to (4.16) (up to a slight modification of a sign in ρ) and it goes to infinity subexponentially with n. Namely, $|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}|/p^{n(1-R)} = O(\exp(n^{1-\omega}))$. On the other hand, a simple computation shows that the square root to the power n is $O(\exp(-Dn^{2\lambda R}))$ for some constant D. Hence the whole quantity tends to 0 as n grows to infinity when $2\lambda R > 1-\omega$, that is $\omega > 1-2\lambda R$. The hypotheses $R > 1/2$ and $\lambda > (1+R)^{-1} > 1/2$ guarantee that the condition on ω does not contradict (4.11).

3. We treat separately also the case of the $\mathbf{z}$'s such that $\mathbf{z} \equiv 2\mathbf{x} \bmod p$. Does this kind of $\mathbf{z}$ induce any decoding error? For what $\mathbf{x}$'s? The strategy to answer these questions is the same that we have adopted in the previous two points. Let us start by considering $\mathbf{z} = 2\mathbf{x}$. There is a decoding error due to $\mathbf{z}$ if $||\alpha\mathbf{y} - \mathbf{2x}||^2 \leq ||\alpha\mathbf{y} - \mathbf{x}||^2$. It is very simple to algebraically show that this is equivalent to say that $(3-2\alpha)||\mathbf{x}||^2 - 2\alpha\mathbf{x}\mathbf{w}^T > 0$. It is then sufficient to show that $||\mathbf{x}||^2 - 2\mathbf{x}\mathbf{w}^T > 0$ with probability tending to 1 when n tends to infinity. If the event $\mathcal{E}_1'$ occurs, then

$$\begin{aligned}
||\mathbf{x}||^2 - 2\mathbf{x}\mathbf{w}^T &\geq nP(1-\varepsilon')^2 - 2f(n)\sigma||\mathbf{x}|| \\
&\geq nP(1-\varepsilon')^2 - 2f(n)\sigma\sqrt{nP}(1-\varepsilon') \\
&= nP(1-\varepsilon')^2\left(1 - \frac{2f(n)\sigma}{\sqrt{nP}(1-\varepsilon')}\right) \\
&> nP(1-\varepsilon')^2\left(1 - \frac{2f(n)}{\sqrt{n}(1-\varepsilon')}\right) \\
&\sim nP(1-\varepsilon')^2,
\end{aligned}$$

where the last inequality is due to the fact that $\text{SNR} = P/\sigma^2 > 1$; the lower bound is clearly asymptotically positive and we are done.

We have proved that $\mathbf{z} = 2\mathbf{x}$ typically does not induce any error. Can we say the same for all the other $\mathbf{z} \equiv 2\mathbf{x} \bmod p$? The only case that can lead to bad decoding is the one of $\mathbf{z} \equiv 2\mathbf{x} \bmod p$ such that $||\alpha\mathbf{y} - \mathbf{z}||^2 < ||\alpha\mathbf{y} - 2\mathbf{x}||^2$ (otherwise, the previous computation concerning $\mathbf{z} = 2\mathbf{x}$ is sufficient). Let $\mathbf{z} = 2\mathbf{x} + p\mathbf{k}$ for some $\mathbf{k} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$. Then $\mathbf{z}$ can be closer to $\alpha\mathbf{y}$ than $2\mathbf{x}$ only if there exists i such that

$$|\alpha y_i - 2x_i| > |\alpha y_i - 2x_i - pk_i|,$$

for some $k_i \geq 1$. This condition implies $|\alpha w_i - (2-\alpha)x_i| = |\alpha y_i - 2x_i| > p/2$, which in turn implies that at least one between $|\alpha w_i|$ and $(2-\alpha)|x_i|$ has to be greater than $p/4$. Now, one can conclude with an argument practically identical to the one applied for the $\mathbf{z}$ of $p\mathbb{Z}^n$ above here.

4. What about the $\mathbf{z}$'s such that $\mathbf{z} \equiv \mathbf{x} \bmod p$? Even if a $\mathbf{z}$ of this kind is closer than $\mathbf{x}$ to $\alpha\mathbf{y}$, its syndrome $H\mathbf{z}^T$ is equal to $\mathbf{s}$, the syndrome of $\mathbf{x}$, and this does not give a decoding error. For this reason, we can actually omit these $\mathbf{z}$'s from the total sum and not consider them.

Concretely, with the previous four points we have shown that we can restrict the sum in (4.27) to the set

$$S = \{\mathbf{x} \in (\mathcal{B}_{\text{eff}} \cap \mathbb{Z}^n) \setminus p\mathbb{Z}^n : \mathbf{z} \equiv \mu\mathbf{x} \bmod p \text{ produces no error}, \forall \mu \in \{0,1,2\}\}.$$

Recall that $H = [(H')^T \mid (H_f)^T]^T$ is the parity-check matrix of Λ, while H_f is the submatrix of H that defines Λ_f. Hence, if $\mathbf{s} = (\mathbf{m} \mid \mathbf{0})^T \in \mathbb{F}_p^{n(R_f-R)} \times \mathbb{F}_p^{n(1-R_f)}$, then the sum that we need to estimate is less than

$$\begin{aligned}
&\sum_{\mathbf{x}\in S} \sum_{\substack{\mathbf{z}\in\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{X_{\mathbf{s}} = \mathbf{x}, \mathbf{z} \in (\Lambda_f \cap \mathcal{B}) \setminus \{\mathbf{x}\}\} \\
&\quad \le \sum_{\mathbf{x}\in S} \sum_{\substack{\mathbf{z}\in\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{H\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p, \mathbf{z} \in \mathcal{B} \setminus \{\mathbf{x}\}\} \\
&\quad = \sum_{\mathbf{x}\in S} \mathcal{P}\{H'\mathbf{x}^T \equiv \mathbf{m}^T \bmod p\} \\
&\qquad\qquad \sum_{\substack{\mathbf{z}\in\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p, \mathbf{z} \in \mathcal{B} \setminus \{\mathbf{x}\}\} \\
&\quad = \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} \\
&\qquad\qquad \sum_{\substack{\mathbf{z}\in\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p, \mathbf{z} \in \mathcal{B} \setminus \{\mathbf{x}\}\} \\
&\quad = \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} \\
&\qquad\qquad \sum_{\substack{\mathbf{z}\in\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{\mathbf{z} \in \mathcal{B} \setminus \{\mathbf{x}\}\},
\end{aligned}$$

where the last equality is true because the events related to the random choice of H_f and the event related to the random noise are independent.

Recall that $\mathcal{B}$ is a random object, that depends on $\mathbf{x}$ and $\mathbf{w}$. We have already observed that $\mathbf{x}$ lies inside it with very high probability. Given this, $\mathbf{z}$ cannot be simultaneously inside the ball and further than twice the radius of $\mathcal{B}$ from $\mathbf{x}$. For this reason we restrict our sum to the $\mathbf{z}$'s inside the sphere $\mathcal{B}' = B_{\mathbf{x},n}(2\rho_{\text{dec}})$. We will show that

$$\lim_{n\to\infty} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} \tag{4.28}$$

$$\sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{\mathbf{z} \in \mathcal{B} \setminus \{\mathbf{x}\}\} = 0. \tag{4.29}$$

There are now two possible situations. If $\mathbf{z} \not\equiv \mu\mathbf{x} \bmod p$ for every $\mu \in \mathbb{F}_p$, then

$$\begin{aligned}&\mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\\ &= \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\\ &= \left(\frac{1}{p}\right)^{2n(1-R_f)}.\end{aligned}$$

If instead $\mathbf{z} \equiv \mu\mathbf{x} \bmod p$ for some $\mu \in \mathbb{F}_p$, the fact that $\mathbf{x}$ belongs to Λ_f automatically implies that $\mathbf{z}$ belongs to Λ_f, too. Hence,

$$\begin{aligned}&\mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, H_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\\ &= \mathcal{P}\{H_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p\}\\ &= \left(\frac{1}{p}\right)^{n(1-R_f)}.\end{aligned}$$

Now, let S' be the subset of S of all the points $\mathbf{x}$ of the random constellation for which there exists at least one $\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}$ such that $\mathbf{z} \equiv \mu\mathbf{x} \bmod p$ (with $\mu \neq 0, 1, 2$ by definition of S). Summarising what we have elaborated till now, we are left to show that

$$\lim_{n\to\infty} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} \sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\not\equiv\mu\mathbf{x} \bmod p}} \left(\frac{1}{p}\right)^{2n(1-R_f)} \mathcal{P}\{\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}\} = 0 \tag{4.30}$$

and

$$\lim_{n\to\infty} \sum_{\mathbf{x}\in S'} \left(\frac{1}{p}\right)^{n(R_f-R)} \sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\equiv\mu\mathbf{x} \bmod p, \mu\neq 0,1,2}} \left(\frac{1}{p}\right)^{n(1-R_f)} \mathcal{P}\{\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}\} = 0. \tag{4.31}$$

Proof of (4.30)**.** As we have already recalled, $\mathcal{B}$ is a random object. Let $\mathcal{S}$ be the set of all the balls of the space of radius ρ_{dec}. Then

$$\begin{aligned}\sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\not\equiv\mu\mathbf{x} \bmod p}} \mathcal{P}\{\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}\} &= \sum_{\substack{\mathbf{z}\in(\mathcal{B}'\cap\mathbb{Z}^n)\smallsetminus\{\mathbf{x}\}\\ \mathbf{z}\not\equiv\mu\mathbf{x} \bmod p}} \sum_{\substack{B\in\mathcal{S}\\ \mathbf{z}\in B}} \mathcal{P}\{\mathcal{B} = B\}\\ &= \sum_{B\in\mathcal{S}} \sum_{\substack{\mathbf{z}\in(\mathcal{B}'\cap B\cap\mathbb{Z}^n)\smallsetminus\{\mathbf{x}\}\\ \mathbf{z}\not\equiv\mu\mathbf{x} \bmod p}} \mathcal{P}\{\mathcal{B} = B\}\\ &\leq \sum_{B\in\mathcal{S}} |B \cap \mathbb{Z}^n| \mathcal{P}\{\mathcal{B} = B\}\\ &\leq \text{Vol}(B_{\mathbf{0},n}(\rho_{\text{dec}} + \sqrt{n}/2)) \sum_{B\in\mathcal{S}} \mathcal{P}\{\mathcal{B} = B\} \qquad (4.32)\\ &= \text{Vol}(B_{\mathbf{0},n}(\rho_{\text{dec}} + \sqrt{n}/2)).\end{aligned}$$

where, of course, (4.32) comes from Lemma 2.3.

Going back to (4.30) and using what we have just deduced, we have

$$\sum_{\mathbf{x}\in S}\left(\frac{1}{p}\right)^{n(R_f-R)}\sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\not\equiv\mu\mathbf{x} \bmod p}}\left(\frac{1}{p}\right)^{2n(1-R_f)}\mathcal{P}\{\mathbf{z}\in\mathcal{B}\smallsetminus\{\mathbf{x}\}\}$$
$$\leq\left(|\mathbb{Z}^n\cap\mathcal{B}_{\text{eff}}|\left(\frac{1}{p}\right)^{n(1-R)}\right)\left(\text{Vol}(B_{\mathbf{0},n}(\rho_{\text{dec}}+\sqrt{n}/2))\left(\frac{1}{p}\right)^{n(1-R_f)}\right). \tag{4.33}$$

As we previously seen, the left factor can be shown to go to infinity subexponentially in n. On the other hand, the right term exponentially decreases to 0, just like (4.26) does. As a result, the dominating term is the latter and the whole product vanishes when n tends to infinity.

Proof of (4.31)**.** We have

$$\sum_{\mathbf{x}\in S'}\left(\frac{1}{p}\right)^{n(R_f-R)}\sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\equiv\mu\mathbf{x} \bmod p,\mu\neq 0,1,2}}\left(\frac{1}{p}\right)^{n(1-R_f)}\mathcal{P}\{\mathbf{z}\in\mathcal{B}\smallsetminus\{\mathbf{x}\}\}$$
$$\leq\sum_{\mathbf{x}\in S'}\left(\frac{1}{p}\right)^{n(R_f-R)}\sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n\\ \mathbf{z}\equiv\mu\mathbf{x} \bmod p,\mu\neq 0,1,2}}\left(\frac{1}{p}\right)^{n(1-R_f)}$$
$$\leq\sum_{\mathbf{x}\in S'}\left(\frac{1}{p}\right)^{n(1-R)}|\{\mathbf{z}\in\mathcal{B}':\mathbf{z}\equiv\mu\mathbf{x} \bmod p,\exists\mu\in\mathbb{F}_p\smallsetminus\{0,1,2\}\}|. \tag{4.34}$$

Recall that for any fixed $\mu\in\mathbb{F}_p$ Lemma 4.2 provides the following upper bound:

$$|\{\mathbf{z}\in\mathcal{B}':\mathbf{z}\equiv\mu\mathbf{x} \bmod p|\leq 1+\frac{16\rho_{\text{dec}}^2}{p^2}\left(\frac{32n\rho_{\text{dec}}^2}{p^2}\right)^{16\rho_{\text{dec}}^2/p^2},$$

hence

$$|\{\mathbf{z}\in\mathcal{B}':\mathbf{z}\equiv\mu\mathbf{x} \bmod p,\exists\mu\in\mathbb{F}_p\smallsetminus\{0,1,2\}|\leq p+\frac{16\rho_{\text{dec}}^2}{p}\left(\frac{32n\rho_{\text{dec}}^2}{p^2}\right)^{16\rho_{\text{dec}}^2/p^2}$$
$$=O\left(n^{(1-2\lambda R_f)n^{(1-2\lambda R_f)}}\right).$$

Let us call $h(n)$ this last term, subexponential in n. Going on from (4.34), we get

$$\sum_{\mathbf{x}\in S'}\left(\frac{1}{p}\right)^{n(1-R)}|\{\mathbf{z}\in\mathcal{B}':\mathbf{z}\equiv\mu\mathbf{x} \bmod p,\exists\mu\in\mathbb{F}_p\smallsetminus\{0,1,2\}\}|\leq\frac{|S'|h(n)}{p^{n(1-R)}}, \tag{4.35}$$

which vanishes asymptotically in n because of Lemma 4.5.

Putting together the estimations of P_1 and P_2, we can derive that

$$\lim_{n\to\infty}P_e(\mathbf{s})=0,$$

quod erat demonstrandum.

□

4.3 Achieving capacity with LDA lattices

We will now adapt the results of the previous section to LDA lattices, that we have introduced in Chapter 3 and that we have shown to be Poltyrev-capacity-achieving (see Theorem 3.1 and Theorem 3.2). In this section, we prove that they can achieve the capacity of the AWGN channel under lattice decoding under similar hypotheses to the ones of Theorem 4.1.

The geometrical approach to demonstrate our result as well as the encoding and decoding scheme will be the very same that we have used for the more general Construction A ensemble of Section 4.2. We will go once again along the same steps that have led to the proof of Theorem 4.1. Some of these will need to be modified and adapted to the LDPC structure of the parity-check matrices of the LDA lattices. In particular, we will extensively employ the expansion properties of the random Tanner graph associated with them.

Like in Theorem 3.2, also for this finite-constellation result the degree of the parity-check equations associated with the LDA lattices is constant. Nevertheless, for some technical reasons and because of some simplifications that lighten the mathematical analysis, the parameters involved in the proof are not completely optimised and the resulting lower bounds for the degree are not the best possible. The same holds for the value of the parameter λ and, as a consequence, the size of the prime number p will have to be bigger than it is in all the other results of this dissertation.

4.3.1 The random LDA lattice codes ensemble

Once again, our lattice codes are given by Voronoi constellations of Construction A lattices. Anyway, this time the parity-check matrix associated with the lattices is sparse, giving rise to LDA lattice codes. So, let

$$H = \left(\frac{H'}{H_f}\right)$$

be a binary matrix of dimension $n(1-R) \times n$ for some $0 < R < 1$. H_f is its lower submatrix, formed by its last $n(1-R_f)$ rows, for some $R < R_f < 1$. Moreover, we take H and H_f to be the skeleton matrices of two (α, A, β, B)-good graphs. So, we suppose that they both have fixed row degree Δ_V and that their column degrees are $\Delta_V/(1-R)$ and $\Delta_V/(1-R_f)$ respectively. Of course, Δ_V has to be big enough to guarantee the expansion properties. We will make some more remarks about the constants α, A, β and B later in Section 4.3.3.

Now, we build the random ensemble of LDA Voronoi constellations by substituting the ones in the skeleton matrix H by random variables uniformly distributed in $\{0, 1, \ldots, p-1\}$, for some prime number p. The random matrix that we obtain is called

$$\mathbf{H} = \left(\frac{\mathbf{H}'}{\mathbf{H}_f}\right)$$

and it is the analogue of the parity-check matrix that generated the more general Construction A lattices of Theorem 4.1 (see (4.4)). The random LDA lattice asso-

ciated with $\mathbf{H}$ (respectively $\mathbf{H}_f$) is called Λ (respectively Λ_f). Of course, they are nested lattices ($\Lambda \subseteq \Lambda_f$) and the Voronoi constellations that we will deal with are given by Λ_f/Λ. Notice that nothing has substantially changed with respect to the construction of the random ensemble of Section 4.2.1, except that our new random matrix is sparse and guarantees some expansion properties of the graphs associated with H' and H_f.

