Re-proving Channel Polarization Theorems: An Extremality and Robustness Analysis

THÈSE N^{O} 6403 (2015)

PRÉSENTÉE LE 16 JANVIER 2015

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS

LABORATOIRE DE THÉORIE DE L'INFORMATION

PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Mine ALSAN

acceptée sur proposition du jury:

Prof. M. Odersky, président du jury
Prof. E. Telatar, directeur de thèse
Prof. E. Arıkan, rapporteur
Prof. A. Guillén i Fàbregas, rapporteur
Prof. A. Lapidoth, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2014

Aileme.

To my parents Nuran and Nejat, my sister Bahar, and my grandmother Sabiha.

The single biggest problem in communication
is the illusion that it has taken place.
— George Bernard Shaw.

Yet, picture yourself in a place where you don't even
have that illusion. Then, listen to its silence.
...
A man once told me
"I will not talk to you ever again!"
after I complained about the situation in which I
was placed by this man and another women to sit in a room
with another man who was suspected by this other women not to talk with
women due to his religious and cultural background based on the observation that
it took three years for the man to talk with the other women by looking into her eyes
and about whom the man who will not talk to me ever again once asked me
whether the man was talking or not inside the room and then told me
that I should talk with the man sitting in the room
because the man was very nice
but just shy.
...
The man in the room later told me
"I thought I should keep the room quiet".
...
At the end, I found myself in an absurd place where you don't even
have that illusion. Only humor was left in its silence.
— Une Drôle de Silence

Acknowledgements

This thesis was supported by Swiss National Science Foundation under grant number 200021-125347/1.

I would like to thank my thesis director Prof. Emre Telatar and committee members Prof. Erdal Arıkan, Prof. Albert Guillén i Fàbregas, and Prof. Amos Lapidoth for reading my thesis and providing valuable comments on my research. I would also like to thank Prof. Martin Odersky for agreeing to be the president of my committee. I was lucky to have Arıkan, the father of polar codes, comment on my thesis, and I am thankful to Fàbregas for providing detailed comments which helped improve the presentation of my thesis. I am grateful to Dr. Olivier Lévêque for editing my French translation of the abstract. In addition, I would like to thank Prof. Bixio Rimoldi and Prof. Micheal Gastpar for acting as jury during my candidacy exam in the begining of my PhD. Finally, I am thankful to Damir Laurenzi and Giovanni Cangiani for always kindly helping with the computer support and to Françoise Behn for providing a professional administrative support.

I would like to express my deepest gratitude to Emre for his guidance and friendship throughout my studies and tell you a bit about his role in shaping the course of my PhD. After all, I owe an explanation to the many who asked me over time questions about him such as how it was to work with a 'celebrity' (and why he was not returning their emails (and many other questions which would not be appropriate to share...)). Let me start by a confession. I remember a musical history professor commenting about the personality of a well known classical music genius: "Today, we are all impressed by his music, yet if you knew him in person, you would probably not sit in the same table to have a cup of tea with this person". For Emre, people knowing him would universally agree on the opposite: the one thing you would definitely want to do would be to sit around a white tableclothed table and have (not tea of course)... In reality, he fooled me to join his lab for PhD not by his impressive knowledge and teaching, but via his generous partage of culinary art pieces accompanied with grape juice. Shame on me! Joking aside, he is the most humble and constructive person one could imagine. I am not writing to curry favor. We had serious disagreements at times due to an issue I faced in the work environment that was caused by 'foreign powers' and that was completely irrelevant to my work and our student/supervisor interaction. The behind-the-scenes of this issue will be left as a mystery here as there is no negative acknowledgments (NACKs)

chapter, however, I thank Emre for listening, trying to be fair and constructive despite our disagreements. Even if I chose to follow my own way, I learned a lot from his approach to conflict resolution and I know that I would benefit immensely from his wisdom in my future carrier. I understood how much it is valuable to interact with people having constructive and elegant minds. If everyone was as careful as him in his communications, academia would be a much pleasant workplace for women. Let me lighten the mood with another confession. Emre is perhaps the only person who became the object of my selective attention. I would fondly remember the meetings we had during my master's thesis and the beginning of my PhD. These meetings would start with an opening and closing ceremony, where we would chit chat from this and that (Turkey, family, home decoration...). In between, Emre would perform his intellectual show on board and I would follow my eyes wide open. He would explain me something he probably just had in his mind, like a recent or old result, a research topic, an exam question, or I would input him with a notion I have heard in a talk or read somewhere that confused me, or even the raised cosine filter, and he would explain the whole topic on board from scratch with clarity. If I don't understand a point, he would retry in a slower pace telling the story slightly differently! At the end, he would humbly apologize for taking my time by telling all of these as they would mostly be irrelevant to my work. As for me, I would make sure before leaving that I have a ticket for the performance next week. In fact, I would listen ultra attentive. Although you can mock about it, my efforts paid well, many of the original results of my thesis came out by treating the content of our discussions like a hidden gem. The ironic part is that Emre would be himself surprised (which I am quite proud of!). I thank him for letting me follow my curiosity instead of pressuring on getting results. This seems to be rarely possible in today's academic practice. This does not mean, however, that I am the only owner of this work. On the contrary, I had access to one of the most up-to-date 'mathematical toolboxes'. This thesis would not have been possible without Emre's contributions and proof ideas. Last but not least, I also had a lot of freedom on the publication process. Thanks to the generous funding provided by SNSF and the lab, I had the opportunity to attend many international conferences and, this should stay between us so I will whisper, see all around the world!

Next in line are my colleagues and friends from work. It was a real pleasure to share the office with Marc who would welcome me everyday with his courteous salutation in French. I don't think I have ever told him how peaceful sharing the office with him was and I thank him sincerely; he was the day watch and I was the night watch of our territory, we sent away the enemies singing them "Hojotoho!, Hojotoho!"(Die Walküre or the funny version)... He is a person who knows the true meaning of mutual respect; he would always kindly ask permission for anything (except walking barefoot). Before he left for an internship, he made sure I won't feel alone in his absence by hanging to our door the poster of his climber 'girlfriend' who would stare at me the whole day with a teasing smile. Besides, I truly appreciate his

activity centered lifestyle, his organization skills, and his fight for freedom: as the mathematician, he is fighting harder than anyone else in the lab to free probabilities (sorry for this one!). It was equally a pleasure to be office neighbors with Adrian and Young, two of the nicest people around. I am really happy our acquaintance with Adrian from the master turned into a good friendship. I have a small collection of little souvenirs he brought back returning from his trips abroad; this became a tradition. I am grateful to him for not covering the sensor of my mouse with a sticker and for inviting me to many activities. Through him, I joined two volleyball teams; being in Lausanne would not have been that much fun without our Wednesday Kebab&Smash nights. Sportive activities are a key to make colleagues friends. I am happy that I got to know better Stefano after he joined the volleyball teams. I thank him for not throwing me out of his place during his kind dinner invitations following my 'aggressive' Mafia playing and for being the only member of the lab who lasted on the dance floor during the celebration of my new title. Not to mention that he did not forget his friends, even after he became famous for developing a technology! Back to the neighbors, I am thankful to Young for his help at various occasions. I was really happy to find Ko-Ko at his office the night before handing the first draft of my thesis. He may not have shared my happiness though; at midnight, he helped me for almost an hour punch holes to 6×200 pages before binding them with plastic ring binders . Whether for a lunch or coffee break, outdoor activity, Christmas dinner, or while attending a conference, I enjoyed sharing time with Lyudmila (the mommy), Andrei, Karol, Saeid, Iris, László, Emre A., Mani, Rajai, Serj, Rafah, Marco. I am thankful to the previous members of LTHI Ayfer and Mohammed for giving insider information at various times before and after I joined the lab. This paragraph would not be complete if I don't acknowledge from the bottom of my heart the creaking fussball table in our cafeteria for enabling the personal and professional development of the many. I am proud to announce that with Adrian Tarniceriu, Andrei Giurgiu, Karol Kruzelecki, Marc Desgroseilliers, Stefano Rosati, and Young Jun Ko, we elevated the fussball experience from just fun to another level: sport, competition, sportsmanship, and more importantly self-expression. I hereby give the title of Fussdoc without reserve to each and one of us!