4.3.2 The encoding and decoding scheme

The encoding and decoding scheme that we apply to LDA Voronoi constellations is the same that we have described in Section 4.2.2 and summarised in Figure 4.2 at the beginning of Section 4.2 for the case of more general Construction A lattices. Nothing changes at all and the fact that the lattices that we deal with now are LDA does not affect the information transmission scheme.

4.3.3 Comments on the expansion properties of the Tanner graphs

The expansion properties of the Tanner graph related to the random matrix $\mathbf{H}$ will be the same that we have already considered in Section 3.3, deriving from Lemma 3.3. In other terms, we choose the skeleton matrices H and H_f to be associated with a (α, A, β, B)-good Tanner graphs, like we have done to prove that LDA lattices are Poltyrev-capacity-achieving (Theorem 3.2). Anyway, in this context we will not employ the properties concerning α and β, while we will need some further conditions on A and B and on the constants ε and ϑ that appear in the definition of (α, A, β, B)-goodness (cf. Definition 3.3). These, in turn, imply some hypotheses on Δ_V. Let us make explicit all of these requirements and also recall the hypotheses on the constants needed to guarantee the expansion:

1. A and B have to respect (3.24) both for R_f and R, which becomes:

$$A > 2 \quad \text{and} \quad B > \frac{2}{1-R_f} > \frac{2}{1-R}. \tag{4.36}$$

2. Moreover, we impose that

$$B(1-R)^2 > B(1-R_f)^2 > A, \tag{4.37}$$

   the first inequality being a consequence of the fact that $R_f > R$.

3. We fix ε and ϑ as functions of A and B:

$$\varepsilon = \frac{1-R}{A+1-R} > \frac{1-R_f}{A+1-R_f} \tag{4.38}$$

$$\vartheta = \frac{1}{B(1-R_f)+1} > \frac{1}{B(1-R)+1}. \tag{4.39}$$

4. Δ_V satisfies (3.33), (3.34) and (3.35) with both R_f and R.

Some of these conditions may appear curious for now, but notice that they give rise to some important consequences that we will largely employ in the sequel. In particular, the third condition together with (α, A, β, B)-goodness qualitatively state that every "quite small" subset of variable nodes of the Tanner graph associated with H or H_f has a "quite large" set of neighbours among the parity-check nodes and vice versa. Indeed, consider for example a set of parity-check equation nodes of size a bit greater than $n(1-R)/(B(1-R)+1)$ in the graph associated with H; it is not a "big" set of nodes, but its neighbourhood has size at least $B(1-R)n/(B(1-R)+1)$, that is, it consists of almost the totality of the variable nodes.

Attention: from now on, till the end of the chapter, we will always suppose that the graphs that we work with are (α, A, β, B)-good and that all the involved constants obey the conditions listed above.

4.3.4 LDA lattices achieve capacity - Detailed proof

A useful lemma

In the sequel we will often need to compare the volumes of two (or more) spheres with the same radius, but different dimensions. This lemma contains once for all the computation that leads to this comparison.

Lemma 4.6. *Consider the two balls $B_{\mathbf{c},n}(\rho)$ and $B_{\mathbf{c}',n-m}(\rho)$, with the same given radius ρ, but with different dimensions n and $n-m$. Suppose also that $0 \leq m \leq n/2$. Then, if $\rho > \sqrt{n}/2$,*

$$\frac{|\mathbb{Z}^{n-m} \cap B_{\mathbf{c}',n-m}(\rho)|}{|\mathbb{Z}^n \cap B_{\mathbf{c},n}(\rho)|} \lesssim \frac{(\sqrt{n})^{n+1}}{(\sqrt{n-m})^{n-m+1}} \left(\sqrt{2\pi e}\right)^{-m} \rho^{-m} \left(1 + \frac{2\sqrt{n}}{2\rho - \sqrt{n}}\right)^n.$$

Recall: the notation $f(n) \lesssim g(n)$, that we have already used so far, indicates that $f(n) \sim h(n) \leq g(n)$ for some $h(n)$; or, equivalently, that $f(n) \leq h'(n) \sim g(n)$, for some $h'(n)$.

Proof. The proof of the lemma is a simple application of Lemma 2.3:

$$\begin{aligned}
\frac{|\mathbb{Z}^{n-m} \cap B_{\mathbf{c}',n-m}(\rho)|}{|\mathbb{Z}^n \cap B_{\mathbf{c},n}(\rho)|} &\leq \frac{\mathrm{Vol}\left(B_{\mathbf{c}',n-m}\left(\rho + \frac{\sqrt{n-m}}{2}\right)\right)}{\mathrm{Vol}\left(B_{\mathbf{c},n}\left(\rho - \frac{\sqrt{n}}{2}\right)\right)} \leq \frac{\mathrm{Vol}\left(B_{\mathbf{c}',n-m}(\rho)\right)}{\mathrm{Vol}\left(B_{\mathbf{c},n}(\rho)\right)} \frac{\left(1 + \frac{\sqrt{n}}{2\rho}\right)^n}{\left(1 - \frac{\sqrt{n}}{2\rho}\right)^n} \\
&\sim \frac{(\sqrt{n})^{n+1}}{(\sqrt{n-m})^{n-m+1}} \left(\sqrt{2\pi e}\right)^{-m} \left(\frac{2\rho + \sqrt{n}}{2\rho - \sqrt{n}}\right)^n \rho^{-m} \\
&= \frac{(\sqrt{n})^{n+1}}{(\sqrt{n-m})^{n-m+1}} \left(\sqrt{2\pi e}\right)^{-m} \left(1 + \frac{2\sqrt{n}}{2\rho - \sqrt{n}}\right)^n \rho^{-m}.
\end{aligned}$$

□

The typical norm of a constellation point

The next lemma states that our Voronoi LDA constellation points have the same typical norm of the more general Construction A constellation points of Section 4.2.4. The proof of the lemma traces the demonstration of Lemma 4.3, but needs to be adapted to the LDA setting in which we work. This requires some tricky combinatorial analysis of the structure of the Tanner graph associated with the random lattices. The most interesting argument is probably the variance estimation that starts from (4.48) and goes on till the end of the proof. Similar reasonings will be used in the proof of Theorem 4.2, too.

Like in Section 4.2.4, let ρ_{eff} denote the quantity

$$\rho_{\text{eff}} = \frac{\sqrt{n}p^{(1-R)}}{\sqrt{2\pi e}}.$$

As before, it is the asymptotic effective radius of the lattice associated with the parity-check matrix $\mathbf{H}$.

Lemma 4.7. *In the setting that we have fixed till now in Section 4.3, consider* $\mathbf{s} = (s_1, s_2, \ldots, s_{n(R_f-R)}, 0, \ldots, 0)$ *to be any non-zero syndrome associated with an uncoded message and a constellation point. Suppose that* $p = n^\lambda$ *for some* $\lambda > 0$ *and let* ω *be a positive constant such that*

$$\omega < \min\{\lambda(1-R), 2\lambda R, 1\}.$$

If $\mathbf{x}$ *is the random LDA constellation point whose syndrome is* $\mathbf{s}$ *(cf.* (4.6)*) and if* λ *satisfies*

$$\lambda > \max\left\{\frac{1}{2R}, \frac{1}{1-R}, \frac{2}{A-2}, \frac{2}{B(1-R)-2}, 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1}\right\}, \quad (4.40)$$

then

$$\lim_{n\to\infty} \mathcal{P}\left\{\rho_{\text{eff}}\left(1 - \frac{1}{n^\omega}\right) \leq ||\mathbf{x}|| \leq \rho_{\text{eff}}\left(1 + \frac{1}{n^\omega}\right)\right\} = 1.$$

Proof. Let X_ρ be the random variable that counts the number of points with syndrome $\mathbf{s}$ in the n-dimensional ball $B_{\mathbf{0},n}(\rho)$ centred at $\mathbf{0}$ with radius ρ. For any $\rho \geq 0$ and for any $\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$, consider the random variable

$$X_{\mathbf{x}} = \begin{cases} 1, & \text{if } \mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p \\ 0, & \text{otherwise} \end{cases}.$$

Consequently,

$$X_\rho = \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} X_{\mathbf{x}}.$$

Let $\mathbf{h}_i$ be the i-th row of $\mathbf{H}$, s_i be the i-th coordinate of $\mathbf{s}$ and let $\mathbf{x}$ be a given point of $(\mathbb{Z}^n \setminus p\mathbb{Z}^n) \cap B_{\mathbf{0},n}(\rho)$.

- If $\mathrm{Supp}(\mathbf{x}) \cap \mathrm{Supp}(\mathbf{h}_i) \neq \emptyset$, then $\mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p\} = 1/p$.
- If $\mathrm{Supp}(\mathbf{x}) \cap \mathrm{Supp}(\mathbf{h}_i) = \emptyset$ and $s_i = 0$, then $\mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p\} = 1$.
- If $\mathrm{Supp}(\mathbf{x}) \cap \mathrm{Supp}(\mathbf{h}_i) = \emptyset$ and $s_i \neq 0$, then $\mathcal{P}\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p\} = 0$.

Let ℓ be the number of rows of $\mathbf{H}$ whose support intersects the support of $\mathbf{x}$ and, just like in (3.59), let

$$T_{\mathbf{x}} = \{i \in \{1, 2, \ldots, n(1-R)\} : \mathrm{Supp}(\mathbf{h}_i) \cap \mathrm{Supp}(\mathbf{x}) \neq \emptyset\}.$$

$T_{\mathbf{x}}$ is identified with the set of the parity-check equation nodes of the Tanner graph associated with $\mathbf{H}$ whose support intersects the support of $\mathbf{x}$. Recall that we denote P the set of all the parity-check equation nodes, so $T_{\mathbf{x}} \subseteq P$.

The events $\{\mathbf{h}_i\mathbf{x}^T \equiv s_i \bmod p\}_{i=1,\ldots,n(1-R)}$ are independent, so

$$\mathcal{P}\{X_{\mathbf{x}} = 1\} = \begin{cases} 0, & \text{if } \exists i : \mathrm{Supp}(\mathbf{x}) \cap \mathrm{Supp}(\mathbf{h}_i) = \emptyset \text{ and } s_i \neq 0 \\ \left(\frac{1}{p}\right)^{\ell}, & \text{otherwise} \end{cases}.$$

Moreover, $\mathcal{P}\{X_{\mathbf{x}} = 1\} = 0$ for every $\mathbf{x} \in p\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$, since its syndrome is $\mathbf{0} \neq \mathbf{s}$. This is in particular the case for $\ell = 0$, which coincides with $\mathbf{x} = \mathbf{0} \in p\mathbb{Z}^n$.

Like for Lemma 4.3, we will split the proof into two parts. First of all, we deduce that

$$\lim_{n\to\infty} \mathcal{P}\left\{X_{\rho_{\mathrm{eff}}\left(1-\frac{1}{n^\omega}\right)} > 0\right\} = 0. \tag{4.41}$$

Later, that

$$\lim_{n\to\infty} \mathcal{P}\left\{X_{\rho_{\mathrm{eff}}\left(1+\frac{1}{n^\omega}\right)} = 0\right\} = 0. \tag{4.42}$$

Proof of (4.41)**.** When $\rho = \rho_{\mathrm{eff}}\,(1 - 1/n^\omega)$,

$$\begin{aligned} \mathbb{E}[X_\rho] &= \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}} = 1\} \\ &\leq \sum_{\ell=1}^{n(1-R)} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho) \\ |T_{\mathbf{x}}|=\ell}} \left(\frac{1}{p}\right)^{\ell} \\ &= \sum_{\ell=1}^{\lfloor nB(1-R)/(B(1-R)+1)\rfloor} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho) \\ |T_{\mathbf{x}}|=\ell}} \left(\frac{1}{p}\right)^{\ell} \qquad (4.43) \\ &\quad + \sum_{\ell=\lfloor nB(1-R)/(B(1-R)+1)+1\rfloor}^{n(1-R)} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho) \\ |T_{\mathbf{x}}|=\ell}} \left(\frac{1}{p}\right)^{\ell}. \qquad (4.44) \end{aligned}$$

In order to estimate the first sum, notice that $\ell \leq nB(1-R)/(B(1-R)+1)$ implies that the complement $J_{\mathbf{x}}$ of $T_{\mathbf{x}}$ in the set of parity-check nodes P of the Tanner graph has a big enough cardinality:

$$|J_{\mathbf{x}}| = |P \smallsetminus T_{\mathbf{x}}| \geq \frac{n(1-R)}{B(1-R)+1},$$

which implies by the expansion properties of the graph that its neighbourhood has big size, too (recall the conditions listed in Section 4.3.3 and Lemma 3.3):

$$|N(J_{\mathbf{x}})| \geq B\frac{n(1-R)}{B(1-R)+1}.$$

If we call $S_{\mathbf{x}}$ the complement of $N(J_{\mathbf{x}})$ in the set of variable nodes of the graph V, we have

$$|S_{\mathbf{x}}| = |V \smallsetminus N(J_{\mathbf{x}})| \leq n - |N(J_{\mathbf{x}})| \leq \frac{n}{B(1-R)+1} \leq \frac{n(1-R)}{A+1-R},$$

because $B(1-R)^2 > A$ by condition (4.37). Then, again by the expansion properties,

$$|N(S_{\mathbf{x}})| \geq A|S_{\mathbf{x}}|.$$

But $S_{\mathbf{x}}$ is the complement in V of $N(J_{\mathbf{x}})$, so $N(S_{\mathbf{x}}) \subseteq T_{\mathbf{x}}$ and we obtain

$$|T_{\mathbf{x}}| \geq A|S_{\mathbf{x}}|.$$

Hence,

$$\begin{aligned}|\{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) : |T_{\mathbf{x}}| = \ell\}| &\leq |\{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) : |S_{\mathbf{x}}| \leq \ell/A\}| \\ &\leq \binom{n}{\lfloor \ell/A \rfloor} p^{\ell/A},\end{aligned}$$

because once we fix the coordinates corresponding to $S_{\mathbf{x}}$, each of them cannot take more values than p: indeed, condition $\lambda > (2R)^{-1}$ (cf. (4.40)) implies that asymptotically $p > 2\rho$, that is, p is bigger than the diameter of the sphere $B_{\mathbf{0},n}(\rho)$. For the same reason, the coordinates of $\mathbf{x}$ corresponding to $N(J_{\mathbf{x}})$ are fixed to be equal to 0 modulo p and this can happen only for one integer value. Going back to (4.43):

$$\begin{aligned}&\sum_{\ell=1}^{\lfloor nB(1-R)/(B(1-R)+1) \rfloor} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \\ |T_{\mathbf{x}}| = \ell}} \left(\frac{1}{p}\right)^{\ell} \\ &\leq \sum_{\ell=1}^{\lfloor nB(1-R)/(B(1-R)+1) \rfloor} \binom{n}{\lfloor \ell/A \rfloor} p^{\ell/A} \left(\frac{1}{p}\right)^{\ell} \\ &\leq \sum_{\ell=1}^{\lfloor nB(1-R)/(B(1-R)+1) \rfloor} n^{\ell/A} p^{\ell/A} \left(\frac{1}{p}\right)^{\ell} \\ &= \sum_{\ell=1}^{\lfloor nB(1-R)/(B(1-R)+1) \rfloor} n^{\ell(1/A - \lambda(1-1/A))},\end{aligned}$$

whose limit is 0 when n tends to infinity, because $\lambda > 2(A-2)^{-1} > (A-1)^{-1}$ (cf. condition (4.40)).

The only thing which is left to finish the proof of (4.41) is the estimation of (4.44). Let $u = n(1-R) - \ell$, then we have to bound the sum

$$\sum_{u=0}^{\lfloor n(1-R)/(B(1-R)+1)\rfloor} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\\ |T_{\mathbf{x}}|=n(1-R)-u}} \left(\frac{1}{p}\right)^{n(1-R)-u}. \tag{4.45}$$

Notice that u is the cardinality of $J_{\mathbf{x}}$ and $u \le n(1-R)/(B(1-R)+1)$ implies that $|N(J_{\mathbf{x}})| \ge B|J_{\mathbf{x}}|$. Once $J_{\mathbf{x}}$ is fixed, at least $B|J_{\mathbf{x}}|$ coordinates of $\mathbf{x}$ are equal to 0 (modulo p). Hence

$$\begin{aligned}|\{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho) : |T_{\mathbf{x}}| = n(1-R)-u\}| &= |\{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho) : |J_{\mathbf{x}}| = u\}|\\ &\le \binom{n(1-R)}{u}|\mathbb{Z}^{n-Bu}\cap B_{\mathbf{0},n-Bu}(\rho)|\\ &\le n^u|\mathbb{Z}^{n-Bu}\cap B_{\mathbf{0},n-Bu}(\rho)|.\end{aligned}$$

Applying Lemma 4.6 and substituting the real value of ρ to obtain (4.46), we deduce that

$$\begin{aligned}(4.45) &\le \sum_{u=0}^{\lfloor n(1-R)/(B(1-R)+1)\rfloor} n^u|\mathbb{Z}^{n-Bu}\cap B_{\mathbf{0},n-Bu}(\rho)|\left(\frac{1}{p}\right)^{n(1-R)-u}\\ &= \sum_{u=0}^{\lfloor n(1-R)/(B(1-R)+1)\rfloor} n^u p^u \frac{|\mathbb{Z}^{n-Bu}\cap B_{\mathbf{0},n-Bu}(\rho)|}{|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)|}|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)|\left(\frac{1}{p}\right)^{n(1-R)}\\ &\lesssim |\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)|\left(\frac{1}{p}\right)^{n(1-R)} \qquad (4.46)\\ &\quad\cdot \sum_{u=0}^{\lfloor n(1-R)/(B(1-R)+1)\rfloor}\left(1-\frac{1}{n^\omega}\right)^{-Bu}\left(\sqrt{\frac{n}{n-Bu}}\right)^{n-Bu+1}\frac{n^u p^u}{p^{(1-R)Bu}} \qquad (4.47)\\ &= |\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)|\left(\frac{1}{p}\right)^{n(1-R)}\cdot\\ &\quad\cdot \sum_{u=0}^{\lfloor n(1-R)/(B(1-R)+1)\rfloor}\left(1-\frac{1}{n^\omega}\right)^{-Bu}\left(1+\frac{Bu}{n-Bu}\right)^{(n-Bu+1)/2} n^{u(1-\lambda(B(1-R)-1))}.\end{aligned}$$

Now, it is easy to show (and we leave the details to the reader) that

$$\left(1-\frac{1}{n^\omega}\right)^{-Bu}\left(1+\frac{Bu}{n-Bu}\right)^{(n-Bu+1)/2} n^{u(1-\lambda(B(1-R)-1))} \le 1$$

and, in particular, it is $o(1)$ whenever $u > 0$, provided that $1-\lambda(B(1-R)-1) < 0$. This is guaranteed by condition $\lambda > 2(B(1-R)-2)^{-1}$ (see (4.40)) and by the fact

that $B(1-R)-1>0$ by condition (4.36). Moreover, we already know that

$$\lim_{n\to\infty} |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)| \left(\frac{1}{p}\right)^{n(1-R)} = 0,$$

since we have already carried out this computation in the proof of Lemma 4.3, from (4.15) to (4.16). Consequently, the whole sum tends to 0 when n tends to infinity. Furthermore, this is true for the sums of (4.43) and (4.44), too, and finally, considering that $\mathcal{P}\{X_\rho > 0\} \leq \mathbb{E}[X_\rho]$, we have

$$\lim_{n\to\infty} \mathcal{P}\left\{X_{\rho_{\text{eff}}\left(1-\frac{1}{n^\omega}\right)} > 0\right\} = 0.$$

Proof of (4.42)**.** Now, let $\rho = \rho_{\text{eff}}\,(1+1/n^\omega)$. We have

$$\begin{aligned}
\mathbb{E}[X_\rho] &= \sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}}=1\} \\
&\geq \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |T_{\mathbf{x}}|=n(1-R)}} \left(\frac{1}{p}\right)^{n(1-R)} \\
&\geq \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ \forall i, x_i\neq 0}} \left(\frac{1}{p}\right)^{n(1-R)} \\
&= |\{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n : x_i\neq 0, \forall i=1,2,\ldots,n\}| \left(\frac{1}{p}\right)^{n(1-R)} \\
&= \left(|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n| - |\{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n : x_i=0, \exists i\}|\right) \left(\frac{1}{p}\right)^{n(1-R)} \\
&\geq \left(|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n| - \sum_{i=1}^{n} |\{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n : x_i=0\}|\right) \left(\frac{1}{p}\right)^{n(1-R)} \\
&= |\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n| \left(1 - n\frac{|\mathbb{Z}^{n-1}\cap B_{\mathbf{0},n-1}(\rho)\smallsetminus p\mathbb{Z}^{n-1}|}{|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n|}\right) \left(\frac{1}{p}\right)^{n(1-R)}.
\end{aligned}$$

Now, we have already computed from (4.18) to (4.19) that

$$\lim_{n\to\infty} |\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n| \left(\frac{1}{p}\right)^{n(1-R)} = +\infty.$$

What about

$$n\frac{|\mathbb{Z}^{n-1}\cap B_{\mathbf{0},n-1}(\rho)\smallsetminus p\mathbb{Z}^{n-1}|}{|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n|}?$$

By Lemma 4.6, introducing the actual value of ρ and recalling that $\lambda > (1-R)^{-1}$ (see (4.40)), we can deduce that

$$n\frac{|\mathbb{Z}^{n-1}\cap B_{\mathbf{0},n-1}(\rho)\smallsetminus p\mathbb{Z}^{n-1}|}{|\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n|} \lesssim \frac{n}{p^{(1-R)}}\left(\sqrt{\frac{n}{n-1}}\right)^n \sim \frac{\sqrt{e}n}{p^{1-R}} \to 0.$$

This allows us to conclude that

$$\lim_{n\to\infty} \mathbb{E}[X_\rho] = +\infty.$$

After that, we need to carry out a detailed estimation of $\mathrm{Var}(X_\rho)$, like we have done in the proof of Lemma 4.3 for the more general Construction A constellations. We have

$$\begin{aligned}
\mathrm{Var}(X_\rho) &= \mathrm{Var}\left(\sum_{\mathbf{x}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} X_{\mathbf{x}}\right) \qquad (4.48)\\
&= \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{z}})\\
&\leq \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathbb{E}[X_{\mathbf{x}}X_{\mathbf{z}}]\\
&= \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}}X_{\mathbf{z}} = 1\}\\
&= \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{X_{\mathbf{x}} = 1, X_{\mathbf{z}} = 1\}\\
&= \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, \mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\}\\
&= \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\setminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, \mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\},
\end{aligned}$$

where the points of $p\mathbb{Z}^n$ are excluded because the probability that their syndrome is $\mathbf{s} \neq \mathbf{0}$ is 0.

Now, let $\mathbf{h}$ be a generic row of $\mathbf{H}$; it represents a parity-check equation and we also write $\mathbf{h} \in P$, where P is the set of vertices of the Tanner graph associated with $\mathbf{H}$ corresponding to the parity-check equations (recall that the notation is the same of Section 3.3.2). This is a little abuse in notation, but it will help us to transfer some arguments onto the Tanner graph and make clearer our demonstration. For a given $\mathbf{x} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$, let $\mathbf{x}_\mathbf{h}$ be the subvector of $\mathbf{x}$ made only of the coordinates of $\mathbf{x}$ itself that belong to the neighbourhood $N(\mathbf{h})$ of $\mathbf{h}$ in the graph. In other words, these are the coordinates of $\mathbf{x}$ that correspond to ones in the row of the skeleton matrix corresponding to $\mathbf{h}$.

Let us fix $\mathbf{x}, \mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$ and a row $\mathbf{h}$ of $\mathbf{H}$ and consider the vector space generated by $\mathbf{x}_\mathbf{h}$ and $\mathbf{z}_\mathbf{h}$, which has dimension 0, 1 or 2 over $\mathbb{R}$. We call the latter $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h})$. Hence, denoting s the syndrome coordinate corresponding to $\mathbf{h}$, we have:

- if $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 0$ and $s = 0$, then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv s \bmod p, \mathbf{h}\mathbf{z}^T \equiv s \bmod p\} = 1$;
- if $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 0$ and $s \neq 0$, then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv s \bmod p, \mathbf{h}\mathbf{z}^T \equiv s \bmod p\} = 0$;
- if $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 1$ and $s = 0$, then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv s \bmod p, \mathbf{h}\mathbf{z}^T \equiv s \bmod p\} = 1/p$;

- if $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 1$ and $s \neq 0$, then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv s \bmod p, \mathbf{h}\mathbf{z}^T \equiv s \bmod p\} = 1/p$ if $\mathbf{z}_\mathbf{h} \equiv \mathbf{x}_\mathbf{h} \bmod p$, otherwise it is 0;
- if $\dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 2$, then $\mathcal{P}\{\mathbf{h}\mathbf{x}^T \equiv s \bmod p, \mathbf{h}\mathbf{z}^T \equiv s \bmod p\} = 1/p^2$.

Summarising, given $\mathbf{x}, \mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)$, let us consider the partition of the set of parity-check equations P given by the following three sets:

$$\begin{aligned} J_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \in P : \dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 0\}, \\ I_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \in P : \dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 1\}, \\ T_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \in P : \dim(\mathbf{x}, \mathbf{z}, \mathbf{h}) = 2\}; \end{aligned}$$

then, recalling that $\mathbf{H}$ has $n(1-R) = |P|$ rows,

$$\mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, \mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\} \leq \left(\frac{1}{p}\right)^{2n(1-R)-|I_{\mathbf{x},\mathbf{z}}|-2|J_{\mathbf{x},\mathbf{z}}|} = \left(\frac{1}{p}\right)^{2|T_{\mathbf{x},\mathbf{z}}|+|I_{\mathbf{x},\mathbf{z}}|}.$$

More precisely, if the equality above does not hold, then the probability is 0. Thanks to what we have just pointed out, we can write

$$\begin{aligned} \mathrm{Var}(X_\rho) &\leq \sum_{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \setminus p\mathbb{Z}^n} \mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, \mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\} \\ &\leq \sum_{\substack{i,j,t \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \setminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i, |J_{\mathbf{x},\mathbf{z}}|=j, |T_{\mathbf{x},\mathbf{z}}|=t}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j} \\ &= \sum_{\substack{i,j,t \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \setminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i, |J_{\mathbf{x},\mathbf{z}}|=j, |T_{\mathbf{x},\mathbf{z}}|=t}} \left(\frac{1}{p}\right)^{2t+i}. \end{aligned}$$

We will split the sum into two different parts, corresponding to two different ranges of j. Namely:

1. $j > \frac{n(1-R)}{B(1-R)+1}$;
2. $j \leq \frac{n(1-R)}{B(1-R)+1}$.

First case: $j > n(1-R)/(B(1-R)+1)$. Since the graph is (α, A, β, B)-good and because of the considerations done in Section 4.3.3 about its expansion properties, we have that

$$|N(J_{\mathbf{x},\mathbf{z}})| \geq nB(1-R)/(B(1-R)+1),$$

which means that the neighbourhood of $J_{\mathbf{x},\mathbf{z}}$ in the Tanner graph consists of almost the totality of the variable nodes. If we call $S_{\mathbf{x},\mathbf{z}}$ its complement in the set of variable nodes V, $S_{\mathbf{x},\mathbf{z}} = V \setminus N(J_{\mathbf{x},\mathbf{z}})$, then

$$|S_{\mathbf{x},\mathbf{z}}| \leq \frac{n}{B(1-R)+1} \leq \frac{n(1-R)}{A+1-R},$$

because $B(1-R)^2 \geq A$ by hypothesis (cf. (4.37)). Moreover, the (α, A, β, B)-goodness of the graph implies that $|N(S_{\mathbf{x},\mathbf{z}})| \geq A|S_{\mathbf{x},\mathbf{z}}|$. Observing that $N(S_{\mathbf{x},\mathbf{z}}) \subseteq I_{\mathbf{x},\mathbf{z}} \cup T_{\mathbf{x},\mathbf{z}}$, we deduce that $|I_{\mathbf{x},\mathbf{z}} \cup T_{\mathbf{x},\mathbf{z}}| = |I_{\mathbf{x},\mathbf{z}}| + |T_{\mathbf{x},\mathbf{z}}| \geq A|S_{\mathbf{x},\mathbf{z}}|$.

Now, notice that by definition of $J_{\mathbf{x},\mathbf{z}}$ the coordinates of $\mathbf{x}$ and $\mathbf{z}$ corresponding to $N(J_{\mathbf{x},\mathbf{z}})$ have to be congruent to 0 modulo p. Hence, they are fixed, because as we recalled before, for every class modulo p there is at most one single value that can be taken by a coordinate of an integer point in the sphere we are dealing with. For this reason, the only coordinates that can take more than one value are the ones corresponding to $S_{\mathbf{x},\mathbf{z}}$. This helps us in counting the number of $\mathbf{x}$'s and $\mathbf{z}$'s such that $|J_{\mathbf{x},\mathbf{z}}| = j$ and $|S_{\mathbf{x},\mathbf{z}}| = n - |N(J_{\mathbf{x},\mathbf{z}})|$:

$$\begin{aligned} |\{\mathbf{x}, \mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n : |J_{\mathbf{x},\mathbf{z}}| = j\}| &\leq \binom{n}{|S_{\mathbf{x},\mathbf{z}}|} |\mathbb{Z}^{|S_{\mathbf{x},\mathbf{z}}|} \cap B_{\mathbf{0},|S_{\mathbf{x},\mathbf{z}}|}(\rho)|^2 \\ &\leq n^{|S_{\mathbf{x},\mathbf{z}}|} p^{2|S_{\mathbf{x},\mathbf{z}}|}, \end{aligned}$$

where once again we have used the fact that the number of possible values taken by a single coordinate of $\mathbf{x}$ or $\mathbf{z}$ is bounded by p. Then, using the fact that

$$2|T_{\mathbf{x},\mathbf{z}}| + |I_{\mathbf{x},\mathbf{z}}| \geq |T_{\mathbf{x},\mathbf{z}}| + |I_{\mathbf{x},\mathbf{z}}| \geq A|S_{\mathbf{x},\mathbf{z}}| = A(n - |N(J_{\mathbf{x},\mathbf{z}})|),$$

we can bound our summation in the following way:

$$\sum_{\substack{i,j,t \\ j>n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i, |J_{\mathbf{x},\mathbf{z}}|=j, |T_{\mathbf{x},\mathbf{z}}|=t}} \left(\frac{1}{p}\right)^{2t+i} \tag{4.49}$$

$$\leq n \sum_{j=\lceil n(1-R)/(B(1-R)+1) \rceil}^{n(1-R)} \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n \\ |J_{\mathbf{x},\mathbf{z}}|=j}} \left(\frac{1}{p}\right)^{A(n-|N(J_{\mathbf{x},\mathbf{z}})|)}$$

$$\leq n \sum_{s=1}^{\lfloor n(1-R)/(A+1-R) \rfloor} \sum_{\substack{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n \\ |S_{\mathbf{x},\mathbf{z}}|=s}} \left(\frac{1}{p}\right)^{As} \tag{4.50}$$

$$\leq n \sum_{s=1}^{\lfloor n(1-R)/(A+1-R) \rfloor} n^s p^{2s} \left(\frac{1}{p}\right)^{As}$$

$$= \sum_{s=1}^{\lfloor n(1-R)/(A+1-R) \rfloor} n^{1+s(1-\lambda(A-2))}$$

$$\leq \sum_{s=1}^{\lfloor n(1-R)/(A+1-R) \rfloor} n^{s(2-\lambda(A-2))} \to 0. \tag{4.51}$$

This geometric series tends to 0 when n goes to infinity, because of condition (4.40):

$$\lambda > \frac{2}{A-2}.$$

Second case: $j \leq n(1-R)/(B(1-R)+1)$**.** Now j is "small" and the expansion properties imply that $|N(J_{\mathbf{x},\mathbf{z}})| \geq B|J_{\mathbf{x},\mathbf{z}}|$.

Before estimating the sum, we will need to investigate the structure of $I_{\mathbf{x},\mathbf{z}}$ and its neighbourhood. For this purpose, consider the graph $\mathcal{G}'_{\mathbf{x},\mathbf{z}}$ that consists of the bipartite subgraph of the whole Tanner graph (called $\mathcal{G}$) given by the parity-check equation nodes of $I_{\mathbf{x},\mathbf{z}}$, the variable nodes of $N(I_{\mathbf{x},\mathbf{z}})$, and the edges connecting them. A priori, $\mathcal{G}'_{\mathbf{x},\mathbf{z}}$ can be made of many different (bipartite) connected components, depending for example on the size of $I_{\mathbf{x},\mathbf{z}}$ (even if $\mathcal{G}$ is connected with very high probability, tending to 1 when n tends to infinity). The set of vertices of each one of these components is made of a subset of $N(I_{\mathbf{x},\mathbf{z}})$ (variable nodes) and a subset of $I_{\mathbf{x},\mathbf{z}}$ (parity-check equation nodes). The connected components can be (trivially) partitioned into two kinds: the ones whose set of parity-check equations has size bigger than $n(1-R)/(B(1-R)+1)$ and the ones for which this does not hold. So, if $\mathcal{C}$ is the generic connected component and $P_{\mathcal{C}} \subseteq P$ is its set of parity-check equation nodes, let us define:

$$\begin{aligned}\mathcal{K}_{\mathbf{x},\mathbf{z}} &= \{\mathcal{C} \subseteq \mathcal{G}'_{\mathbf{x},\mathbf{z}} : |P_{\mathcal{C}}| \leq n(1-R)/(B(1-R)+1)\} \text{ and} \\ \mathcal{M}_{\mathbf{x},\mathbf{z}} &= \{\mathcal{C} \subseteq \mathcal{G}'_{\mathbf{x},\mathbf{z}} : |P_{\mathcal{C}}| > n(1-R)/(B(1-R)+1)\}. \end{aligned} \tag{4.52}$$

Of course, $\mathcal{G}'_{\mathbf{x},\mathbf{z}} = \mathcal{K}_{\mathbf{x},\mathbf{z}} \cup \mathcal{M}_{\mathbf{x},\mathbf{z}}$ and the union is disjoint. If we define

$$\begin{aligned}K_{\mathbf{x},\mathbf{z}} &= \bigcup \{P_{\mathcal{C}} : \mathcal{C} \in \mathcal{K}_{\mathbf{x},\mathbf{z}}\} \subseteq P \text{ and} \\ M_{\mathbf{x},\mathbf{z}} &= \bigcup \{P_{\mathcal{C}} : \mathcal{C} \in \mathcal{M}_{\mathbf{x},\mathbf{z}}\} \subseteq P,\end{aligned}$$

then we can also write $I_{\mathbf{x},\mathbf{z}} = K_{\mathbf{x},\mathbf{z}} \cup M_{\mathbf{x},\mathbf{z}}$ and the union is disjoint, too.