Now it is time to pay homage to my dear Turkish friends at EPFL. I cannot effectively summarize all the times we spent together from week day lunches to week-end tea parties and hikes. If it weren't for Tuna, I wouldn't have applied to EPFL. I am grateful to him for giving me such a useful advice years ago, despite not knowing me very much. I really admire the way he draws his way so consistently towards a 'bright future'. Yasemin and Tuna would be the first ones I would knock the door for help and I am grateful to them for always warmly welcoming me in their house. I am thankful to Elif for listening and sharing opinions on the joys and frustrations of many things during our long lunch breaks and also our coffee breaks while attending the Analysis course together; it was relieving to have the opportunity to talk with another 'women in science and engineering'. However, it was quite a

challenge to answer her questions about French grammar. I, as the authority, would often have to resort to quoting my middle school French professor saying: “Il y a une chose à retenir sur la grammaire française, que les règles sont les exceptions”. I also shared many great moments with Gürkan, my yoga buddy and the most ‘outdoor’ person of the Turkish community, Engin, the dive master who carried his laptop in his backpack an entire day wandering the streets of Venice, and Ahmed, whose part-time life in Barcelona has been a complete mystery to all of us. Thanks to Ahmed and Gürkan, I had my first (and only so far) tent camping and lac swimming experience in Switzerland. It was quite tiring to watch tireless kids ‘fire’ the 1st of August and quite challenging to keep up with them during the hikes. My stomach presents her special thanks to Engin for all the ‘künefe’ trials in which everyone except him would be satisfied with the taste. I am also glad that I made acquaintance via coincidence with Özlem. I thank her for regularly checking on me and making me social again by inviting me to fun activities. I am also glad that my road crossed with many other members of the Turkish community; nice moments were shared with Ali-Galip, Anıl, Ebru, Emrah, Eymen, Gilles (Turkish?), Gözen, Gökhan, Onur, Seliz, Zafer, and many others that I apologize for not mentioning by their names... I am also thankful to Madame Zerrin for helping me join the committee of ACIDE.

At last comes the turn of my old friends who don’t live in Lausanne. Though we don’t find the opportunity to communicate very often, special thanks to Ayda, Berna, Mehmet, Özge, Sinem, and Sophia for inviting me over and all the good memories we shared.

Pages would not be enough to thank my family. Being far away made me understand better how valuable they are to me and how easy my life used to be with them. I would like to thank my parents Nuran and Nejat for their love and constant support. I especially appreciate them for raising us to become diverse, independent, and self-confident persons. I express my heartfelt gratitude to my grandmother Sabiha for giving so much love and care to us and for coming abroad with us to help. She is the most accommodating person I know and she does have a great sense of humor. I am very lucky to have a sister as Bahar. I am so grateful to her for being my best friend today, and in the past, for always being patient with her ‘little’ sister. It has occurred to me while writing that she might be the reason why I chose ‘math&science’ in high school. Her high school math teacher begged her for an approval: “Please tell me that I am a good teacher and you understand what I teach”, because no one else seemed to understand. She thought he was a good teacher, yet I think she did not realize at that point how smart she is (no disrespect to the teacher intended). At the end she decided to major in a social science, and me, I know who to blame now! To tell the truth, I am quite amazed by how she manages to multitask everything at home and at work. I am thankful to Bahar and my brother-in-law Gökhan for visiting me in Lausanne and for always joyfully welcoming me in their home together with the two cutest members of the family.

Abstract

The general subject considered in this thesis is a recently discovered coding technique, polar coding, which is used to construct a class of error correction codes with unique properties. In his ground-breaking work, Arıkan proved that this class of codes, called polar codes, achieve the symmetric capacity — the mutual information evaluated at the uniform input distribution — of any stationary binary discrete memoryless channel with low complexity encoders and decoders requiring in the order of $O(N \log N)$ operations in the block-length N. This discovery settled the long standing open problem left by Shannon of finding low complexity codes achieving the channel capacity.

Polar codes are not only appealing for being the first to 'close the deal'. In contrast to most of the existing coding schemes, polar codes admit an explicit low complexity construction. In addition, for symmetric channels, the polar code construction is deterministic; the theoretically beautiful but practically limited "average performance of an ensemble of codes is good, so there must exist one particular code in the ensemble at least as good as the average" formalism of information theory is bypassed. Simulations are thus not necessary in principle for evaluating the error probability which is shown in a study by Telatar and Arıkan to scale exponentially in the square root of the block-length. As such, at the time of this writing, polar codes are appealing for being the only class of codes proved, and proved with mathematical elegance, to possess all of these properties.

Polar coding settled an open problem in information theory, yet opened plenty of challenging problems that need to be addressed. This novel coding scheme is a promising method from which, in addition to data transmission, problems such as data compression or compressed sensing, which includes all types of measurement processes like the MRI or ultrasound, could benefit in terms of efficiency. To make this technique fulfill its promise, the original theory has been, and should still be, extended in multiple directions. A significant part of this thesis is dedicated to advancing the knowledge about this technique in two directions. The first one provides a better understanding of polar coding by generalizing some of the existing results and discussing their implications, and the second one studies the robustness of the theory over communication models introducing various forms of uncertainty or variations into the probabilistic model of the channel.

The idea behind the design of a polar code is a phenomenon called channel

polarization. This consists of synthesizing two new channels by applying the polar transform to two other channels. In the process, it is observed that while the sum symmetric capacities are preserved, the overall reliability is improved by creating 'variance', i.e., the two new channels are created in such a way that the difference between their symmetric capacities is strictly larger than the difference between the symmetric capacities of the original pair of channels as long as the channels are not already perfect or completely noisy. Consequently, the new synthetic channels polarize: one becomes better and the other worse than the original mediocre channels. This result follows as a corollary to information combining which shows that the extremal bounds of the difference between the symmetric capacities of the created channels are attained by the binary erasure channel and the binary symmetric channel.

The mutual information, though fundamental, is not the only information measure of interest to the information theory community. In the field's literature, 'Gallager's $E_0(\rho)$', for $\rho > -1$, is a well rooted family of information measures appearing in various error exponent problems and also in sequential decoding. The mutual information, determining the theoretical limit of information transmission, and the cutoff rate, another channel parameter which used to be interpreted as the 'practical limit' of information transmission, turn out both to be special cases of $E_0(\rho)/\rho$. In retrospect, Arıkan's discovery came as the offspring of his prior work looking into a method to close the gap between the mentioned two limits.

Based on this account, we study as part of this thesis the evolution of this more general family of information measures under the polar transform. In particular, we prove that the polar transform improves $E_0(\rho)$ for binary input channels. The result helps us understand better why the polar transform yields capacity achieving and low complexity codes: the improvement in $E_0(\rho)$ translates into an improvement in the complexity–error-probability trade-off. This is a concept introduced in the 1996 Shannon Lecture given by Forney. In addition, we prove that even if we change the measure of information from the customary mutual information to $E_0(\rho)$, the binary erasure channel and the binary symmetric channel still remain extremal. Speaking of extremality, we also show independent from any polarization context the extremality of these two channels amongst all binary input channels of a given $E_0(\rho)$ value evaluated at a fixed ρ.