Now, every $P_{\mathcal{C}} \subseteq \mathcal{C} \in \mathcal{K}_{\mathbf{x},\mathbf{z}}$ is such that $|N(P_{\mathcal{C}})| \geq B|P_{\mathcal{C}}|$, so this holds for the whole $K_{\mathbf{x},\mathbf{z}}$, too:

$$|N(K_{\mathbf{x},\mathbf{z}})| \geq B|K_{\mathbf{x},\mathbf{z}}|. \tag{4.53}$$

Another useful observation is that $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| \leq 1$; in other words, there cannot be more than one connected component whose parity-check equation set is too big. Indeed, every one of these sets is such that its neighbourhood has size at least $B(1-R)n/(B(1-R)+1)$. If there were two (or more) connected components in $\mathcal{M}_{\mathbf{x},\mathbf{z}}$, the union of these neighbourhoods would be greater than the whole set of variable nodes of the Tanner graph itself, which is not possible.

We will consider separately the two cases $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$ and $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 1$ and split

the summation into two more parts:

$$\sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j}$$

$$= \sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t \\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=0}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j} + \quad (4.54)$$

$$+ \sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t \\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=1}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j} \quad (4.55)$$

A small remark before proceeding: a priori, we are summing also over the $\mathbf{x}s$ and $\mathbf{z}s$ such that $|I_{\mathbf{x},\mathbf{z}}| = 0 = |J_{\mathbf{x},\mathbf{z}}|$. This implies that $|T_{\mathbf{x},\mathbf{z}}| = n(1-R)$ and that

$$\mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p, \mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\} = \mathcal{P}\{\mathbf{H}\mathbf{x}^T \equiv \mathbf{s}^T \bmod p\}\mathcal{P}\{\mathbf{H}\mathbf{z}^T \equiv \mathbf{s}^T \bmod p\}.$$

The consequence is that in this case $\mathrm{Cov}(X_{\mathbf{x}}, X_{\mathbf{x}}) = 0$ and the actual contribution to the variance of these couples of $\mathbf{x}s$ and $\mathbf{z}s$ is null. Consequently, we will suppose from now on that $i + j$ is always different from 0. We will recall this observation when needed in the sequel.

1. If $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$, then $I_{\mathbf{x},\mathbf{z}} = K_{\mathbf{x},\mathbf{z}}$ and $|N(I_{\mathbf{x},\mathbf{z}})| \geq B|I_{\mathbf{x},\mathbf{z}}|$. For fixed i, j and t, let us estimate the number of $\mathbf{x}$'s and $\mathbf{z}$'s such that $|I_{\mathbf{x},\mathbf{z}}| = i, |J_{\mathbf{x},\mathbf{z}}| = j$ and $|T_{\mathbf{x},\mathbf{z}}| = t$ in this case. We have already pointed out that $|N(J_{\mathbf{x},\mathbf{z}})| \geq B|J_{\mathbf{x},\mathbf{z}}|$, too, because j is "small". This implies that, just like in the previous case, at least $B|J_{\mathbf{x},\mathbf{z}}|$ coordinates of $\mathbf{x}$ are fixed to 0 (modulo p). Choosing these coordinates is equivalent to choosing the parity-check equations of $|J_{\mathbf{x},\mathbf{z}}|$.

   On the other hand, what can we say about $\mathbf{z}$? Observe that, by definition, $\mathbf{x}_{\mathbf{h}}$ and $\mathbf{z}_{\mathbf{h}}$ are multiple modulo p for every parity-check equation $\mathbf{h}$ that corresponds to a vertex of $I_{\mathbf{x},\mathbf{z}}$. Hence, for a fixed $\mathbf{x}$, the $\mathbf{z}$'s that we take into account cannot take more than p different values with respect to $\mathbf{x}$ in the coordinates that correspond to $N(I_{\mathbf{x},\mathbf{z}})$ (and we know that these coordinates are at least $B|I_{\mathbf{x},\mathbf{z}}|$). Fixing them is the same as fixing the parity-check equations of $|I_{\mathbf{x},\mathbf{z}}|$.

   Putting together all of these observations, we obtain:

$$\begin{aligned}
&|\{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n : |I_{\mathbf{x},\mathbf{z}}| = i, |J_{\mathbf{x},\mathbf{z}}| = j, |T_{\mathbf{x},\mathbf{z}}| = t\}| \\
&\leq \binom{n(1-R)}{j} |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho)| \binom{n(1-R)}{i} p^i |\mathbb{Z}^{n-Bi} \cap B_{\mathbf{0},n-Bi}(\rho)| \\
&\leq n^{(j+i)} p^i |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho)||\mathbb{Z}^{n-Bi} \cap B_{\mathbf{0},n-Bi}(\rho)|.
\end{aligned}$$

Let us define the quantity

$$\mathcal{E}(\rho) = \sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)} \left(\frac{1}{p}\right)^{2n(1-R)} = |\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|^2 \left(\frac{1}{p}\right)^{2n(1-R)} \lesssim \mathbb{E}[X_\rho]^2.$$

We will use it in the following estimation:

$$\begin{aligned}
&\sum_{\substack{i,j,t\\ j\le n(1-R)/(B(1-R)+1)\\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\setminus p\mathbb{Z}^n\\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t\\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=0}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j}\\
&= \sum_{\substack{i,j,t\\ j\le n(1-R)/(B(1-R)+1)\\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\setminus p\mathbb{Z}^n\\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t\\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=0}} \frac{\mathcal{E}(\rho)}{\mathcal{E}(\rho)} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j}\\
&= \sum_{\substack{i,j,t\\ j\le n(1-R)/(B(1-R)+1)\\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\setminus p\mathbb{Z}^n\\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t\\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=0}} \frac{\mathcal{E}(\rho)}{|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|^2} p^{i+2j}\\
&\le \sum_{\substack{i,j,t\\ j\le n(1-R)/(B(1-R)+1)\\ i+j+t=n(1-R)}} \frac{|\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho)||\mathbb{Z}^{n-Bi} \cap B_{\mathbf{0},n-Bi}(\rho)|}{|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|^2} \cdot\\
&\qquad\qquad \cdot n^{(j+i)} p^i p^{i+2j} \mathcal{E}(\rho)\\
&\lesssim \sum_{\substack{i,j,t\\ j\le n(1-R)/(B(1-R)+1)\\ i+j+t=n(1-R)}} f(n) p^{-B(1-R)(j+i)} n^{(j+i)} p^{2(i+j)} \mathcal{E}(\rho), \qquad (4.56)
\end{aligned}$$

where the last asymptotical inequality comes from Lemma 4.6 and

$$\begin{aligned}
f(n) &= \frac{(\sqrt{n})^{2(n+1)}}{(\sqrt{n-Bj})^{n-Bj+1}(\sqrt{n-Bi})^{n-Bi+1}} \left(\sqrt{2\pi e}\right)^{-B(j+i)} \cdot\\
&\quad \cdot \left(1 + \frac{2\sqrt{n}}{2\rho - \sqrt{n}}\right)^{2n} \frac{\rho^{-B(j+i)}}{p^{-B(1-R)(j+i)}}\\
&= \frac{(\sqrt{n})^{2(n+1)-B(j+i)}}{(\sqrt{n-Bj})^{n+1-Bj}(\sqrt{n-Bi})^{n+1-Bi}} \left(1 + \frac{2\sqrt{n}}{2\rho - \sqrt{n}}\right)^{2n} \left(1 + \frac{1}{n^\omega}\right)^{-B(j+i)},
\end{aligned}$$

recalling that

$$\rho = \frac{\sqrt{n} p^{(1-R)}}{\sqrt{2\pi e}} \left(1 + \frac{1}{n^\omega}\right).$$

Let us go back to (4.56): besides $f(n)$ and $\mathcal{E}(\rho)$, in the sum we have

$$p^{-B(1-R)(j+i)} n^{(j+i)} p^{2(i+j)} = n^{(j+i)(1-\lambda(B(1-R)-2))}$$

and the exponent is strictly negative because (4.36) and (4.40) impose that

$$B > \frac{2}{1-R} \quad \text{and} \quad \lambda > \frac{2}{B(1-R)-2} > \frac{1}{B(1-R)-2} \tag{4.57}$$

(recall also that $j+i \neq 0$, because the contribution to the variance of couples of $\mathbf{x}$'s and $\mathbf{z}$'s corresponding to $i=j=0$ is actually 0).

What can we say about $f(n)$? First of all that

$$\left(1+\frac{2\sqrt{n}}{2\rho-\sqrt{n}}\right)^{2n} \leq \left(1+\frac{2\sqrt{2\pi e}}{2p^{(1-R)}-\sqrt{2\pi e}}\right)^{2n} \to 1,$$

because we have imposed that $\lambda > (1-R)^{-1}$ (see always (4.40)). Moreover,

$$\left(1+\frac{1}{n^{\omega}}\right)^{-B(j+i)} \leq 1.$$

Now, consider the term

$$f_j(n) = \left(\sqrt{\frac{n}{n-Bj}}\right)^{n-Bj+1} \sim \left(1+\frac{Bj}{n-Bj}\right)^{\frac{n-Bj+1}{2}};$$

it is easy to show (and we leave this task to the reader) that if $j \neq 0$

$$f_j(n) n^{j(1-\lambda(B(1-R)-2))} = o(1),$$

otherwise it is 1. Symmetrically, if $i \neq 0$,

$$f_i(n) n^{i(1-\lambda(B(1-R)-2))} = o(1),$$

otherwise it is 1, but at least one of them is vanishing when n tends to infinity. Consequently, the whole $f(n) \sim f_j(n) f_i(n)$ is always dominated by the main term $n^{(j+i)(1-\lambda(B(1-R)-2))}$. Furthermore, it is straightforward to see that conditions (4.57) are actually sufficient to conclude that

$$\sum_{\substack{i,j,t \\ j \leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} f(n) p^{-B(1-R)(j+i)} n^{(j+i)} p^{2(i+j)} \mathcal{E}(\rho) \lesssim o(1)\mathcal{E}(\rho). \tag{4.58}$$

We will need this inequality later, after the estimation of the variance for the case $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 1$.

2. If $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 1$, then $\mathcal{G}'_{\mathbf{x},\mathbf{z}}$ contains a "big" connected component and $|M_{\mathbf{x},\mathbf{z}}| > n(1-R)/(B(1-R)+1)$, which implies that

$$|N(I_{\mathbf{x},\mathbf{z}} \cup J_{\mathbf{x},\mathbf{z}})| \geq |N(I_{\mathbf{x},\mathbf{z}})| \geq |N(M_{\mathbf{x},\mathbf{z}})| \geq \frac{B(1-R)}{B(1-R)+1} n.$$

If we call $R_{\mathbf{x},\mathbf{z}}$ the complement of $N(I_{\mathbf{x},\mathbf{z}} \cup J_{\mathbf{x},\mathbf{z}})$ in the set V of all the variable nodes of the Tanner graph, we have that

$$|R_{\mathbf{x},\mathbf{z}}| \leq n/(B(1-R)+1) \leq (1-R)n/(A+1-R).$$

Moreover, $N(R_{\mathbf{x},\mathbf{z}}) \subseteq T_{\mathbf{x},\mathbf{z}}$ and the expansion properties of the graph guarantee that $|N(R_{\mathbf{x},\mathbf{z}})| \geq A|R_{\mathbf{x},\mathbf{z}}|$, from which we deduce that

$$|T_{\mathbf{x},\mathbf{z}}| + |J_{\mathbf{x},\mathbf{z}}| \geq |T_{\mathbf{x},\mathbf{z}}| \geq A|R_{\mathbf{x},\mathbf{z}}|.$$

This will help us in counting the number of $\mathbf{x}$'s and $\mathbf{z}$'s such that $|I_{\mathbf{x},\mathbf{z}}|, |J_{\mathbf{x},\mathbf{z}}|$ and $|T_{\mathbf{x},\mathbf{z}}|$ are fixed to be respectively i, j and t. First of all, the same argument of the case $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$ holds for the $\mathbf{x}$'s: at least $B|J_{\mathbf{x},\mathbf{z}}|$ of their coordinates are fixed to be 0 (modulo p) and these coordinates are identified by the parity-check equations of $J_{\mathbf{x},\mathbf{z}}$. Concerning the $\mathbf{z}$'s, given a fixed $\mathbf{x}$, their coordinates are fixed to 0 in the neighbourhood of $J_{\mathbf{x},\mathbf{z}}$ and can take up to p different values in the neighbourhood of $I_{\mathbf{x},\mathbf{z}}$ (this values are the multiples modulo p of the coordinates of $\mathbf{x}$). If we call

$$r = |R_{\mathbf{x},\mathbf{z}}| \qquad \text{and} \qquad k = |K_{\mathbf{x},\mathbf{z}}|,$$

all of this allows us to conclude that

$$\begin{aligned}
&|\{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \setminus p\mathbb{Z}^n : |I_{\mathbf{x},\mathbf{z}}| = i, |J_{\mathbf{x},\mathbf{z}}| = j, |T_{\mathbf{x},\mathbf{z}}| = t\}| \\
&\leq \binom{n(1-R)}{j} |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}| \binom{n(1-R)}{t} p^{k+1} |\mathbb{Z}^r \cap B_{\mathbf{0},r}(\rho)| \\
&\leq n^{j+t} p^{k+1+r} |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}|. \qquad (4.59)
\end{aligned}$$

Now, we would like to estimate $k = |K_{\mathbf{x},\mathbf{z}}|$. Notice that, by definition of $K_{\mathbf{x},\mathbf{z}}$ and thanks to the expansion properties, just like in the case $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$ we have that $|N(K_{\mathbf{x},\mathbf{z}})| \geq B|K_{\mathbf{x},\mathbf{z}}|$. We claim that

$$|K_{\mathbf{x},\mathbf{z}}| \leq \frac{|T_{\mathbf{x},\mathbf{z}}| + |J_{\mathbf{x},\mathbf{z}}|}{AB-1}$$

and we will show this taking into account the two (exhaustive) following cases:

- If $|N(K_{\mathbf{x},\mathbf{z}})| \leq n(1-R)/(A+1-R)$, then

  $$|J_{\mathbf{x},\mathbf{z}}| + |K_{\mathbf{x},\mathbf{z}}| + |T_{\mathbf{x},\mathbf{z}}| \geq |N(N(K_{\mathbf{x},\mathbf{z}}))| \geq A|N(K_{\mathbf{x},\mathbf{z}})| \geq AB|K_{\mathbf{x},\mathbf{z}}|,$$

  from which the claim follows directly.
- If $|N(K_{\mathbf{x},\mathbf{z}})| > n(1-R)/(A+1-R)$, then

  $$\begin{aligned}
  &|J_{\mathbf{x},\mathbf{z}}| + |K_{\mathbf{x},\mathbf{z}}| + |T_{\mathbf{x},\mathbf{z}}| \geq |N(N(K_{\mathbf{x},\mathbf{z}}))| \geq A\frac{n(1-R)}{A+1-R} \\
  &\geq A\frac{n}{B(1-R)+1} = An\left(1 - \frac{B(1-R)}{B(1-R)+1}\right) \geq A\left(n - |N(M_{\mathbf{x},\mathbf{z}})|\right) \\
  &\geq A|N(K_{\mathbf{x},\mathbf{z}})| \geq AB|K_{\mathbf{x},\mathbf{z}}|
  \end{aligned}$$

  and we obtain the same conclusion as before.