Once a deeper understanding of the technique of polar coding is developed, the thesis proceeds with the study of a practical problem related to the design of polar codes: "robustness against channel parameter variations", as stated in Arıkan's original work. Working out this problem is particularly challenging for polar coding as the initial development revealed that polar codes are channel specific designs. However, from an engineering point of view, it is critical that the results of a theory be robust. This is why right after its conception, partial orderings for channels became relevant for designing polar codes. Two channels are ordered if the code designed for one of the channels can be mapped to a code resulting in at most the same decoding error probability when used over the other channel. In fact, it was once more Shannon

who introduced in a note the concept of partial orderings for discrete memoryless channels. In this thesis, we first touch this topic by introducing a rigorous framework in which we propose to study partial orderings for communication channels in the context of stochastic orders known as convex orderings. In this process, we discover a novel partial ordering for binary discrete memoryless channels we call the symmetric convex ordering. Then, the thesis focuses on different communication models proposed in the literature for building more robust systems; chapters are dedicated to extend the original theory of polar coding to the following complex scenarios:

Coding with a given decision rule— In this scenario, we study the performance of mismatched polar decoders. A mismatched polar decoder is a polar successive cancellation decoder which uses, instead of the true channel's law, the metric of a mismatched channel during the decision procedure. We find the transmission capacity of polar coding with mismatched polar decoding. Moreover, we show that this capacity is lower bounded by a certain family of improving lower bounds converging to the polar mismatched capacity; whenever any of these bounds are positive, strictly positive communication rates can be achieved with properly constructed polar codes. We also observe that the block decoding error probability still decays exponentially in the square root of the block-length as in the matched case. It is worth emphasizing that while extending the theory of polar coding to mismatched communication scenarios, the mismatched polar decoder preserves the $O(N \log N)$ low complexity structure of the 'matched' polar decoder. This structural advantage further motivates polar coding in the presence of a decoding mismatch.

Communication over a class of channels— We also investigate in this thesis the design of robust polar codes over a class of channels. Generally in this scenario, the code designer has access only to a partial knowledge about the true channel through the class to which it belongs. The problem is approached from different angles. First by allowing the decoder to know the true channel, we link polar ordering to the symmetric convex ordering, the novel order introduced by this thesis. Then letting instead the encoder know the channel, we extend the results about the mismatched capacity of polar codes to the compound setting by using the notion of one-sided sets of channels introduced by Abbe and Zheng. Taking yet another approach, we show that polar codes using an approximation at the decoder side are robust over the class of binary symmetric channels. Combining this result with simulations, we provide strong evidence that polar codes are 'practically universal' over binary symmetric channels. Finally, we prove that universality can be traded for complexity by showing that multiple runs of the polar decoder implementing a generalized likelihood ratio test give a universal decoding rule for binary input channels satisfying certain mild conditions. Hence, more resources at the decoder is the price for universality.

Communication over non-stationary channels— A further original contribution of this thesis is the extension of the theory of channel polarization over non-stationary memoryless channels. This is a model which is quite useful to capture the effects of

time-varying noise present in real communication systems as it is no longer assumed that the communication channel is stationary during the transmission of information. As the existing proof techniques are not applicable to this scenario, we first reprove the polarization phenomenon by using only elementary methods. Then by using the same method, we show that Arıkan's construction also polarizes non-stationary memoryless channels in the same way it polarizes stationary ones.

Key words: Polar coding, polar codes, channel polarization, mismatched decoding, compound channels, robust code design, generalized likelihood ratio test (GRLT), coding for non-stationary channels, extremal channels, Gallager's E_0, error exponents, information combining.

Résumé

Le sujet principal de cette thèse est une technique de codage récemment découverte, le codage polaire, destinée à construire une famille de codes correcteurs aux propriétés uniques. Dans son travail de fondateur, Arıkan a démontré que cette famille de codes correcteurs, appelés les codes polaires, atteignent la capacité symétrique— l'information mutuelle évaluée sous une distribution d'entrée uniforme— de tout canal binaire discret sans mémoire et stationnaire avec des codeurs et des décodeurs à faible complexité exigeant de l'ordre de $O(N \log N)$ opérations en la longueur du bloc N. Cette découverte a résolu le problème laissé ouvert par Shannon d'inventer des codes qui atteignent la capacité avec une faible complexité.

Les codes polaires ne sont pas seulement intéressants parce qu'ils sont les premiers 'à conclure l'affaire'. Contrairement à la plupart des systèmes de codage existants, les codes polaires admettent une construction explicite de faible complexité. De plus, pour les canaux symétriques, la construction de codes polaires est déterministe. La technique usuelle de la théorie de l'information attestant que "si la performance moyenne d'un ensemble de codes est bonne, alors il doit y avoir au moins un code de l'ensemble aussi bon que la moyenne" est belle en théorie mais limitée en pratique ; les codes polaries contournent cette approche traditionnelle et des simulations ne sont pas en principe nécessaires pour évaluer la probabilité d'erreur d'un code donné. Une étude réalisée par Telatar et Arıkan indique que celle-ci décroît proportionnellement à l'exponentielle de la racine carrée de la longueur du bloc. A ce titre, au moment d'écrire ces lignes, les codes polaires sont la seule famille de codes démontrée, et démontrée avec élégance mathématique, à posséder toutes ces propriétés.

Le codage polaire a résolu un problème ouvert depuis longtemps en théorie de l'information mais, en même temps, a posé plusieurs problèmes difficiles qui doivent être abordés. Ce nouveau schéma de codage est une méthode prometteuse grâce à laquelle, en plus de la transmission de données, des problèmes tels que la compression de données ou l'acquisition comprimée, comprenant tous les types de processus de mesure comme l'IRM ou l'échographie, pourraient en bénéficier en efficacité. Afin de permettre à cette technique de tenir sa promesse, la théorie originale a été, et devra encore être, étendue dans plusieurs directions. Une partie considérable de cette thèse est consacrée à l'avancement des connaissances sur cette technique dans deux directions. La première permet une meilleure compréhension

du codage polaire en généralisant certains résultats existants et en discutant leurs implications, et la seconde étudie la robustesse de la théorie par rapport aux modèles de communication introduisant diverses formes d'incertitude ou de variations dans le modèle probabiliste du canal.

L'idée derrière la conception de codes polaires est un phénomène appelé la polarisation de canal. Ce phénomène consiste en ceci : à synthétiser deux nouveaux canaux en appliquant la transformée polaire à deux autres canaux. Il est observé que, dans le procédé, la somme des capacités symétriques est conservée tandis que la fiabilité globale est améliorée par la création de 'variance ', c'est-à-dire que les deux nouveaux canaux sont créés de manière à ce que la différence entre leurs capacités symétriques soit strictement plus grande que la différence entre les capacités symétriques des canaux originaux tant que ces derniers ne sont pas déjà sans bruit ou complètement bruités. Par conséquent, les nouveaux canaux synthétiques sont polarisés : l'un devient meilleur et l'autre plus mauvais que les canaux médiocres du début. Ce résultat est un corollaire à 'information combining' qui montre que les limites extrémals de la différence entre les capacités symétriques des nouveaux canaux sont atteintes par le canal binaire à effacement et le canal binaire symétrique.

L'information mutuelle, bien que fondamentale, n'est pas la seule mesure d'information d'intérêt pour la communauté de la théorie de l'information. Dans la littérature, les fonctions $E_0(\rho)$ introduites par Gallager, pour $\rho > -1$, constituent une famille de mesures d'information bien enracinée qui apparaît dans divers problèmes d'exposants d'erreurs et aussi dans le décodage séquentiel. L'information mutuelle, qui détermine la limite théorique de la transmission de l'information, et le taux de coupure, un autre paramètre de canal qui a été interprété autrefois comme la 'limite pratique' de la transmission de l'information, se révèlent en tant que cas spéciaux de $E_0(\rho)/\rho$. En rétrospective, la découverte d'Arıkan est le résultat de sa recherche d'une méthode pour combler l'écart entre les deux limites mentionnées.

A cet égard, nous étudions dans le cadre de cette thèse l'évolution de cette famille de mesures d'information sous la transformée polaire. En particulier, nous démontrons que la transformée polaire améliore le paramètre $E_0(\rho)$ des canaux à entrées binaires. Le résultat nous permet de mieux comprendre la raison pour laquelle la transformée polaire donne des codes qui atteignent la capacité et qui sont de faible complexité : l'amélioration du paramètre $E_0(\rho)$ se traduit par une amélioration du compromis entre la complexité et la probabilité d'erreur. Il s'agit d'un concept introduit en 1996 par Forney. De plus, nous démontrons que même si la mesure d'information est changée de l'information mutuelle habituelle à $E_0(\rho)$, le canal binaire à effacement et le canal binaire symétrique restent toujours extrémaux. En parlant des canaux extrémaux, nous caractérisons aussi, indépendamment de tout contexte de polarisation, l'extrémalité de ces deux canaux parmi tous les canaux à entrées binaires ayant une valeur donnée de $E_0(\rho)$ pour une valeur fixe de ρ.

Après avoir développé une compréhension plus profonde de la technique de codage polaire, la thèse procède à l'étude d'un problème pratique relié à la conception

de codes polaires indiqué dans l'étude originale d'Arıkan : "la robustesse contre les variations des paramètres du canal". La résolution de ce problème est particulièrement difficile pour le codage polaire, parce que le développement initial a montré que la conception d'un code polaire est adaptée spécifiquement à la loi de distribution du canal de communication. Cependant, du point de vue de l'ingénieur, il est essentiel que les résultats d'une théorie soit robustes. Ainsi, juste après sa conception, les ordres partiels pour les canaux sont devenus pertinents pour la conception de codes polaires. Deux canaux sont ordonnés si le code destiné à l'un des canaux peut être transformé en un code qui donne au maximum la même probabilité d'erreur de décodage si utilisé sur l'autre canal. Une fois de plus, c'est Shannon qui a introduit dans une note le concept d'ordres partiels pour les canaux discrets sans mémoire. Dans cette thèse, nous traitons ce sujet en introduisant un cadre rigoureux dans lequel nous proposons d'étudier les ordres partiels pour les canaux de communication dans le contexte d'ordres stochastiques appelés ordres convexes. Plus précisément, nous découvrons un nouvel ordre partiel pour les canaux binaires discrets sans mémoire que nous appelons l'ordre convexe symétrique. Par la suite, la thèse examine les différents modèles de communication proposés dans la littérature pour construire des systèmes plus robustes. Des chapitres de cette thèse sont dédiés à étendre la théorie originale du codage polaire aux scénarios complexes suivants :

Codage avec une règle de décision donnée— Dans ce scénario, nous étudions la performance des décodeurs polaires désadaptés. Un décodeur polaire désadapté est un décodeur polaire à annulations successives qui utilise au cours de la procédure de décision la loi de distribution d'un canal désadapté au lieu de la loi du vrai canal. Nous définissons la capacité de transmission avec décodage polaire désadapté. De plus, nous montrons qu'il existe une famille de bornes inférieures à cette capacité et nous faisons la conjecture que la familles de bornes converge vers cette capacité lorsque la longueur du bloc devient grande. Donc, quand l'une de ces bornes est positive, des taux de communication strictement positifs peuvent être atteints avec des codes polaires appropriés. Nous observons également que la probabilité d'erreur de décodage du bloc décroît proportionnellement à l'exponentielle de la racine carrée de la longueur du bloc, comme précédemment. Il faut aussi souligner que, tout en étendant la théorie du codage polaire à des cas de communication avec des canaux désadaptés, le décodeur polaire désadapté préserve la même structure à faible complexité de l'ordre de $O(N \log N)$ que le décodeur polaire 'adapté'. Cette structure à faible complexité motive davantage le codage polaire en présence de décodage désadapté.

Communication sur une famille de canaux— Nous étudions également dans cette thèse la conception de codes polaires robustes sur une famille de canaux. Généralement, dans ce scénario, nous avons seulement accès à une connaissance partielle du vrai canal via la famille à laquelle il appartient. Donc, un code universel doit être conçu pour la famille de canaux. Le problème est abordé sous différents points de vue. D'abord, en permettant au décodeur (mais pas au codeur) de connaître

le vrai canal, nous relions l'ordre polaire à l'ordre convexe symétrique, le nouvel ordre partiel introduit par cette thèse. Ensuite, en permettant au codeur, au lieu du décodeur, d'avoir connaissance du vrai canal, nous étendons les résultats de la thèse sur la capacité de transmission avec décodage polaire désadapté en utilisant la notion de famille de canaux unilatéraux introduite par Abbé et Zheng. Prenant encore une autre approche, nous montrons que les codes polaires utilisant une méthode de calcul approximative au décodeur sont robustes pour la famille de canaux binaires symétriques. En combinant ce résultat avec des simulations, nous fournissons des preuves solides qui montrent que les codes polaires sont 'pratiquement universels' sur les canaux binaires symétriques. Enfin, nous démontrons que l'universalité peut être échangée contre la complexité. Nous montrons que plusieurs appels au décodeur polaire mettant en œuvre un test du rapport de vraisemblance généralisé donnent une règle de décodage universelle sur les canaux à entrées binaires qui satisfont certaines conditions. Par conséquent, il y a besoin de plus de ressources au niveau du décodeur pour atteindre l'universalité.

Communication sur les canaux non-stationnaires— Une autre contribution originale de cette thèse est l'extension de la théorie de la polarisation de canal aux canaux sans mémoire qui sont non-stationnaires. Ce modèle, qui ne suppose plus que le canal de communication est stationnaire durant la transmission de l'information, est très utile pour capter les effets des variations temporelles du bruit présent dans les systèmes de communication réels. Comme les techniques de preuve existantes ne sont pas applicables à ce scénario, nous reprouvons à nouveau le phénomène de polarisation pour le cas stationnaire en utilisant uniquement des méthodes élémentaires. Ensuite, en nous servant de la même méthode, nous montrons que la construction d'Arıkan polarise également les canaux non-stationnaires sans mémoire, de la même manière qu'elle polarise ceux qui sont stationnaires.

Mots clefs : Codage polaire, codes polaires, polarisation de canal, décodage désadapté, famille de canaux, conception de code robuste, test du rapport de vraisemblance généralisé (GLRT), codage pour canaux non-stationnaires, canaux extrémaux, E_0 de Gallager, exposants d'erreur, 'information combining'.

Contents

List of Figures

Basic Notations and Conventions

$:=$	equal by definition
$\square$	end of a proof
$\mathcal{A}, \mathcal{X}, \mathcal{Y}, \mathcal{W}$	sets
$\mathrm{cl}(\mathcal{W})$	closure of the set $\mathcal{W}$
$\overline{\mathcal{W}}$	convex closure of the set $\mathcal{W}$
$\mathbb{F}_2$	binary set $\mathbb{F}_2 := \{0, 1\}$
$x \in \mathcal{X}$	x is an element of the set $\mathcal{X}$
$\mathbf{x} := x_1^N$	sequence $(x_1, \ldots x_N)$ of elements of a set $\mathcal{X}$
$x_{1,o}^N, x_{1,e}^N$	odd and even indexed components of the sequence $(x_1, \ldots x_N)$
$\mathcal{X} \times \mathcal{Y}$	Cartesian product of the sets $\mathcal{X}$ and $\mathcal{Y}$
$\mathcal{X}^N$	N-th Cartesian power of the set $\mathcal{X}$
$\lvert\mathcal{A}\rvert := \#\mathcal{A}$	number of elements of the set $\mathcal{A}$
$\mathcal{A} \subset \mathcal{A}'$	$\mathcal{A}$ is a subset of $\mathcal{A}'$ (not necessarily proper)
$\mathcal{A}_N$	a subset of $\{1, \ldots, N\}$
$\mathcal{A}_N^c$	complement of the set $\mathcal{A}_N$, i.e., $\{1, \ldots, N\} \setminus \mathcal{A}_N$
$P(x)$	probability distribution over $x \in \mathcal{X}$
$(\Omega, \mathscr{F}, P)$	probability space: sample space Ω, sigma-field $\mathscr{F}$, probability P
$P_{\text{unif}}(x)$	uniform distribution over $x \in \mathcal{X}$
$P_{\text{tilt}}^{\alpha}(x)$	tilted probability distribution, i.e., $P_{\text{tilt}}^{\alpha}(x) := \dfrac{P(x)^{\alpha}}{\sum_{x \in \mathcal{X}} P(x)^{\alpha}}$