If we apply this estimation to (4.59), also recalling that $|J_{\mathbf{x},\mathbf{z}}|+|T_{\mathbf{x},\mathbf{z}}| \geq A|R_{\mathbf{x},\mathbf{z}}|$, we get

$$\begin{aligned}&|\{\mathbf{x},\mathbf{z} \in \mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho) \smallsetminus p\mathbb{Z}^n : |I_{\mathbf{x},\mathbf{z}}| = i, |J_{\mathbf{x},\mathbf{z}}| = j, |T_{\mathbf{x},\mathbf{z}}| = t\}| \\ &\quad \leq n^{t+j} p^{(t+j)/(AB-1)+1} p^{(t+j)/A} |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}|.\end{aligned}$$

We can now go back to the main estimation and, again, introduce the quantity $\mathcal{E}(\rho)$:

$$\begin{aligned}&\sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |I_{\mathbf{x},\mathbf{z}}|=i,|J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t \\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=1}} \left(\frac{1}{p}\right)^{2n(1-R)-i-2j} \\ &\quad = \sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \sum_{\substack{\mathbf{x},\mathbf{z}\in\mathbb{Z}^n\cap B_{\mathbf{0},n}(\rho)\smallsetminus p\mathbb{Z}^n \\ |J_{\mathbf{x},\mathbf{z}}|=j,|T_{\mathbf{x},\mathbf{z}}|=t \\ |\mathcal{M}_{\mathbf{x},\mathbf{z}}|=1}} \left(\frac{1}{p}\right)^{n(1-R)-2j+t+j} \\ &\quad = \sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} n^{t+j} p^{(t+j)/(AB-1)+1} p^{(t+j)/A} \frac{|\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho)|}{|\mathbb{Z}^n \cap B_{\mathbf{0},n}(\rho)|} \cdot \\ &\qquad\qquad \cdot p^{2j-t-j} \sqrt{\mathcal{E}(\rho)} \\ &\quad \lesssim \sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \frac{g(n)p}{\sqrt{\mathcal{E}(\rho)}} \left(\frac{p^2}{p^{(1-R)B}}\right)^j \left(\frac{np^{1/(AB-1)}p^{1/A}}{p}\right)^{t+j} \mathcal{E}(\rho),\end{aligned}$$

where we have applied Lemma 4.6 to obtain the asymptotical estimation and $g(n)$ is the analogue of $f(n)$:

$$g(n) = \left(\sqrt{\frac{n}{n-Bj}}\right)^{n-Bj+1} \left(1 + \frac{2\sqrt{n}}{2\rho - \sqrt{n}}\right)^n \left(1 + \frac{1}{n^\omega}\right)^{-Bj}.$$

Now, very similarly to what happens in the case $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$ (we omit some easy details), conditions

$$B > \frac{2}{1-R} \qquad \text{and} \qquad \lambda > 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1}$$

(cf. (4.36) and (4.40)) allow us to deduce that

$$\sum_{\substack{i,j,t \\ j\leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R)}} \frac{g(n)p}{\sqrt{\mathcal{E}(\rho)}} \left(\frac{p^2}{p^{(1-R)B}}\right)^j \left(\frac{np^{1/(AB-1)}p^{1/A}}{p}\right)^{t+j} \mathcal{E}(\rho) \lesssim o(1)\mathcal{E}(\rho). \tag{4.60}$$

Notice that $\mathcal{E}(\rho)$, is known to tend subexponentially to infinity when n grows and so does its square root.

Putting together (4.51), (4.58) and (4.60), we obtain that

$$\mathrm{Var}(X_\rho) \lesssim o(1)\mathcal{E}(\rho).$$

By the means of the Chebyshev's inequality (cf. Lemma 2.1) and taking into account the fact that $\mathcal{E}(\rho) \leq \mathbb{E}[X_\rho]^2$, we finally arrive to the end of the proof:

$$\begin{aligned}\mathcal{P}\{X_\rho = 0\} &\leq \mathcal{P}\{|X_\rho - \mathbb{E}[X_\rho]| \geq \mathbb{E}[X_\rho]\} \\ &\leq \frac{\mathrm{Var}(X_\rho)}{\mathbb{E}[X_\rho]^2} \\ &\lesssim \frac{o(1)\mathcal{E}(\rho)}{\mathbb{E}[X_\rho]^2} \\ &\lesssim \frac{o(1)\mathbb{E}[X_\rho]^2}{\mathbb{E}[X_\rho]^2} \longrightarrow 0,\end{aligned}$$

that is,

$$\lim_{n\to\infty} \mathcal{P}\left\{X_{\rho_{\mathrm{eff}}\left(1+\frac{1}{n^\omega}\right)} = 0\right\} = 0.$$

□

The proof that capacity is achieved with LDA lattices

Now that we have proved that in the case of LDA Voronoi constellations the sent point has the same typical norm of the constellation points of the more general Construction A, we are ready to prove the result that LDA lattices can achieve the capacity of the AWGN channel under lattice decoding. As we have already said, the transmission scheme is the same of Section 4.2.2 and the proof of the theorem is then very similar to the one of Theorem 4.1. Nevertheless, we will have to adapt it to the LDPC structure that gives rise to LDA lattices, just like we had to adapt the proof of the previous lemma.

Theorem 4.2. *The random ensemble of nested LDA lattices that we have described till now in this section achieves the capacity of the AWGN channel under MMSE lattice decoding, when* $\mathrm{SNR} > 1$, $R > 1/2$ *and* $p = n^\lambda$ *for some constant* λ *such that*

$$\lambda > \max\left(\frac{1}{1-R_f}, \frac{2}{A-2}, \frac{2}{B(1-R_f)-2}, \frac{1}{2R_f}, 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1}\right) \quad (4.61)$$

and such that condition (4.40) *is satisfied, too.*

Remark: the proof of this theorem strongly relies on the techniques that we have already applied in the proofs of Theorem 4.1 and Lemma 4.7. For this reason, we will skip some detail and some technical computation. Everything which is not completely developed is a straightforward modification of some well-referenced computation that was already done in another context. We strongly recommend to get familiar with the arguments used in the demonstrations of Theorem 4.1 and Lemma 4.7 before reading the sequel in depth.

Proof. The geometric and probabilistic strategy to prove this theorem is the same that we have applied to prove Theorem 4.1. Namely, the beginnings of the two proofs are identical and almost everything coincides; the small differences can be easily solved by a slight adaptation of what is done in the proof of Theorem 4.1. For this reason, we claim that the only thing that we need to prove is that

$$\lim_{n\to\infty} \left(\sum_{\mathbf{x}\in S} \mathcal{P}\{\mathbf{H}'\mathbf{x}^T \equiv \mathbf{m}^T \bmod p\} \right. \tag{4.62}$$

$$\left. \cdot \sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\ \mu=0,1,2}} \mathcal{P}\{\mathbf{H}_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, \mathbf{H}_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{\mathbf{z}\in\mathcal{B}\setminus\{\mathbf{x}\}\} \right) = 0. \tag{4.63}$$

This formula is the equivalent of (4.28) and (4.29). For the notation, we recall that:

- $\mathcal{B}_{\text{eff}}$ is the n-dimensional ball centred at $\mathbf{0}$ with radius

$$\rho_{\text{eff}}\left(1+\frac{1}{n^\omega}\right) = \frac{\sqrt{n}p^{(1-R)}}{\sqrt{2\pi e}}\left(1+\frac{1}{n^\omega}\right),$$

  where ω is the same constant of Lemma 4.7.

- $\mathbf{m}$ is the upper part of the syndrome $\mathbf{s}$ (cf. Figure 4.2 at the beginning of Section 4.2).

- $\mathcal{B}'$ is the n-dimensional ball centred at $\mathbf{x}$, with radius equal to

$$2\rho_{\text{dec}} = 2\sqrt{\alpha n}\sigma(1+\xi),$$

  where α is the Wiener coefficient used for MMSE scaling, σ is the noise variance per dimension and ξ is a fixed small constant (actually, in the proof of Theorem 4.1, we had ε instead of ξ; here we change the notation to avoid any misunderstanding with the ε involved in the definition of an (α, A, β, B)-good graph).

- $\sigma = \sigma_{\max}(1-\delta)$, for some $0 < \delta < 1$ and $\sigma_{\max}$ is defined like in (4.22).

- $\mathcal{B}$ is the *decoding sphere*, centred at $\alpha\mathbf{y}$ (the MMSE-scaled channel output) with radius ρ_{dec}.

- S is defined as the set of the $\mathbf{x}$'s in $(\mathcal{B}_{\text{eff}}\cap\mathbb{Z}^n)\setminus p\mathbb{Z}^n$ for which for all $\mu\in\{0,1,2\}$, all $0 \le \nu < 1$, and all $\mathbf{z}$ such that $|\{i\in\{1,2,\ldots,n\} : x_i \equiv \mu z_i \bmod p\}| \ge n - n^\nu$, the point $\mathbf{z}$ itself does not induce a decoding error ($\mathbf{x}$ being the channel input). Showing that the previous sum can be restricted to the $\mathbf{x}$'s in S slightly generalises what is done to define the S of the proof of Theorem 4.1; it also requires similar techniques to the ones that will be employed later on to conclude this proof.

First of all, let us deduce something about the non-zero subsyndrome $\mathbf{m}$: how many are the $\mathbf{m} \in \mathbb{F}_p^{n(R_f-R)}$ such that $m_i \neq 0$ for every i? We have:

$$\begin{aligned}|\{\mathbf{m} \in \mathbb{F}_p^{n(R_f-R)} : m_i \neq 0, \forall i\}| &= (p-1)^{n(R_f-R)} \\ &= \left(1-\frac{1}{p}\right)^{n(R_f-R)} p^{n(R_f-R)} \\ &= \left(1-\frac{1}{n^\lambda}\right)^{n(R_f-R)} p^{n(R_f-R)} \to p^{n(R_f-R)},\end{aligned}$$

because $\lambda > (1-R_f)^{-1} > 1$ (recall hypothesis (4.61)). This means that the proportion of $\mathbf{m}$'s that contain some 0 coordinates is vanishing with respect to the total number of subsyndromes. For this reason, the contribution to the average error probability of this messages is vanishing and we only need to show (4.62) for the $\mathbf{m}$'s such that $m_i \neq 0$ for every i. From now on, we make this hypothesis, which implies that

$$\mathcal{P}\{\mathbf{H}'\mathbf{x}^T \equiv \mathbf{m}^T \bmod p\} \leq \left(\frac{1}{p}\right)^{n(R_f-R)},$$

since the intersection of the supports of $\mathbf{x}$ and any row of $\mathbf{H}'$ is never empty. Note that, if the inequality is strict, then the probability is 0.

Now, we would like to express the other probabilities of (4.63) that $\mathbf{x}$ and $\mathbf{z}$ have a certain subsyndrome in the same form as in the proof of Lemma 4.7. For this purpose, given a fixed $\mathbf{x}$ and a fixed $\mathbf{z}$, let

$$\begin{aligned}J^f_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \text{ row of } \mathbf{H}_f : \dim(\mathbf{x},\mathbf{z},\mathbf{h}) = 0\}, \\ I^f_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \text{ row of } \mathbf{H}_f : \dim(\mathbf{x},\mathbf{z},\mathbf{h}) = 1\}, \\ T^f_{\mathbf{x},\mathbf{z}} &= \{\mathbf{h} \text{ row of } \mathbf{H}_f : \dim(\mathbf{x},\mathbf{z},\mathbf{h}) = 2\},\end{aligned}$$

where the definition of $\dim(\mathbf{x},\mathbf{z},\mathbf{h})$ is the same that we have given in the proof of Lemma 4.7. Hence, if we define

$$Z = \{\mathbf{z} \in \mathcal{B}' \cap \mathbb{Z}^n : \mathbf{z} \not\equiv \mu\mathbf{x} \bmod p, \forall \mu \in \{0,1,2\}\},$$

we obtain

$$\begin{aligned}&\sum_{\mathbf{x}\in S} \mathcal{P}\{\mathbf{H}'\mathbf{x}^T \equiv \mathbf{m}^T \bmod p\} \cdot \\ &\qquad \cdot \sum_{\substack{\mathbf{z}\in\mathcal{B}'\cap\mathbb{Z}^n \\ \mathbf{z}\not\equiv\mu\mathbf{x},\mu=0,1,2}} \mathcal{P}\{\mathbf{H}_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, \mathbf{H}_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}\} \\ \leq \sum_{\substack{i,j,t \\ i+j+t=n(1-R_f)}} &\sum_{\substack{\mathbf{x}\in S \\ \mathbf{z}\in Z \\ |I^f_{\mathbf{x},\mathbf{z}}|=i,|J^f_{\mathbf{x},\mathbf{z}}|=j,|T^f_{\mathbf{x},\mathbf{z}}|=t}} \left(\frac{1}{p}\right)^{n(R_f-R)} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j} \mathcal{P}\{\mathbf{z} \in \mathcal{B} \smallsetminus \{\mathbf{x}\}\}.\end{aligned}$$

Now, observe that if $\mathcal{S}$ is the set of all the balls of the space of radius ρ_{dec} (the same radius as $\mathcal{B}$), then

$$\begin{aligned}
&\sum_{\substack{\mathbf{x}\in S\\ \mathbf{z}\in Z\\ |I^f_{\mathbf{x},\mathbf{z}}|=i,|J^f_{\mathbf{x},\mathbf{z}}|=j,|T^f_{\mathbf{x},\mathbf{z}}|=t}} \mathcal{P}\{\mathbf{z}\in\mathcal{B}\setminus\{\mathbf{x}\}\} \\
&\quad= \sum_{\substack{\mathbf{x}\in S\\ \mathbf{z}\in Z\\ |I^f_{\mathbf{x},\mathbf{z}}|=i,|J^f_{\mathbf{x},\mathbf{z}}|=j,|T^f_{\mathbf{x},\mathbf{z}}|=t}} \sum_{\substack{B\in\mathcal{S}\\ \mathbf{z}\in B}} \mathcal{P}\{\mathcal{B}=B\} \\
&\quad= \sum_{B\in\mathcal{S}} \sum_{\substack{\mathbf{x}\in S\\ \mathbf{z}\in B\cap Z\\ |I^f_{\mathbf{x},\mathbf{z}}|=i,|J^f_{\mathbf{x},\mathbf{z}}|=j,|T^f_{\mathbf{x},\mathbf{z}}|=t}} \mathcal{P}\{\mathcal{B}=B\}. \qquad (4.64)
\end{aligned}$$

Then, for a fixed $\mathbf{x}$, if we call

$$Z_{ijt} = \max_{B\in\mathcal{S}} |\{\mathbf{z}\in B\cap Z\} : |I^f_{\mathbf{x},\mathbf{z}}|=i, |J^f_{\mathbf{x},\mathbf{z}}|=j, |T^f_{\mathbf{x},\mathbf{z}}|=t\}|,$$

we can go on from (4.64) and write

$$\begin{aligned}
(4.64) &\le \sum_{\mathbf{x}\in S} Z_{ijt} \sum_{B\in\mathcal{S}} \mathcal{P}\{\mathcal{B}=B\} \\
&= \sum_{\mathbf{x}\in S} Z_{ijt}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\sum_{\mathbf{x}\in S} \mathcal{P}\{\mathbf{H}'\mathbf{x}^T \equiv \mathbf{m}^T \bmod p\}\cdot \\
&\qquad\qquad \cdot \sum_{\mathbf{z}\in Z} \mathcal{P}\{\mathbf{H}_f\mathbf{x}^T \equiv \mathbf{0}^T \bmod p, \mathbf{H}_f\mathbf{z}^T \equiv \mathbf{0}^T \bmod p\}\mathcal{P}\{\mathbf{z}\in\mathcal{B}\setminus\{\mathbf{x}\}\} \\
&\le \sum_{\substack{i,j,t\\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j} \\
&= \sum_{\substack{i,j,t\\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(1-R)} Z_{ijt} \left(\frac{1}{p}\right)^{n(1-R_f)-i-2j}.
\end{aligned}$$

Now, exactly like in the variance computation in the proof of Lemma 4.7, we will split this sum into two parts, depending on the values of j. Namely, we will study the convergence of the sum for:

1. $j > n(1-R_f)/(B(1-R_f)+1)$.
2. $j \le n(1-R_f)/(B(1-R_f)+1)$.

First case: $j > n(1-R_f)/(B(1-R_f)+1)$**.** This case develops in the same manner as the "first case" of the proof of Lemma 4.7: we have

$$\begin{aligned}
&\sum_{\substack{i,j,t \\ j>n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j} \\
&\qquad \leq \sum_{\substack{i,j,t \\ j>n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j} \\
&\qquad \leq \sum_{\substack{i,j,t \\ j>n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} Z_{ijt} \left(\frac{1}{p}\right)^{2t+i}. \qquad (4.65)
\end{aligned}$$

The very same argument used in the proof of Lemma 4.7 implies also that this sum is bounded by the same bound as (4.49) (with R_f instead of R, of course). Notice that in this particular case no problem comes from the fact that the radii of the balls in the previous sum are different from the radii of the balls in (4.49); this is because all of them are asymptotically smaller than p (thanks to condition (4.61): $\lambda > (1-R_f)^{-1} > (1-R)^{-1}$)) and this somehow rough condition is the only one that we need for our estimation. In conclusion,

$$\lim_{n\to\infty} \sum_{\substack{i,j,t \\ j>n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j} = 0,$$

thanks to the hypothesis $\lambda > 2(A-2)^{-1}$ (cf. (4.61) and the application of the condition in (4.51)).

Second case: $j \leq n(1-R_f)/(B(1-R_f)+1)$**.** Once again, we take inspiration from the proof of Lemma 4.7. There, the "second case" was splitted into two more cases. We will do the same here and bound in two different ways the sum on the $\mathbf{x}$'s and $\mathbf{z}$'s, depending on the fact that $\mathcal{M}^f_{\mathbf{x},\mathbf{z}}$ is 0 or 1 (for the definition of $\mathcal{M}^f_{\mathbf{x},\mathbf{z}}$, see the corresponding definition of $\mathcal{M}_{\mathbf{x},\mathbf{z}}$ in the proof of Lemma 4.7 (cf. (4.52)).