$P^N(\mathbf{x})$	N-th power of $P(x)$, i.e., $P^N(\mathbf{x}) := \prod_{i=1}^{N} P(x_i)$
A, U, X, Y, Δ	random variables
$\mathbf{1}\{\cdot\}$	indicator function of the set $\{\cdot\}$
$X \sim P(x)$	X is distributed according to $P(x)$
$\mathbb{P}[\,\{\cdot\}\,]$	probability of the set $\{\cdot\}$, the brackets might be dropped
$\mathbb{P}_Q[\,\{\cdot\}\,]$	probability of the set $\{\cdot\}$ weighted with respect to the distribution Q
$\mathbb{E}[\Delta]$	expectation of the random variable Δ
$\mathbb{E}_Q[.]$	expectation taken with respect to the distribution Q
$\lvert\,\Delta\,\rvert$	absolute value of the random variable Δ
$W\colon \mathcal{X} \to \mathcal{Y}$	channel mapping with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$
$W(y\lvert x)$	transition probabilities of $W\colon \mathcal{X} \to \mathcal{Y}$, for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$
$W^N(y_1^N\lvert x_1^N)$	N independent uses of W, i.e., $W^N(y_1^N\lvert x_1^N) := \prod_{i=1}^{N} W(y_i\lvert x_i)$
$PW(x, y)$	joint input-output distribution $P(x)W(y\lvert x)$ over $(x, y) \in \mathcal{X} \times \mathcal{Y}$
$PW(y)$	marginal output distribution, i.e., $PW(y) := \sum_{y \in \mathcal{Y}} PW(x, y)$
$\widehat{P}(x) := \widehat{P}_{x_1^N}(x)$	N-type (empirical distribution) of the sequence x_1^N over $x \in \mathcal{X}$
$\widehat{P}(y\lvert x)$	noise composition, i.e., conditional type $\widehat{P}(y\lvert x) := \widehat{P}_{y_1^N\lvert x_1^N}(y\lvert x)$
$\mathcal{P}^N(\mathcal{X})$	set of all N-types on $\mathcal{X}^N$
$\mathbb{R}, \mathbb{Z}, \mathbb{N}$	set of real, integer, and natural numbers
$\operatorname{sign}(m)$	sign function
$\lceil m \rceil$	smallest integer not less than m
$\lfloor m \rfloor$	largest integer not greater than m
$\max\{a, b\}$	maximum of the two numbers a and b
$\min\{a, b\}$	minimum of the two numbers a and b
$\lvert\, m \,\rvert^{\dagger}$	positive part of $m \in \mathbb{R}$, i.e., $\lvert\, m \,\rvert^{\dagger} := \max\{m, 0\}$

$\mathbf{u} \leq \mathbf{v}$	component-wise inequality of two sequences
$\mathbf{u} \vee \mathbf{v}$	component-wise maximum of two sequences
$\mathbf{u} \wedge \mathbf{v}$	component-wise minimum of two sequences
$\mathbf{u} \oplus \mathbf{v}$	component-wise modulo-2 sum of two sequences
$\exp_2, \ln, \log$	exponential with base 2, natural logarithm, logarithm to the base 2
$p * q$	$p * q := p(1-q) + q(1-p)$
$h_2(p)$	binary entropy function: $h_2(p) := -p \log p - (1-p)\log(1-p)$, $\forall p \in [0,1]$
$h_2^{-1}(.)$	inverse of the binary entropy function $h_2(p)$
$\binom{n}{k}$	binomial coefficient
$F^{\otimes N}$	N-th Kronecker product of the matrix F
$O(.)$	big O notation for order of complexity
$\mathrm{Div}(P\|\|Q)$	Kullback-Leibler divergence: $\mathrm{Div}(P\|\|Q) := \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}$
$\mathrm{Div}(W\|\|V\|P)$	conditional divergence: $\mathrm{Div}(W\|\|V\|P) := \mathbb{E}\left[\mathrm{Div}(W(y\|x)\|\|V(y\|x))\right]$
$\simeq$	asymptotically equal to

Acronyms

a.s.	*almost surely*
B-DMC	*binary discrete memoryless channel*
BEC	*binary erasure channel*
BSC	*binary symmetric channel*
DMC	*discrete memoryless channel*
i.i.d.	*independent and identically distributed*
GRLT	*generalized likelihood ratio test*
MMI	*maximum mutual information*

Chapter 1

Introduction

Suppose we would like to transmit a message to a destination over a noisy communication medium. My 6 years old nephew Batuhan and my 4 years old niece Yasemin would say write the message 10 times and send it by a dragon, a wagon, ultimate alien, our penguin friends (les manchots in french), ... and the post office. After all, the imaginary world is for free! We, on the other hand, have to be more realistic.

We start thinking and try to remember: What did we learn in Information Theory and Coding? A mathematical model of communication! Upon that, we decide to consider the communication system described in Figure 1.1. The communication medium is modeled as a 'probabilistic box' represented in the figure by a stationary *discrete memoryless channel* (DMC) $W\colon \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ denotes the input alphabet, $\mathcal{Y}$ the finite output alphabet, and $W(y|x)$ the channel transition probabilities for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. If we assign to the channel an input random variable $X \sim P(x)$, then ℓ consecutive uses of the channel will describe the input-output mapping $X_1^\ell \to Y_1^\ell$, for $X_1^\ell \in \mathcal{X}^\ell$ and $Y_1^\ell \in \mathcal{Y}^\ell$. The communication channel is *memoryless* in the sense that the conditional probabilities satisfy

$$\mathbb{P}[Y_i = y_i|Y_1 = y_1, \dots Y_{i-1} = y_{i-1}, X_1 = x_1, \dots X_i = x_i] = \mathbb{P}[Y_i = y_i|X_i = x_i],$$

for any $i = 1, 2, \dots$, and the channel is *stationary* due to the following property:

$$\mathbb{P}[Y_i = y|X_i = x] = W(y|x).$$

Figure 1.1: Communication over a DMC W.

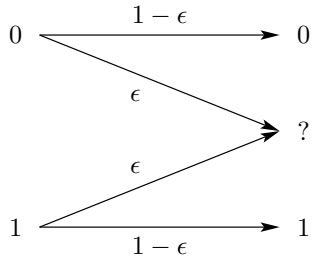

Figure 1.2: Binary erasure channel.

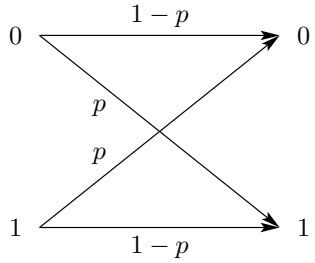

Figure 1.3: Binary symmetric channel.

This means that, at each time instant, the law of the DMC determines the behavior of the medium independently of the past. A popular channel model is the *binary erasure channel* (BEC). As shown in Figure 1.2, the BEC either transmits correctly the input bit or erases it. The erasure probability of the channel is usually denoted by $\epsilon \in [0, 1]$. Another popular model is the *binary symmetric channel* (BSC) shown in Figure 1.3. No erasure occurs this time, but each input can be flipped at the output. The chances of a flip happening is determined by the crossover probability of the BSC denoted by $p \in [0, 1]$.

The toy model of Figure 1.1 in our pocket, we set off for reliable communication. The Professor said that redundancy should be added to the data. But, how much? We suddenly recall that the channel capacity is the fundamental measure for reliable communication, and we decide to compute the *capacity of the channel* defined as [1]

$$C(W) = \max_{P(x)} I(X;Y),$$

where the maximization is over all possible channel input distributions $P(x)$, for $x \in \mathcal{X}$, and the hard to forget formula of the *mutual information* between $X \sim P(x)$ and $Y \sim PW(y)$ is

$$I(X;Y) = I(P;W) := \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} P(x)W(y|x) \log \left(\frac{W(y|x)}{\sum_{x'\in\mathcal{X}} P(x')W(y|x')} \right).$$

So, we will need to find the maximizing input distribution (a computer might be

helpful). Trial and error, imagine we did. To add the redundancy, we will design[1] the *encoder* which will map one of the 2^k *messages* identified by a k-bit data sequence u_1^k into a codeword $\mathcal{E}nc(u_1^k) = x_1^\ell$ of length $\ell(u_1^k)$. All the codeword mappings will compose the *codebook* $\mathcal{C}$. Some tools for designing the encoder and the codebook come to our minds: block codes (ℓ is the same for all the possible k-bit messages, and it would be wise to chose $\ell \geq k$ and $k/\ell \leq C(W)$.), typical sequences, Huffman codes, no, that was for source coding, random codes, linear codes, Reed Solomon codes, convolutional codes, ... Imagine we built an encoding scheme from our toolbox. Then, after we transmit the codeword and receive the output of the channel y_1^ℓ, we will have to decode the output with the *decoder* to produce an estimate $\hat{u}_1^k = \mathcal{D}ec(y_1^\ell)$ of the data. Let us list some options: optimal maximum likelihood decoding, minimum distance decoding, typicality decoding, or if we were lucky enough, we might have also heard about sequential decoding or iterative decoding. Which one to chose so that we recover the initial message with low error probability? Undecided, we realize it is time to make a good review of all these notions. Relax, due to space and time limitations, we will not do that here. Instead, let us catch the latest in modern coding trends: Polar coding.

1.1 Structural Components of Polar Coding

Just above, we have quite simplified the procedure; apologies to the great masters. In this section, we introduce the structural components of the polar coding scheme described in [2].

Channel Combining

Suppose that a binary sequence of random variables U_1^N is drawn *independent and identically distributed* (i.i.d.) from the *uniform distribution* P_{unif}, and another binary sequence X_1^N is generated by applying to U_1^N a one-to-one transformation $G_N \colon \{0,1\}^N \to \{0,1\}^N$. It can be easily inferred from this that the sequence $X_1^N = U_1^N G_N$ is also i.i.d. with uniform distribution.

Assume we make N independent channel uses of a *binary discrete memoryless channel* (B-DMC) $W : \mathbb{F}_2 \to \mathcal{Y}$ to transmit X_1^N and we obtain Y_1^N. Then, we have

$$W^N(y_1^N|x_1^N) = \prod_{i=1}^{N} W(y_i|x_i), \tag{1.1}$$

[1] At this point, the subject matter calls for duty the engineers.

by independence, and

$$W_N(y_1^N|u_1^N) := W^N(y_1^N|u_1^N G_N), \tag{1.2}$$

by the one-to-one transformation. Moreover, as $(X_1, Y_1), \ldots, (X_N, Y_N)$ are i.i.d. pairs of random variables, the combined channel satisfies

$$I(U_1^N; Y_1^N) = I(X_1^N; Y_1^N) = \sum_{i=1}^{N} I(X_i; Y_i) = N I(P_{\text{unif}}; W). \tag{1.3}$$

Channel Splitting

Once we combined the channels which map $X_i \to Y_i$, for $i = 1, \ldots, N$, we can re-split the combined channel describing $U_1^N \to Y_1^N$ by applying the chain rule to obtain

$$I(U_1^N; Y_1^N) = \sum_{i=1}^{N} I(U_i; Y_1^N \mid U_1^{i-1}) = \sum_{i=1}^{N} I(U_i; Y_1^N U_1^{i-1}), \tag{1.4}$$

where the second equality follows by the independence of U_1^{i-1} from U_i. Therefore, instead of the N independent channel uses of W we described in (1.3), we can consider from the chain rule perspective N successive uses of the *synthetic channels* mapping $U_i \to Y_1^N U_1^{i-1}$, for $i = 1, \ldots, N$, and we can write

$$N I(P_{\text{unif}}; W) = \sum_{i=1}^{N} I(P_{\text{unif}}; W_N^{(i)}),$$

where $W_N^{(1)}(y_1^N \mid u_1), \ldots, W_N^{(N)}(y_1^N u_1^{N-1} \mid u_N)$ denote the transition probabilities of these synthetic channels, respectively.

Channel Polarization

Following the splitting operation, channel polarization depicts an extremal case of the transition probabilities of the synthetic channels.

Definition 1.1. [2] *Channel polarization* is defined as the case when each of the channels $W_N^{(i)} \colon \mathbb{F}_2 \to \mathcal{Y}^N \times \mathbb{F}_2^{i-1}$, for $i = 1, \ldots, N$, converges either to a perfect channel or a completely noisy channel. Under this event, the empirical distribution of the mutual information terms in summation (1.4) satisfies

$$\frac{1}{N} \ \#\{i \in \{1, \ldots, N\} : \ I(U_i; Y_1^N U_1^{i-1}) \in (\gamma, 1-\gamma)\} \xrightarrow[N\to\infty]{} 0, \tag{1.5}$$

for any $\gamma \in (0, 1)$.

These two extremal situations are interesting for the following reason: In the case of an almost perfect channel with $I \approx 1$, no coding would be necessary and the data could be transmitted uncoded over the channel since no information loss occurs, and in the case of an almost completely noisy channel with $I \approx 0$, no coding scheme would be helpful since all the information would be lost during transmission and any design would be to no end. Polar codes exploit exactly this idea.

Low Complexity Polar Encoding

The described channel combining operation makes use of linear label maps. To reduce the complexity of this mapping, polar coding employs a special type of encoding structure to transform the inputs into codewords. The matrix G_N is generated in a recursive fashion from the following *basic polar transformation matrix*:

$$F_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{1.6}$$

The recursion is applied through the Kronecker products of this basic matrix by computing

$$F_N = F_2^{\otimes N},$$

where $F_2^{\otimes N}$ denotes the N-th Kronecker product of F_2. As a result, the recursion constraints the block-length to $N = 2^n$, where $n = 0, 1, \ldots$ is the number of times the recursion is applied. Finally, for systematic convenience during the decoding procedure, a permutation matrix B_N known as bit-reversal is applied to obtain

$$G_N = B_N F_N.$$

The main advantage of this special structure is that it can be implemented efficiently in $O(N \log N)$ time complexity [2, Theorem 5], while leading to channel polarization [2, Theorem 1]. The purpose of this section being to explore the 'infrastructure' of polar coding, we deliberately leave the explanation of the polarization argument to the beginning of Chapter 4.

Definition 1.1 implies that, after a long sequence of polar transformations, polarization divides the channels $W_N^{(1)}(y_1^N \mid u_1), \ldots, W_N^{(N)}(y_1^N u_1^{N-1} \mid u_N)$ into two clusters. Those having their mutual information close to 1 are classified as the good channels. The *information set* of polar coding, denoted by $\mathcal{A}_N \subset \{1, 2, \ldots, N\}$, refers to the indices of the channels in this category. The set of the remaining indices $\mathcal{A}_N^c := \{1, 2, \ldots, N\} \setminus \mathcal{A}_N$ is called the *frozen set*. As data can only be transmitted reliably over the indices in $\mathcal{A}_N$, to transmit a k-bit message, the encoder of polar codes has to (i) know or compute $\mathcal{A}_N$, and (ii) fix the value of u_i, for all $i \in \mathcal{A}_N^c$. With this knowledge, the encoder can take the initial k-bit message, embed this

data into the positions of the sequence u_1^N belonging to $\mathcal{A}_N$ (assuming N is large enough so that the size of the information set $|\mathcal{A}_N| \geq k$), and fill the remaining positions with the apriori fixed values. The codeword can then be found by computing $x_1^N = u_1^N G_N$.