1. If $|\mathcal{M}^f_{\mathbf{x},\mathbf{z}}| = 0$, a trivial adaptation of the argument used in the proof of Lemma

4.7 for the corresponding case $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| = 0$ says that

$$\sum_{\substack{i,j,t \\ j\le n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j}$$

$$\le \sum_{\substack{i,j,t \\ j\le n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} n^{i+j}p^i |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho_{\text{eff}}(1+1/n^\omega))| \cdot$$

$$\cdot |\mathbb{Z}^{n-Bi} \cap B_{\mathbf{0},n-Bi}(\rho_{\text{dec}})| \left(\frac{1}{p}\right)^{n(R_f-R)} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j}.$$

Let us define

$$\mathcal{Q}(\rho_{\text{eff}}, \rho_{\text{dec}}) = |\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}| \left(\frac{1}{p}\right)^{n(1-R)} |\mathbb{Z}^n \cap \mathcal{B}| \left(\frac{1}{p}\right)^{n(1-R_f)}.$$

Then

$$\sum_{\substack{i,j,t \\ j\le n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j}$$

$$\le \sum_{\substack{i,j,t \\ j\le n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \frac{|\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho_{\text{eff}}(1+1/n^\omega))||\mathbb{Z}^{n-Bi} \cap B_{\mathbf{0},n-Bi}(\rho_{\text{dec}})|}{|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}||\mathbb{Z}^n \cap \mathcal{B}|} \cdot$$

$$\cdot n^{i+j}p^{2i+2j}\mathcal{Q}(\rho_{\text{eff}}, \rho_{\text{dec}})$$

$$\lesssim \sum_{\substack{i,j,t \\ j\le n(1-R_f)/(B(1-R_f)+1) \\ i+j+t=n(1-R_f)}} \left(\sqrt{\frac{n}{n-Bj}}\right)^{n-Bj+1} \left(\sqrt{\frac{n}{n-Bi}}\right)^{n-Bi+1} \cdot$$

$$\cdot \frac{n^j p^{2j}}{p^{Bj(1-R)}} \frac{n^i p^{2i}}{p^{Bi(1-R_f)}} (1+\xi)(1-\delta)^{-Bi}\mathcal{Q}(\rho_{\text{eff}}, \rho_{\text{dec}})$$

$$= o(1)\mathcal{Q}(\rho_{\text{eff}}, \rho_{\text{dec}}),$$

because condition (4.61) states

$$\lambda > \max\left(\frac{2}{B(1-R)-2}, \frac{2}{B(1-R_f)-2}\right) = \frac{2}{B(1-R_f)-2}.$$

Now, notice that we have already shown in the proof of Theorem 4.1 that

$$\lim_{n\to\infty} \mathcal{Q}(\rho_{\text{eff}}, \rho_{\text{dec}}) = 0;$$

indeed, it is bounded from above by (4.33), which was shown to be vanishing when n tends to infinity. This concludes the analysis of the case $|\mathcal{M}^f_{\mathbf{x},\mathbf{z}}| = 0$.

2. Now, let $|\mathcal{M}^f_{\mathbf{x},\mathbf{z}}| = 1$ and suppose for now that $t < n^\nu$ for some $\nu < 1$. Consider the set of parity-check equation nodes of the Tanner graph associated with $\mathbf{H}_f$ given by $I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}}$ and the bipartite subgraph $\mathcal{H}_{\mathbf{x},\mathbf{z}}$ that it induces, whose set of parity-check equation nodes is $I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}}$, whose set of variable nodes is $N(I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}})$ and whose edges are all the edges of the original Tanner graph beween these two sets. A priori this graph is not connected; if we denote $\mathcal{C}$ one of its connected component and $P_\mathcal{C}$ its set of parity-check equation nodes, we can partition $\mathcal{H}_{\mathbf{x},\mathbf{z}}$ into the disjoint union of the two following graphs:

$$\mathcal{L}^f_{\mathbf{x},\mathbf{z}} = \{\mathcal{C} \subseteq \mathcal{H}_{\mathbf{x},\mathbf{z}} : |P_\mathcal{C}| \leq n(1-R_f)/(B(1-R_f)+1)\} \text{ and}$$
$$\mathcal{D}^f_{\mathbf{x},\mathbf{z}} = \{\mathcal{C} \subseteq \mathcal{H}_{\mathbf{x},\mathbf{z}} : |P_\mathcal{C}| > n(1-R_f)/(B(1-R_f)+1)\}.$$

As a consequence, $I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}}$ is the disjoint union of

$$L^f_{\mathbf{x},\mathbf{z}} = \bigcup \{P_\mathcal{C} : \mathcal{C} \in \mathcal{L}_{\mathbf{x},\mathbf{z}}\} \qquad \text{and} \qquad D^f_{\mathbf{x},\mathbf{z}} = \bigcup \{P_\mathcal{C} : \mathcal{C} \in \mathcal{D}_{\mathbf{x},\mathbf{z}}\}.$$

The first observation that we can make is that since $j \leq n(1-R_f)/(B(1-R_f)+1)$ and $t < n^\nu$, then $|\mathcal{D}_{\mathbf{x},\mathbf{z}}| = 1$. Indeed, $|\mathcal{D}_{\mathbf{x},\mathbf{z}}| \leq 1$ for the same reason for which $|\mathcal{M}_{\mathbf{x},\mathbf{z}}| \leq 1$ in the proof of Lemma 4.7 (see what follows equation (4.53)); moreover, $\mathcal{D}_{\mathbf{x},\mathbf{z}} \neq \emptyset$ because otherwise $L^f_{\mathbf{x},\mathbf{z}} = I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}}$ and these two conditions would hold (at least asymptotically):

(a) $L^f_{\mathbf{x},\mathbf{z}}$ has size $n(1-R_f) - |T^f_{\mathbf{x},\mathbf{z}}| \geq n - n^\nu$.

(b) $|N(L^f_{\mathbf{x},\mathbf{z}})| \geq B|L^f_{\mathbf{x},\mathbf{z}}| \geq B(n-n^\nu) > n$.

The second one is clearly a nonsense and proves that $|\mathcal{D}_{\mathbf{x},\mathbf{z}}| = 1$.

We go on with this analysis and we claim that $L^f_{\mathbf{x},\mathbf{z}}$ cannot be too big: $|L^f_{\mathbf{x},\mathbf{z}}| \leq t/(AB-1)$. Indeed, the expansion properties imply that $|N(L^f_{\mathbf{x},\mathbf{z}})| \geq B|L^f_{\mathbf{x},\mathbf{z}}|$. So,

- if $|N(L^f_{\mathbf{x},\mathbf{z}})| \leq n/(A+1-R_f)$, then

$$|L^f_{\mathbf{x},\mathbf{z}}| + |T^f_{\mathbf{x},\mathbf{z}}| \geq |N(N(L^f_{\mathbf{x},\mathbf{z}}))| \geq A|N(L^f_{\mathbf{x},\mathbf{z}})| \geq AB|L^f_{\mathbf{x},\mathbf{z}}|,$$

  from which the claim follows directly;

- if instead $|N(L^f_{\mathbf{x},\mathbf{z}})| \leq n/(A+1-R_f)$, we can observe that

$$\begin{aligned}
&|L^f_{\mathbf{x},\mathbf{z}}| + |T^f_{\mathbf{x},\mathbf{z}}| \geq |N(N(L^f_{\mathbf{x},\mathbf{z}}))| \geq A\frac{n(1-R_f)}{A+1-R_f} \\
&\geq A\frac{n}{B(1-R_f)+1} = An\left(1 - \frac{B(1-R_f)}{B(1-R_f)+1}\right) \geq A\left(n - |N(D^f_{\mathbf{x},\mathbf{z}})|\right) \\
&\geq A|N(L^f_{\mathbf{x},\mathbf{z}})| \geq AB|L^f_{\mathbf{x},\mathbf{z}}|,
\end{aligned}$$

  that allows us to deduce the claim once again.

Substantially, we have just proved that when $j \leq n(1-R_f)/(B(1-R_f)+1)$ and $t < n^\nu$, then the parity-check equations associated with the big connected component of $I^f_{\mathbf{x},\mathbf{z}} \cup J^f_{\mathbf{x},\mathbf{z}}$ are almost all the equations of the matrix H_f; the size of what is left (the set $|T^f_{\mathbf{x},\mathbf{z}}|$ plus the equations of the "small" connected components) is $O(n^\nu)$.

The next thing we would like to prove is that the number of $\mathbf{x} \in \mathcal{B}_{\text{eff}}$ such that there exists some $\mathbf{z} \in \mathcal{B}'$ satisfying the hypothesis on $|T^f_{\mathbf{x},\mathbf{z}}|$ and $|J^f_{\mathbf{x},\mathbf{z}}|$ is very small with respect to the total number of $\mathbf{x}$. Hence, let

$$\mathcal{N} = \left| \left\{ \mathbf{x} \in \mathcal{B}_{\text{eff}} : \exists \mathbf{z} \in \mathcal{B}' \text{ such that } |T^f_{\mathbf{x},\mathbf{z}}| < n^\nu \text{ and } |J^f_{\mathbf{x},\mathbf{z}}| \leq \frac{n(1-R_f)}{B(1-R_f)+1} \right\} \right|.$$

Let us start observing that $\mathbf{x}$ and $\mathbf{z}$ have to be multiple modulo p on all the coordinates of $N(D^f_{\mathbf{x},\mathbf{z}})$. Indeed, this holds by definition of $I^f_{\mathbf{x},\mathbf{z}}$ on the coordinates of $N(I^f_{\mathbf{x},\mathbf{z}} \cap D^f_{\mathbf{x},\mathbf{z}})$ and by the fact that they are fixed to 0 modulo p for both $\mathbf{x}$ and $\mathbf{z}$ on the coordinates of $N(J^f_{\mathbf{x},\mathbf{z}} \cap D^f_{\mathbf{x},\mathbf{z}})$. In other terms, there exists $\mu \in \{3, 4, \ldots, p-1\}$ - recall that the values $0, 1$ and 2 are excluded by the definition of S and Z - such that

$$\begin{aligned} |\{l \in \{1,2,\ldots,n\} : x_l \equiv \mu z_l \bmod p\}| &\geq n - |N(L^f_{\mathbf{x},\mathbf{z}} \cup T^f_{\mathbf{x},\mathbf{z}})| \\ &\geq n - B|L^f_{\mathbf{x},\mathbf{z}} \cup T^f_{\mathbf{x},\mathbf{z}}| \\ &\geq n - |T^f_{\mathbf{x},\mathbf{z}}| \left(1 + \frac{1}{AB-1}\right) \\ &\geq n - 2Bn^\nu. \end{aligned}$$

Let us define

$$\mathcal{N}' = |\{\mathbf{x} \in B_{\mathbf{0},\lceil n-2Bn^\nu \rceil}(\rho_{\text{eff}}(1+1/n^\omega)) : \exists \mathbf{z} \in B_{\mathbf{x},\lceil n-2Bn^\nu \rceil}(2\rho_{\text{dec}}), \mathbf{x} \equiv \mu \mathbf{z} \bmod p\}|;$$

the previous estimation directly implies that

$$\mathcal{N} \leq \binom{n}{\lfloor 2Bn^\nu \rfloor} |\mathbb{Z}^{\lfloor 2Bn^\nu \rfloor} \cap B_{\mathbf{0},\lfloor 2Bn^\nu \rfloor}(\rho_{\text{eff}}(1+1/n^\omega))| \mathcal{N}' \leq n^{2Bn^\nu} p^{2Bn^\nu} \mathcal{N}'.$$

Now, let us consider the following summation, in which t is fixed to be smaller than n^ν:

$$\sum_{\substack{i,j,t \\ j \leq n(1-R_f)/(B(1-R_f)+1) \\ t < n^\nu \\ i+j+t = n(1-R_f)}} \sum_{\mathbf{x} \in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j}. \tag{4.66}$$

$Z_{ijt} = 0$ except for a subset of points $\mathbf{x}$ of size $\mathcal{N}$. Moreover, once an $\mathbf{x}$ is fixed, the $\mathbf{z}$(s that guarantee the configuration of i, j and t are identified by:

(a) the choice of $O(n^\nu)$ coordinates.

(b) the choice of at most p value for each of those coordinates.

Then,

$$\begin{aligned}(4.66) &\leq \sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t<n^\nu \\ i+j+t=n(1-R_f)}} \mathcal{N}(np)^{O(n^\nu)} \left(\frac{1}{p}\right)^{n(R_f-R)+2n(1-R_f)-i-2j} \\ &\leq |\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}| \left(\frac{1}{p}\right)^{n(1-R)} \sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t<n^\nu \\ i+j+t=n(1-R_f)}} \frac{(np)^{O(n^\nu)}\mathcal{N}'}{|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}|} \left(\frac{1}{p}\right)^{t-j} \\ &\leq \left(|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}| \left(\frac{1}{p}\right)^{n(1-R)}\right) \frac{n^2 n^{O(n^\nu)}\mathcal{N}'}{|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}|}.\end{aligned}$$

Now, the term in the parentheses is known to grow subexponentially to infinity with n (see for example the comments concerning (4.33)). On the other hand, Lemma 4.5 provides an upper bound for $\mathcal{N}'$ and it is possible to show that the ratio on the right decreases exponentially to 0 (it is the analogue computation to the one done at the end of the proof of Theorem 4.1). Hence, the latter is the dominating term and whole quantity is vanishing when n tends to infinity.

We are left to study the sum corresponding to $t > n^\nu$:

$$\sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t>n^\nu \\ i+j+t=n(1-R_f)}} \sum_{\mathbf{x}\in S} \left(\frac{1}{p}\right)^{n(R_f-R)} Z_{ijt} \left(\frac{1}{p}\right)^{2n(1-R_f)-i-2j}. \tag{4.67}$$

For this estimation, we rely once again on the similar computations already

done in the proof of Lemma 4.7: based on that example, we can easily find:

$$
\begin{aligned}
(4.67) \leq & \sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t>n^\nu \\ i+j+t=n(1-R_f)}} \left(np^{1/(AB-1)+1/A}\right)^{t+j} p \left(\frac{1}{p}\right)^{n(1-R)+n(1-R_f)-i-2j} \cdot \\
& \cdot |\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho_{\text{eff}}(1+1/n^\omega))| \\
\leq & |\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}| \left(\frac{1}{p}\right)^{n(1-R)} \sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t>n^\nu \\ i+j+t=n(1-R_f)}} \left(\frac{np^{1/(AB-1)+1/A}}{p}\right)^{t+j} pp^{2j} \cdot \\
& \cdot \frac{|\mathbb{Z}^{n-Bj} \cap B_{\mathbf{0},n-Bj}(\rho_{\text{eff}}(1+1/n^\omega))|}{|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}|} \\
\lesssim & \left(|\mathbb{Z}^n \cap \mathcal{B}_{\text{eff}}| \left(\frac{1}{p}\right)^{n(1-R)}\right) \sum_{\substack{i,j,t \\ j\leq n(1-R_f)/(B(1-R_f)+1) \\ t>n^\nu \\ i+j+t=n(1-R_f)}} \left(\sqrt{\frac{n}{n-Bj}}\right)^{n-Bj+1} \cdot \\
& \cdot p \left(\frac{np^{1/(AB-1)+1/A}}{p}\right)^{t+j} \left(\frac{p^2}{p^{(1-R_f)B}}\right)^{j}.
\end{aligned}
$$

Now, the term outside the sum goes subexponentially to infinity when n tends to infinity, as we pointed out before; on the other hand, the sum converges to 0 thanks to conditions (4.36) and (4.61):

$$
B > \frac{2}{1-R_f} \qquad \text{and} \qquad \lambda > 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1};
$$

taking ν big enough, it also dominates the left term. So, the whole quantity tends to 0 when n tends to infinity.

Conclusion. Putting together the "first case" and the two distinguished parts of the "second case", we conclude that the limit in (4.62) holds true and this ends the proof of the theorem. □

Chapter 5

Applications and numerical experiments

This dissertation has dealt so far with information-theoretical results about the capacity-achieving properties of lattices and most of them concerned LDA lattices. This analysis integrates the work started by de Buda in 1975 [dB75] and carried out successfully till nowadays by the international community, as summarised in Section 2.3.2. Though, these results are non-contructive and from a practical point of view not many lattice coding schemes have been proposed that have a chance of approaching capacity. For large dimensions, the promising lattices are:

- Low-Density Parity-Check (LDPC) Lattices, based on an underlying binary code structure: the binary code is chosen to be amenable to an iterative decoding algorithm and belongs to the LDPC families [SBP06, BC08, SS13];
- Low-Density Lattice Codes (LDLC), that are constructed directly so as to be decodable by a scheme inspired by the LDPC techniques [SFS08];
- Turbo Lattices, based on binary, iteratively decodable turbo codes [SSP12];
- Polar Lattices, that, as the name suggests, take inspiration from polar codes [YLW13, YL12].