Low Complexity Polar Code Construction

The description of the encoding operation has just revealed that a polar code can be constructed by specifying the information set $\mathcal{A}_N$ and fixing the inputs of the frozen set. Specifying $\mathcal{A}_N$, in turn, requires computing either directly the transition probabilities of the synthetic channels, or some information measures of these probabilities which would help to assess the quality of the channels. However, observe that as the block-length N grows large, the cardinalities of the output alphabets of the synthetic channels $W_N^{(i)} : \mathbb{F}_2 \to \mathcal{Y}^N \times \mathbb{F}_2^{i-1}$ grow exponentially large. Tracking the exact transition probabilities is therefore in general a high complexity operation. Nevertheless, an algorithm to compute efficiently the information sets of polar codes is proposed in [3]; the idea is to replace the exact computations with approximations. Currently, the best known complexity bound for the code construction is $O(N \log(2N))$. A discussion at the end of Chapter 6 will bring an enlightening look into how $\mathcal{A}_N$ can be efficiently computed via approximations.

Let us here illustrate the design principle over an exceptional channel model, the BEC, for which the exact computations can be carried out efficiently. An example which discusses the effects of channel combining and splitting over the BEC can also be found in [4, Example 3]. Let W be a BEC of erasure probability $\epsilon \in [0, 1]$. Suppose two copies of W are combined through the basic polar transformation matrix F_2 given in (1.6). After combining, the channel inputs are given by $X_1 = U_1 \oplus U_2$ and $X_2 = U_2$, and the splitting method synthesizes the channels mapping $U_1 \to Y_1 Y_2$ and $U_2 \to Y_1 Y_2 U_1$. We observe that $U_1 = X_1 \oplus X_2$ will be erased if and only if either channel inputs X_1 or X_2 are erased. Therefore, the channel describing $U_1 \to Y_1 Y_2$ is a BEC, and its erasure probability is given by $\epsilon^- = 2\epsilon - \epsilon^2$. Similarly, the channel describing $U_2 \to Y_1 Y_2 U_1$ is also a BEC, and its erasure probability is given by $\epsilon^+ = \epsilon^2$, as given the correct value of U_1, the input $U_2 = X_2$ will be erased if and only if both channel inputs are erased. It follows from this analysis that being a BEC is a property preserved by the polar transform. Hence, one can compute $\mathcal{A}_N$ efficiently in $O(N \log N)$ time by using the $\epsilon^{\pm}$ recursion of the BEC.

Low Complexity Polar Decoding

Due to the nature of the channel splitting operation, the synthetic channels lend themselves to a particular decoding procedure referred to as successive cancellation decoding. Let $L_W(y) := W(y|1)/W(y|0)$, for $y \in \mathcal{Y}$, denote the *likelihood ratio* of

the channel W. After u_1^N is encoded and transmitted, the *successive cancellation decoder* of polar codes will decode the received channel output sequence y_1^N using the following estimators:

$$\hat{u}_i = \begin{cases} u_i, & \text{if } \ i \in \mathcal{A}_N^c \\ f^{(i)}(y_1^N, \hat{u}_1^{i-1}), & \text{if } \ i \in \mathcal{A}_N \end{cases}, \tag{1.7}$$

where Arıkan chooses for $f^{(i)}(y_1^N, \hat{u}_1^{i-1})$ the *maximum likelihood* decoding rule for the i-th synthetic channel $W_N^{(i)}$ given by

$$f^{(i)}(y_1^N, \hat{u}_1^{i-1}) := \begin{cases} 0, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) < 1 \\ 1, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) > 1 \\ *, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) = 1 \end{cases}, \tag{1.8}$$

with $*$ chosen from the set $\{0, 1\}$ by a fair coin flip.

At the i-th stage of decoding, to estimate the input u_i of the channel with law $W_N^{(i)}(y_1^N u_1^{i-1} \mid u_i)$, the polar decoder should have, in principle, correctly estimated the inputs u_1^{i-1} of the previous stages. For otherwise, the decoder would be computing the likelihood ratio of the channel $L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1})$ for the wrong outputs. Though no genie exists to give the correct estimates, the analysis carried in [2] shows that this decoder performs with vanishing error probability. This result is illustrated by upper bounding the best achievable block error probability under successive cancellation decoding of polar coding with block-length N and information set $\mathcal{A}_N$ as follows:

$$\begin{aligned} P_{\text{e}}(W, \mathcal{A}_N) &= \mathbb{P}\left[\bigcup_{i \in \mathcal{A}_N} \left\{\hat{U}_1^{i-1} = u_1^{i-1}, \hat{U}_i \neq u_i\right\}\right] \\ &= \mathbb{P}\left[\bigcup_{i \in \mathcal{A}_N} \left\{\ \hat{U}_1^{i-1} = u_1^{i-1}, f^{(i)}(y_1^N, \hat{U}_1^{i-1}) \neq u_i\right\}\right] \\ &\leq \mathbb{P}\left[\bigcup_{i \in \mathcal{A}_N} \left\{f^{(i)}(y_1^N, u_1^{i-1}) \neq u_i\right\}\right] \leq \sum_{i \in \mathcal{A}_N} P_{\text{e, ML}}(W_N^{(i)}), \end{aligned}$$

where we define

$$\begin{aligned} P_{\text{e, ML}}(W) := &\sum_{\substack{y: \\ L_W(y) > 1}} W(y|0) + \frac{1}{2} \sum_{\substack{y: \\ L_W(y) = 1}} W(y|0) \\ &+ \sum_{\substack{y: \\ L_W(y) < 1}} W(y|1) + \frac{1}{2} \sum_{\substack{y: \\ L_W(y) = 1}} W(y|1). \end{aligned} \tag{1.9}$$

So, $P_{\text{e, ML}}(W_N^{(i)})$ is the error probability of the 'genie-aided' decoder for the i-th synthetic channel. The upper bound shows that as long as the synthetic channel at

hand is almost perfect, the contribution of this channel to the overall decoding error probability is almost negligible.

From the above description, it is not obvious that the polar decoder can be implemented with low complexity operations. In fact, it is once more the construction via the basic polar transformation matrix F_2 which leads to a low complexity recursive decoder implementation. The recursion implementing the polar decoder described above in $O(N \log N)$ complexity [2, Theorem 5] is given by [2, Eqs. (74) and (75)]:

$$L_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}) = \frac{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) + L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2})}{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}) + 1}, \quad (1.10)$$

and

$$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) = \left[L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})\right]^{1-2\hat{u}_{2i-1}} \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}), \quad (1.11)$$

where the notation $L_N^{(i)}(y_1^N, \hat{u}_1^{i-1})$ refers to $L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1})$, and the '$o$' and '$e$' subscripts to odd and even indexed components of the sequences, respectively.

1.2 State of the Art

Reliable communication is one of the major problems considered by the field of information and coding theory. Towards this objective, channel codes are designed where channel encoders construct redundant channel input sequences, such that upon reception, the corresponding channel decoders can overcome the effects of noise introduced during transmission and recover the transmitted message. An information rate is achievable over the channel if no matter how small the decoding error probability is, we can find a code with a sufficiently large block-length of that rate. Shannon defined the channel capacity as the highest achievable rate that information can be sent through the channel. The error probability analysis over random code ensembles used with typicality decoders, tailored to jointly typical sequences, provides a simple way to show that the mutual information between the channel input and output is a fundamental quantity in determining the channel capacity [1].

The *reliability function* $E(R)$ provides a finer measure on the quality of the channel: for any R less than the channel capacity, it is possible to find a sequence of codes of increasing block-length, each of which of rate at least R, and whose block decoding error probability decays exponentially to zero as the block-length increases — $E(R)$ is the largest possible rate of this decay. Using a random ensemble with maximum likelihood decoding, Gallager's classical trea-

tise [5] gives a lower bound to $E(R)$, the random coding exponent $E_r(R)$ in the form $E_r(R) = \max_{\rho\in[0,1]} [E_0(\rho) - \rho R]$, where $E_0(\rho)$ is known as 'Gallager's E_0'[2]. Remarkably, this lower bound is tight for rates above the critical rate $E_0'(1)$.