As it may appear natural at this point, we ask ourselves the questions: can we find an iterative decoding scheme for LDA lattices? Is their sparse parity-check matrix underlying structure suitable for practical algorithmic implementation? Of course, the answer is yes. Moreover, this is exactly the reason why we have turned our attention to this kind of LDPC-based Construction A family. From a chronological point of view, the experimental results of [dPBZB12] come even before and motivate the theoretical achievements of [dPBZB13, dPBZ13] and Chapter 3 and Chapter 4. In the sequel of the chapter, we describe and discuss our LDA-adapted decoder of infinite constellations, giving some specifics and some comments on the choice of parameters for the LDA family we experiment with; we also discuss some simulation results, compared to the ones of the other lattice families provided in the literature.

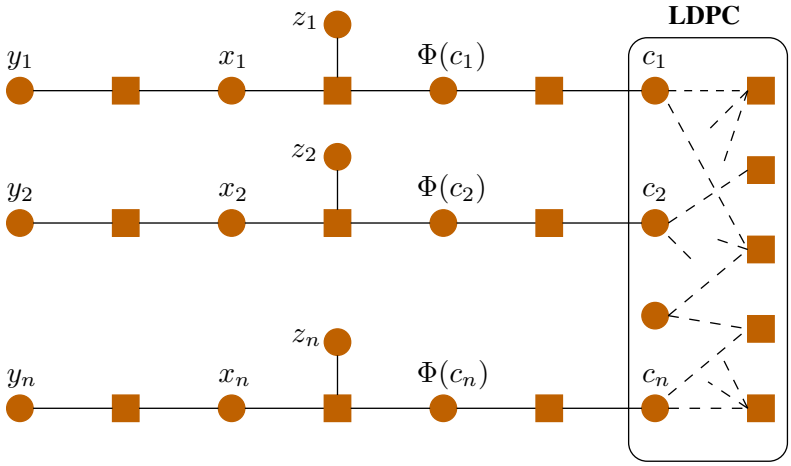

Figure 5.1: Factor graph of an LDA lattice.

5.1 Iterative decoding of infinite LDA constellations

In small dimensions, typically less than 100, optimal decoding algorithms for lattice constellations are manageable, such as Sphere Decoding [VB99][BB03]. For higher dimensions ($n \geq 1000$), there is no method to handle decoding of lattices besides iterative message passing algorithms. For this reason, we propose an iterative decoding scheme for LDA lattices. Even if it is suboptimal with repsect to ML decoding and we are not able to study its performance from a purely theoretical point of view, it still reveals to be an interesting algorithm, thanks to its complexity and the satisfactory numerical simulation results.

The complexity of iterative message passing is *linear in* n. In our setting, since we work with non-binary LDPC codes underlying Construction A, a critical parameter is also the size p of the finite field containing the linear code. Indeed, the p-ary LDPC code C defining an LDA lattice Λ can be decoded via *Belief Propagation* (BP) or *Min-Sum Decoding* [RU08]. The results that we show in Section 5.2 are obtained with BP. Decoding of an LDPC checknode in C is made via the *Forward-Backward algorithm* on the syndrome trellis [BCJR74], that has p^2 transitions in its largest section. For large p, checknode decoding should be done via Fast Fourier Transform [HR76] to make it faster. We describe below the factor graph of Λ and the messages propagating on its edges.

5.1.1 Factor graph for LDA lattices

The *factor graph* [KFL01] is derived from the lattice structure given in Section 2.2.1 (namely, see (2.1)):

$$\Lambda = \Phi(C) + p\mathbb{Z}^n,$$

where Φ is the embedding of C into $\mathbb{Z}^n$. Messages and constraints are given for a Construction A lattice over $\mathbb{Z}$ transmitted over a memoryless AWGN channel, with noise variance per dimension equal to σ^2. The channel input is denoted by

$\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and the output is $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. It is straightforward to extend them to Construction A over the Gaussian integers $\mathbb{Z}[i]$ (through the formula: $\Lambda = \Phi(C) + \mathbb{Z}[i]^n$) and other types of memoryless channels. As shown in Figure 5.1, the constraints are:

- The channel, where the output conditional distribution is $y_i \sim \mathcal{N}(x_i, \sigma^2)$, $i = 1, \ldots, n$.
- The lattice constraint given by Construction A, i.e. by the fact that Λ is the union of cosets of $p\mathbb{Z}^n$. We have $x_i = \Phi(c_i) + pz_i$, where $z_i \in \mathbb{Z}$, $c_i \in \mathbb{F}_p$, and $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in C$.
- The embedding Φ of $\mathbb{F}_p$ into the Euclidean space. For $(L, L') = (\mathbb{Z}, p\mathbb{Z})$ and $\ell = n$ (the notation comes from Definition 2.13), the isomorphism $\Phi(c_i)$ is simply defined as the element of $\{-(p-1)/2, -(p-3)/2, \ldots, (p-1)/2\}$ that projects onto c_i modulo p (of course, with $p \neq 2$). We will write c_i instead of $\Phi(c_i)$ in order to simplify the notation.
- The LDPC constraint given by $H\mathbf{c}^T \equiv \mathbf{0} \bmod p$, where H is the (sparse) parity-check matrix of the LDPC code C.

5.1.2 Probabilistic messages for Construction A

Now, let us find the expressions of messages propagating from left to right in the factor graph. The left-to-right message produced by x_i is

$$p(x_i|y_i) \propto \exp\left(-\frac{(y_i - x_i)^2}{2\sigma^2}\right), \quad \forall x_i \in \mathbb{Z}, \tag{5.1}$$

the symbol $\propto$ meaning "proportional to". Since we have $x_i = c_i + pz_i \equiv c_i \mod p$, the left-to-right message received by c_i is

$$p(c_i|y_i) = \sum_{\substack{x_i \in \mathbb{Z} \\ x_i \equiv c_i}} p(x_i|y_i). \tag{5.2}$$

Let us move to describing the messages propagating right to left. The right-to-left message produced by c_i is the LDPC *extrinsic information* $p(c_i|C, \mathbf{y} \setminus \{y_i\})$ determined by multiplying all messages from its neighbouring checknodes [RU08]. The outgoing message from z_i is 1 in the absence of *a priori* information. As shown later for our practical implementation, there is a hidden constraint producing an *a priori* information $\pi(z_i)$. Thus, the right-to-left message received by x_i would be

$$p(x_i|C, \mathbf{y} \setminus \{y_i\}) \propto \pi(z_i) \cdot p(c_i|C, \mathbf{y} \setminus \{y_i\})$$

From the above description and the fact that the *a posteriori* probability (APP) of a variable node v is determined by the product of the two messages in the two opposite directions on any edge connected to v (belief propagation on an a cyclic graph [RU08]), we can state the following lemma.

Lemma 5.1. *Let $\Lambda = C[n,k]_p + p\mathbb{Z}^n$ be an LDA lattice and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a lattice point. A message passing decoder should maximise $APP(x_i)$, for $i = 1, \ldots, n$, where the* a posteriori *probability for a lattice component is given by*

$$APP(x_i) \propto p(x_i|y_i) \cdot \pi(z_i) \cdot p(c_i|C, \mathbf{y} \setminus \{y_i\}). \tag{5.3}$$

5.1.3 Implementation

The summation over $\mathbb{Z}$ in (5.2) decays very quickly around y_i because of the exponential behaviour given in (5.1). Consider the real interval $\mathcal{W}_i = [y_i - m\sigma, y_i + m\sigma]$ where σ^2 is the AWG noise variance per dimension and $m \in \mathbb{R}^+$. We choose m such that the probability of the transmitted x_i being outside $\mathcal{W}_i$ is less than ε, i.e. $2Q(m) < \varepsilon$ where $Q(\cdot)$ is the Gaussian tail function. For example, $m = 6.467$ and $\varepsilon = 10^{-10}$. The observation for a code symbol becomes

$$p(c_i|y_i) \approx \sum_{\substack{x_i \in \mathcal{W}_i \\ x_i \equiv c_i}} p(x_i|y_i). \tag{5.4}$$

Limiting the search for a lattice component $x_i = c_i + pz_i$ to $\mathcal{W}_i$ brings an *a priori* on z_i. For a given symbol value c_i and a given channel observation y_i, the search for the unknown z_i is now restricted to $[(y_i - c_i - m\sigma)/p, (y_i - c_i + m\sigma)/p]$. The number of admissible integer translations z_i is $\mu_i(y_i, c_i)$ given by

$$\mu_i(y_i, c_i) = |\{x_i \in \mathcal{W}_i : x_i \equiv c_i \mod p\}|.$$

Consequently, the prior on z_i is given by $\pi(z_i) = 1/\mu_i(y_i, c_i)$. The implementation can be further simplified if p is large enough. Indeed, taking $2m\sigma \leq p$ yields $\mu_i(y_i, c_i) = 1$, for all y_i and all c_i. Since we consider variance noise values up to Poltyrev capacity (cf. Definition 2.19), the latter condition is satisfied when $2m\sigma_{\max} \leq p$, where $\sigma^2_{\max}$ is given by (3.1), which translates into $p^R \geq 2m/\sqrt{2\pi e}$.

Summarising, we decode an LDA (over $\mathbb{Z}$) lattice point coordinate-wise as follows, for a fixed index $i = 1, \ldots, n$:

- *Initialisation*: compute $p(x_i|y_i)$ (5.1) for all $x_i \in \mathcal{W}_i$ and add them as described in (5.4) to get the p values of $p(c_i|y_i)$.
- *Iterations*: apply Belief Propagation with input $p(c_i|y_i)$ to compute the p values of $p(c_i|C, \mathbf{y} \setminus \{y_i\})$.
- *Final decision*: for every $x_i \in \mathcal{W}_i$, compute the product in (5.3) and find the $x_i = \hat{x}_i$ that maximises it.

An alternative strategy for the final decision consists in taking as $\hat{x}_i$ the representant of the class modulo p that is the closest to y_i and that maximises the extrinsic probability $p(c_i|C, \mathbf{y} \setminus \{y_i\})$. Notice that, when p is large enough (or when the noise is weak enough, too), the width of the window $\mathcal{W}_i$ is smaller than p itself and the classes modulo p are represented by at most one integer around y_i, as anticipated before, and the two different strategies for the final decision eventually coincide.

5.2 Optimisation and decoding performance

In this section, we present some details on the choice of the LDPC codes for the construction of the LDA lattices that we have tested; after that, we conclude with some simulation results and the comparison with the performance of already known lattice families.

The core of the lattice is of course the p-ary LDPC code and its choice may be optimised. In the classical binary setting, an LDPC code is identified by its parity-check matrix and, equivalently, by the associated Tanner graph. When the entries of the parity-check matrix are non-binary, the Tanner graph is built as usual, and in addition, a label is associated to every edge; this label is equal to the corresponding non-zero entry in the parity-check matrix of the code (see for example [SFS08]).

Optimising the choice of the p-ary code coincides with optimising the related labelled Tanner graph. In the binary case, this is often reduced to choosing a graph without small cycles. In the case of p-ary LDPC codes, we also choose in a clever way the non-zero p-ary entries of the parity-check matrix (that is, the p-ary labels of the graph edges). This aspect has a significant impact on iterative decoding, at least for relatively small values of p, and has not been previously considered. The "non-triviality" of the graph labels guarantees the existence of better codes with respect to their binary equivalents, resulting in a more powerful and improved Construction A.

5.2.1 Choice of the coefficients for the parity-check equations

In order to make a good choice for the coefficients of the parity-check matrix H of the LDPC code, we investigate the *single parity-check (SPC) code* defined by each parity-check equation (the rows of H). Formally, we define

$$C_{\mathrm{SPC}} = \{\mathbf{c} = (c_1, c_2, \ldots, c_\Delta) \in \mathbb{F}_p^\Delta : h_1 c_1 + h_2 c_2 \ldots + h_s c_s \equiv 0 \bmod p\}$$

as the SPC code associated with the *non-zero* coefficients $h_1, h_2, \ldots, h_\Delta \in \mathbb{F}_p \setminus \{0\}$ of a row of H. We say that this row has *degree* equal to Δ.

Note that the message-passing decoder applies MAP decoding to the individual SPC codes. Contrary to the binary case, there are many choices for an SPC code and they may have a strong influence over MAP decoding. In particular, (5.1) shows that the minimum *Euclidean* distance of the SPC code will be an important parameter and we choose to optimise it. The Euclidean minimum distance is defined as

$$\mathrm{d}_{\min}(C_{\mathrm{SPC}}) = \min_{\mathbf{c} \in C_{\mathrm{SPC}} \setminus \{\mathbf{0}\}} ||\Phi(\mathbf{c})||$$

(where Φ, as before, is the embedding of the code into $\mathbb{Z}$). Experiments confirm that coefficients h_i's that maximise $d_{\min}(C_{\mathrm{SPC}})$ yield a significantly improved performance over random h_i's for Construction A over with $\mathbb{Z}$.

We will focus for a moment on this kind of lattices and show how to implement the good choice of the coefficients in the particular case for which we show the

simulation results later. With these parameters, one can see that $\mathrm{d}_{\min}(C_{\mathrm{SPC}})$ cannot be greater than $\sqrt{3}$. The condition $d_{\min}(C_{\mathrm{SPC}}) \neq 1$ is an immediate consequence of the fact that all the h_i's are non-zero. We can find how to avoid a Euclidean minimum distance of $\sqrt{2}$ as follows: let $(c_1, c_2, \ldots, c_\Delta)$ be a point of C_{SPC} of smallest Euclidean norm;

$$\begin{aligned}\mathrm{d}_{\min}(C_{\mathrm{SPC}}) = \sqrt{2} &\iff \sqrt{c_1^2 + c_2^2 + \ldots + c_\Delta^2} = \sqrt{2} \\ &\iff c_i, c_j = \pm 1,\ \exists\ i, j \in \{1, \ldots, \Delta\} \\ &\qquad \text{and } c_k = 0\ \forall k \neq i, j.\end{aligned}$$

Also, $\mathbf{c}$ must satisfy the parity-check equation, that becomes

$$\pm h_i \pm h_j = 0,\ h_i = \pm h_j.$$

This means that the condition

$$h_i \neq \pm h_j,\ \forall i, j \in \{1, \ldots, \Delta\} \tag{5.5}$$

suffices to impose $\mathrm{d}_{\min}(C_{\mathrm{SPC}}) > \sqrt{2}$.

Our simulations have directed us towards the choice $\Delta = 5$: in this case the first value of p for which we may have $\mathrm{d}_{\min} > \sqrt{2}$ is $p = 11$ and experimentally, this has turned out to be the optimum choice of p for regular LDPC codes.

5.2.2 Tanner graph construction

Generally, random graphs give good performance, provided that one manually removes all 4-cycles and guarantees a girth of at least 6. We have anyway preferred to use LDPC codes whose corresponding graph is built by means of the Progressive Edge-Growth algorithm (PEG) [HEA05]. This algorithm builds the graph edge by edge, in an iterative manner that *locally maximises the current girth of the graph during construction.* Experimentally, we have seen that PEG-obtained graphs allow to reach better symbol error rates (SER), thanks to a "deeper" error floor region with respect to random graphs. At the same time, in the waterfall region of random graphs, PEG-obtained graphs have very similar performance.

5.2.3 Simulation results

We will show here some simulation results and compare them with what is known in the literature about other families of lattices used for the transmission of information. As we have already specified, we decode infinite LDA constellations by the means of the iterative algorithm that we have just presented. We will evaluate the performance of LDA lattices as a function of the noise variance: the best lattices are the ones which attain small symbol error rates for values of the noise variance that

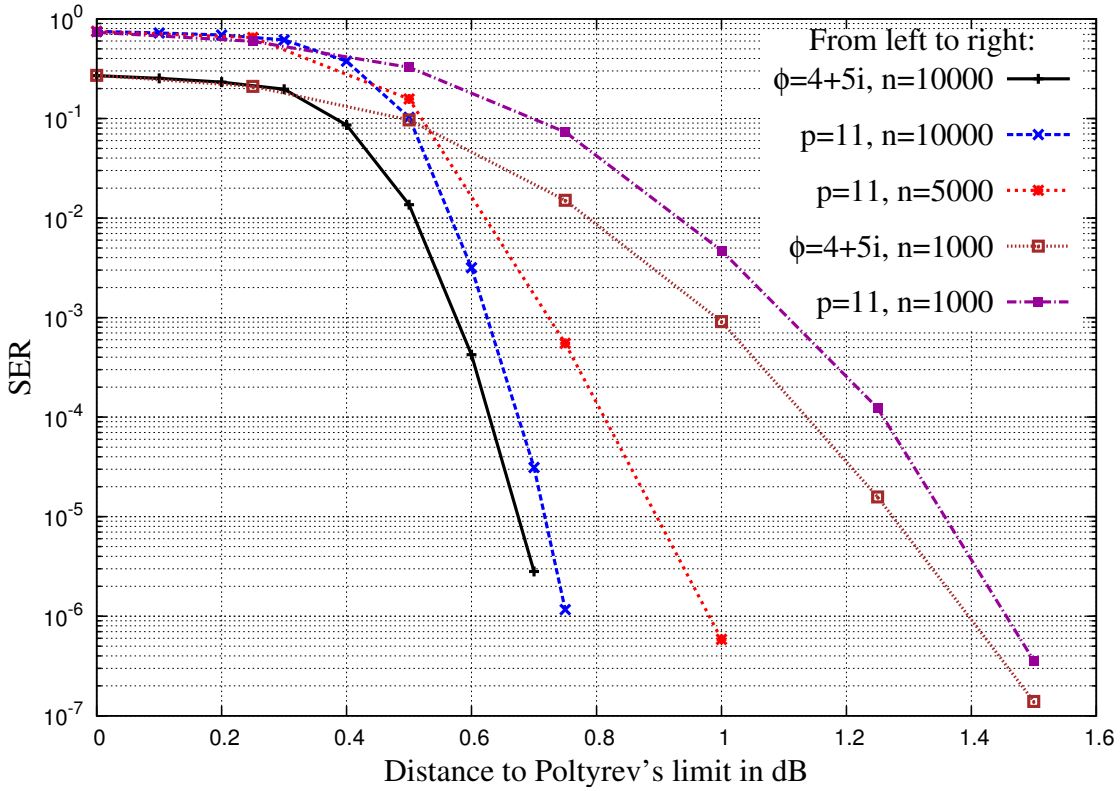

Figure 5.2: Symbol error rate versus distance to Poltyrev limit for LDA lattices.

are close to $\sigma^2_{\max}$. We will speak of distance from capacity, meaning the distance (in dB) of the channel noise variance from $\sigma^2_{\max}$. We recall that

$$\sigma^2_{\max} = \begin{cases} \frac{1}{2\pi e} p^{2(1-R)}, \text{ for } (L, L') = (\mathbb{Z}, p\mathbb{Z}), \\ \frac{1}{2\pi e} p^{(1-R)}, \text{ for } (L, L') = (\mathbb{Z}[i], \phi\mathbb{Z}[i]). \end{cases}$$

Figure 5.2 presents our best experimental results. The values of the parameters that we have fixed in the simulations are the ones that experimentally have given the best performance till now. The number of decoding iterations has been fixed to at most 200 in all simulations.

As far as LDA lattices obtained by p-ary Construction A over $\mathbb{Z}$ are concerned, we have only investigated regular LDPC codes and similarly to the case of binary LDPCs constructed as binary images of q-ary LDPCs [PFD06], we have found that a degree 2 per variable node yields the best results. As mentioned before, the most interesting case to come up was that of a $(2, 5)$-regular code with $p = 11$.

As described in Section 5.2.2, the graph is built using the PEG algorithm, with the slight modification with respect to [HEA05] that the check nodes degree distribution is fixed, too. The non-zero entries of the parity-check equations are chosen as described in Section 5.2.1. Fig. 5.2 shows that for $n = 1000$, we attain a SER of less than 10^{-6} at 1.5 dB from Poltyrev capacity. This corresponds to an improvement of about 0.2 dB with respect to the performance of LDLC [SFS08] at a SER of 10^{-5}.

With a similar lattice in dimension $n = 5000$, we attain a SER of less than 10^{-6} at 1 dB from Poltyrev capacity, which corresponds to an improvement of more than 0.2 dB with respect to Irregular LDPC lattices and of about 0.8 dB with respect to Regular LDPC lattices (see [BC08]).

In dimension $n = 10000$, our LDA $\mathbb{Z}$-lattice provides a SER of 10^{-6} at 0.75 dB from Poltyrev capacity, which is better than what LDLC do [SFS08].

An even more interesting result is given by the performance of LDA $\mathbb{Z}[i]$-lattices (construction A with $(L, L') = (\mathbb{Z}[i], \phi\mathbb{Z}[i])$. As in the previous examples, the Tanner graph is $(2, 5)$-regular, while the prime ideal used for the modulo operation is $(4+5i)$, corresponding to $p = 41$. In (real) dimension $n = 1000$ ($\ell = 500$), a SER of about 10^{-5} is attained at 1.25 dB from Poltyrev capacity, equalling the performance of Turbo Lattices [SSP12], while, for $n = 10000$ ($\ell = 5000$), the same SER is attained at about 0.7 dB from Poltyrev capacity.

Chapter 6

Conclusion and ideas for future work

This thesis is devoted to the application of lattices to the transmission of information over the AWGN channel. We have considered several aspects of this problem, that can be summarised as follows:

- We have started by introducing the family of LDA lattices, that put together the strength of Construction A and LDPC codes (cf. Definition 3.1).
- We have analysed the theoretical asymptotical performance of two families of LDA lattices, when infinite constellations are taken into accounts. Namely, we have proved that there exist two Poltyrev-capacity-achieving families of LDA lattices (cf. Theorem 3.1 in Section 3.2.3 and Theorem 3.2 in Section 3.3.4).
  - The first of these families is characterised by a logarithmically growing parity-check equation degree. The field size of the LDPC codes underlying the LDA family is a prime number p, whose order of magnitude is n^λ, for some constant λ that can be chosen a priori as small as wanted. Recall, anyway, that for the proof we need to let n tend to infinity.
  - The second family is characterised by a constant parity-check equation degree. So, the Poltyrev-capacity-achieving result goes a little beyond the previous one and requires a different random ensemble of LDA lattices. This result strictly depends on the expansion properties of the Tanner graphs associated with the LDA lattices. Lemma 3.3 quantitatively expresses these properties. Once again, p is taken to be of the order of n^λ, though this time λ has to satisfy some lower bounds, that also imply some lower bounds on the parity-check equation degree. Nevertheless, these bounds are very reasonable and allow us to take a small p with respect to n and still have a small parity-check equation degree.
- We have gone beyond the works of Erez and Zamir [EZ04], Ordentlich and Erez [OE12], and Ling and Belfiore [LB13], giving a new proof that finite Voronoi constellations of Construction A lattices can achieve the capacity of the AWGN channel under (MMSE-scaled) lattice decoding (cf. Theorem 4.1 in Section 4.2.4), when SNR > 1. Our proof does not require the sharing of

common randomness (dither) between the sender and the receiver and we give an explicit encoding method of uncoded messages into lattice points for this communication scheme. In this setting, the prime number p can be of the same size as n^λ for any positive constant λ big enough (the lower bound variating between 1/2 and 2/3 depending on the underlying code structure). This is an improvement with respect to other proofs that use Construction A, too.

- We have adapted the previous result to LDA lattices, that is we have shown that they are Shannon-capacity-achieving, too (cf. Theorem 4.2 in Section 4.3.4). Once again, we have exploited the expansion properties of the Tanner graphs associated with the LDA lattices. The value of the parameter λ in this case is lower bounded by some bigger constants, as well as the column and row degrees of the parity-check matrices associated with the lattice family. These values depend on the rate of the underlying LDPC codes and on the expansion constants that we fix.

- We have described an iterative decoding algorithm based on Belief Propagation for LDA lattices that relies on their LDPC-code underlying structure. We have also shown some satisfactory decoding simulation results of some particular LDA lattices and compared them to the performance of other lattice families proposed in the literature.

Finally, there are many directions in which the research on LDA lattices and their applications can evolve. Here is a list of useful starting point:

- First of all, other simulations and numerical experiments could be carried out, to optimize the parameters involved in the construction of well-performing LDA lattices. Namely, for fixed dimensions, there is no deterministic way of deciding what is the best prime number p, what are the best row and column degrees for the parity-check matrices, what is the best rate of the underlying LDPC code, what are the optimal labels for the Tanner graph...

- It would be interesting to experiment with other kinds of Construction A, for example extending the construction to the Eisenstein integers $\mathbb{Z}[\omega]$.

- The decoding algorithm implementation can be further optimised, allowing simulations in higher dimensions than 1000 or 10000.

- Of course, a comparison to classical coded modulations with finite constellations should be done. Our theoretical results suggest that an interesting test would be to investigate the performance of LDA Voronoi constellations under our iterative decoding. Now, the problem in this case is how to precisely design the constellation and encode the messages into lattice points. Our coding scheme explicitely says how to do it, but it requires a lattice decoder, which is computationally too complex to be applied in the high dimensions we are interested in. On the other hand, our iterative decoder is not precise enough to be good for encoding. To come up with a solution to this problem would be a definitely valuable result.

- The main application of LDA lattices in this paper was error correction on a Gaussian channel, but other numerous potential applications exist such as physical layer network coding and physical layer security. An interesting research domain seems to be the application of LDA lattices to coding over the wiretap channel.

Bibliography

[Bal98] Paolo Baldi. *Calcolo delle Probabilità e Statistica*. McGraw-Hill Libri Italia, Milano, 1998.

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296:625–635, 1993.

[Bas81] Leonid A. Bassalygo. Asymptotically optimal switching circuits. *Problems of Information Transmission*, 17(3):206–211, 1981.

[BB03] Loïc Brunel and Joseph J. Boutros. Lattice decoding for joint detection in direct-sequence CDMA systems. *IEEE Transactions on Information Theory*, 49(4):1030–1037, April 2003.

[BC08] Ihn-Jung Baik and Sae-Young Chung. Irregular low-density parity-check lattices. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 2479–2483, 2008.

[BCJR74] Lalit R. Bahl, John Cocke, Frederick Jelinek, and Josef Raviv. Optimal decoding for linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, 20(2):284–287, March 1974.

[CS82] John H. Conway and Neil J. A. Sloane. Voronoi regions of lattices, second moments of polytopes, and quantization. *IEEE Transactions on Information Theory*, 28(2):211–226, March 1982.

[CS99] John H. Conway and Neil J. A. Sloane. *Sphere Packings, Lattices and Groups*, volume 290 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 3rd edition, 1999.

[CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, 1991.

[dB75] Rudi de Buda. The upper error bound of a new near-optimal code. *IEEE Transactions on Information Theory*, 21(4):441–445, July 1975.

[dB89] Rudi de Buda. Some optimal codes have structure. *IEEE Journal on Selected Areas in Communications*, 7(6):893–899, August 1989.

[dPBZ13] Nicola di Pietro, Joseph J. Boutros, and Gilles Zémor. New results on construction a lattices based on very sparse parity-check matrices. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 1675–1679, 2013.

[dPBZB12] Nicola di Pietro, Joseph J. Boutros, Gilles Zémor, and Loïc Brunel. Integer low-density lattices based on construction a. In *IEEE Information Theory Workshop Proceedings (ITW)*, pages 422–426, 2012.

[dPBZB13] Nicola di Pietro, Joseph J. Boutros, Gilles Zémor, and Loïc Brunel. New results on low-density integer lattices. In *Information Theory and Applications Workshop Proceedings (ITA)*, pages 1–6, 2013.

[Ebe13] Wolfgang Ebeling. *Lattices and Codes*. Springer Spektrum, Wiesbaden, 3rd edition, 2013.

[ELZ05] Uri Erez, Simon Litsyn, and Ram Zamir. Lattices which are good for (almost) everything. *IEEE Transactions on Information Theory*, 51(10):3401–3416, October 2005.

[EM05] Uri Erez and Gadi Miller. The ML decoding performance of LDPC ensembles over $\mathbb{Z}_q$. *IEEE Transactions on Information Theory*, 51(5):1871–1879, May 2005.

[Ere02] Uri Erez. *Coding with Known Interference ans Some Results on Lattices for Digital Communication*. PhD thesis, Tel-Aviv University, 2002.

[ESZ05] Uri Erez, Shlomo Shamai, and Ram Zamir. Capacity and lattice strategies for cancelling known interference. *IEEE Transactions on Information Theory*, 51(11):3820–3833, November 2005.

[EZ04] Uri Erez and Ram Zamir. Achieving $\frac{1}{2}\log(1 + \mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding. *IEEE Transactions on Information Theory*, 50(10):2293–2314, October 2004.

[For89] G. David Forney, Jr. Multidimensional constellations – part II: Voronoi constellations. *IEEE Journal on Selected Areas in Communications*, 7(6):941–958, August 1989.

[For92] G. David Forney, Jr. Trellis shaping. *IEEE Transactions on Information Theory*, 38(2):281–300, March 1992.

[Gal63] Robert G. Gallager. *Low-Density Parity-Check Codes*. M. I. T. Press, 1963.

[Ger79] Allen Gersho. Asymptotically optimal block quantization. *IEEE Transactions on Information Theory*, 25(4):373–380, 1979.

[GZ07] Philippe Gaborit and Gilles Zémor. On the construction of dense lattices with a given automorphism group. *Annales de l'Institut Fourier*, 57(4):1051–1062, 2007.

[HEA05] Xiao-Yu Hu, Evangelos Eleftheriou, and Dieter M. Arnold. Regular and irregular progressive edge-growth algorithm. *IEEE Transactions on Information Theory*, 51(1):386–398, January 2005.

[HR76] Carlos R. P. Hartmann and Luther D. Rudolph. An optimum symbol-by-symbol decoding rule for linear codes. *IEEE Transactions on Information Theory*, 22(5):514–517, September 1976.

[IZF13] Amir Ingber, Ram Zamir, and Meir Feder. Finite-dimensional infinite constellations. *IEEE Transactions on Information Theory*, 59(3):1630–1656, March 2013.

[KFL01] Frank R. Kschischang, Brendan J. Frey, and Hans-Andrea Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, February 2001.

[KL78] Gregory A. Kabatiansky and Vladimir I. Levenshtein. On bounds for packings on a sphere and in space. *Problems of Information Transmission*, 14(1):1–17, 1978.

[LB13] Cong Ling and Jean-Claude Belfiore. Achieving the AWGN channel capacity with lattice Gaussian coding. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 1416–1420, July 2013.

[LLBS12] Cong Ling, Laura Luzzi, Jean-Claude Belfiore, and Damien Stehlé. Semantically secure lattice codes for the gaussian wiretap channel, 2012. Available at http://arxiv.org/abs/1210.6673v3.

[Loe97] Hans-Andrea Loeliger. Averaging bounds for lattices and linear codes. *IEEE Transactions on Information Theory*, 43(6):1767–1773, November 1997.

[LSZ93] Tamás Linder, Christian Schlegel, and Kenneth Zeger. Corrected proof of de Buda's theorem. *IEEE Transactions on Information Theory*, 39(5):1735–1737, September 1993.

[Mac99] David J. C. MacKay. Good error correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory*, 45(2):399–431, March 1999.

[MdOB90] Marcio Magalães de Oliveira and Gérard Battail. A capacity theorem for lattice codes on Gaussian channels. In *SBT/IEEE International Telecommunications Symposium Proceedings*, Rio de Janeiro, Brazil, September 3-6 1990.

[MR04] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *IEEE Annual Symposium on Foundations of Computer Science Proceedings*, pages 372–381, Rome, Italy, October 2004.

[MS77] Florence J. MacWilliams and Neil J. A. Sloane. *The Theory of Error-Correcting Codes*, volume 16 of *North-Holland Mathematical Library*. North-Holland Publishing Company, Amsterdam, New York, Oxford, 1977.

[OE12] Or Ordentlich and Uri Erez. A simpler proof for the existence of "good" pairs of nested lattices, 2012. Available at http://arxiv.org/abs/1209.5083.

[PFD06] Charly Poulliat, Marc Fossorier, and David Declercq. Design of non binary LDPC codes using their binary images: Algebraic properties. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 93–97, 2006.

[Pol94] Gregory Polyrev. On coding without restrictions for the AWGN channel. *IEEE Transactions on Information Theory*, 40(2):409–417, March 1994.

[Rog59] Claude A. Rogers. Lattice coverings of space. *Mathematica*, 6:33–39, 1959.

[Rog64] Claude A. Rogers. *Packing and Covering*. Cambridge University Press, Cambridge (U. K.), 1964.

[RU08] Tom Richardson and Rüdiger Urbanke. *Modern Coding Theory*. Cambridge University Press, New York, 2008.

[SBP06] Mohammad-Reza Sadeghi, Amir H. Banihashemi, and Daniel Panario. Low-density parity-check lattices: Construction and decoding analysis. *IEEE Transactions on Information Theory*, 52(10):4481–4495, October 2006.

[SFS08] Naftali Sommer, Meir Feder, and Ofir Shalvi. Low-density lattice codes. *IEEE Transactions on Information Theory*, 54(4):1561–1585, April 2008.

[SS13] Mohammad-Reza Sadeghi and Amin Sakzad. On the performance of 1-level LDPC lattices. In *Iran Workshop on Communication and Information Theory Proceedings(IWCIT)*, pages 1–5, 2013.

[SSP12] Amin Sakzad, Mohammad-Reza Sadeghi, and Daniel Panario. Turbo lattices: Construction and error decoding performance, 2012. Available at http://arxiv.org/abs/1108.1873v3.

[UR98] Rüdiger Urbanke and Bixio Rimoldi. Lattice codes can achieve capacity on the AWGN channel. *IEEE Transactions on Information Theory*, 44(1):273–278, January 1998.

[VB99] Emanuele Viterbo and Joseph J. Boutros. A universal lattice code decoder for fading channels. *IEEE Transactions on Information Theory*, 45(5):1639–1642, July 1999.

[YL12] Yanfei Yan and Cong Ling. A construction of lattices from polar codes. In *IEEE Information Theory Workshop Proceedings (ITW)*, pages 124–128, 2012.

[YLW13] Yanfei Yan, Cong Ling, and Xiaofu Wu. Polar lattices: Where Arıkan meets Forney. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 1292–1296, 2013.

[ZF96] Ram Zamir and Meir Feder. On lattice quantization noise. *IEEE Transactions on Information Theory*, 42(4):1152–1159, July 1996.