A particular aspect of the methods to prove the above outcomes is that they are based on averaging over code ensembles. We know there are codes which are good with respect to the achievability criteria. However, no real system exists without cost. This last one is dictated by the storage capabilities and the computational complexity limitations in the case of communication systems. Consider a block code with exponentially many codewords in the block-length. Quickly we find ourselves out of memory for small values of the block-length, and of making a prohibitive number of pairwise comparisons to identify the correct message sent through the channel.

Coding theorists have spent decades to find good codes with low encoding and decoding complexities. A historical account on the evolution of channel coding can be found in [6]. Let us highlight here a few developments of interest to polar coding. In the previous section, we observed that polar codes are linear: The encoder transforms the input message u_1^N to a codeword $x_1^N = u_1^N G_N$ using a generator matrix. As a class of codes having an efficient structure for the codeword generation process, the class of linear codes gives a partial solution to the low complexity coding problem. What's more is that the properties of this class of codes established linearity as a prerequisite rather than an option. Quoting as phrased in [6]: "In practice, practically all codes are linear". On the other hand, to help reduce the complexity at the decoder side, a line of research focused on devising sequential decoding algorithms, such as Fano's algorithm [7] suitable for decoding convolutional codes invented by Elias [8]. The theoretical analysis of sequential decoding showed that the function $E_0(\rho)$ that appears as an auxiliary function on the road to deriving $E_r(R)$ turns out to be of independent interest on its own right. In particular, $E_0(\rho)/\rho$ is the largest rate for which a sequential decoder can operate while keeping the ρ-th moment of the decoder's computation effort per symbol bounded [9].

In 2006, Arıkan [4] proposed channel combining and splitting as a method to improve $E_0(1)$, the cutoff rate of the channel, using multiple sequential decoders in a successive cancellation configuration. Afterwards, polar codes emerged in his seminal paper [2] as an appealing error correction method based on channel combining and splitting. Although algebraic coding techniques and probabilistic coding techniques such as convolutional codes have found applications in the real world before polar codes were invented, none of these techniques are proved to achieve the symmetric capacity with low complexity and with an explicit construction.

[2]The functions $E_0(\rho)$ and $E_r(R)$ are in fact parameters which depend on the channel. They are defined rigorously in Chapter 2, see (2.1) and (2.9).

Beside their computational complexity, typicality decoders and maximum likelihood decoders present another obstacle: they require the exact channel knowledge to function. To study the situations in which such a complete knowledge is missing, the study of reliable communication under channel uncertainty becomes relevant. The qualitative notion of channel uncertainty is made more concrete by the definition of more complex channel models. For instance, in the compound channel model, the unknown channel is restricted to belong to a given class of channels, and in this case, the selected code should ensure good performance for all the channels in the class. Blackwell et. al. [10] studied this problem and defined the capacity of a class of channels. More details on this topic are discussed in Section 8.1. Another more complex model is the *arbitrarily varying channel* model [11]. Here the channel law is allowed to vary over each use of the channel, notably in a malicious fashion. In addition to partial or missing channel information, the use of sub-optimal decoders might also arise due to practical implementation constraints. Although the use of such decoders may result in capacity loss, this is in general unavoidable. The paper [12] focus on additive decoding rules called d-decoders and analyze the transmission capacity with a given d-decoder. More details on this topic are discussed in Sections 7.1 and 7.4.1.

The need to communicate reliably without using the channel knowledge has led to the concept of universal coding. Using the method of types, Csiszár [13] proved that universally attainable transmission rates can be obtained for any discrete memoryless channel with finite input and output alphabets by using the *maximum mutual information* (MMI) decoder. This is a decoder which declares an input message if and only if the corresponding codeword maximizes the empirical mutual information computed from the joint type (empirical distribution) of the codeword and the received channel output sequence. The decoder is, unfortunately, too costly to be implementable in practice. Hence from a practical perspective, universal coding with low complexity decoders is the 'next' problem in line which needs to be undertaken.

1.3 Thesis Outline

The first two chapters of this thesis are the only ones whose contents are independent from any polarization context. Chapter 2 is a preliminary which introduces Gallager's E_0 and derives some basic results which will be used later, especially in the subsequent two chapters. Chapter 3 describes certain extremalities for B-DMCs when the information measure is E_0. We show that the BEC and the BSC are the extremal channels of the considered problems. As Chapter 3 is independent of the polarization context, it can also be read at last.

Chapter 4 starts with a quick overview of the main channel polarization result concerning the evolution of the synthetic channels' symmetric capacity parame-

ters under the polar transform. Three properties —polarization, conservation, and extremality— and the 'Conservation and Convergence Laws' they imply are particularly emphasized. From that chapter onward, these properties and laws form the leitmotiv of this thesis. While the problems we tackle change, the answers we are searching for tell variations of the recurrent theme. For instance Chapter 5 proves, in the same spirit, the properties and laws for the variational distance between the transition probabilities $W(y|0)$ and $W(y|1)$ of symmetric B-DMCs.

Chapter 6 introduces a new partial ordering for B-DMCs called the symmetric convex ordering. Convex orderings[3] are stochastic orders which are often encountered in statistics and actuarial sciences. Using this framework, the chapter shows that the polar transform preserves symmetric convex orderings, and as a consequence, the information sets of symmetric convex ordered channels are also ordered.

Starting from Chapter 7, the thesis begins to examine the robustness of polar coding against various perturbations of the original model. In Chapter 7, the decision rule used by the polar decoder, i.e., $f^{(i)}$ given in (1.8), is no longer assumed to be necessarily matched to the true channel. The chapter investigates the transmission capacity of polar coding with mismatched polar decoding.

In Chapter 8, the problem of designing universal polar codes over a set of compound B-DMCs is studied. The goal of this chapter is to identify the conditions a class of channels should satisfy so that a specific information set (preferably as large as possible) and polar decoder design can be used over all the channels in the class without degrading performance.

Subsequently, Chapter 9 removes the stationarity assumption from the probabilistic box modeling the communication medium, and instead, assumes the medium is modeled with a non-stationary memoryless channel. A polar coding theorem is proved, one last time in the thesis, for non-stationary B-DMCs with the help of a new proof technique. Chapter 10, the final chapter of this thesis, gives an overview of the contributions of this thesis and identifies future directions for research.

An important remark is in order before we close this introduction. Most of the results presented in this thesis are derived for a slightly more general polar transform which synthesizes two channels by using the exact same kernel F_2 given in (1.6) but in order to combine two independent but not necessarily identical B-DMCs. The transform will be defined more formally in Chapter 4.

[3]We use the term *convex orderings* to refer to any stochastic order based on second order properties such as the convex ordering, the increasing convex ordering, etc.

Chapter 2

A General Measure of Information

Given a DMC $W\colon \mathcal{X} \to \mathcal{Y}$, fix a distribution Q on its input alphabet. Let

$$E_0(\rho) = E_0(\rho, Q, W) := -\log \sum_{y\in\mathcal{Y}} \left[\sum_{x\in\mathcal{X}} Q(x)W(y \mid x)^{\frac{1}{1+\rho}}\right]^{1+\rho}, \tag{2.1}$$

for $\rho > -1$.

What's Coming, Doc?

In this chapter, we will study the above quantity, *'Gallager's E_0'*, review some of its basic properties, and derive some facts related to the E_0 function of B-DMCs evaluated under the uniform input distribution. (The choice of the uniform input distribution will be advocated.) In between, we will also mention some applications in information theory involving this parameter. The reader should not, however, expect a complete tutorial coverage of the subject. The included material is selected primarily so as to provide the necessary background for Chapter 3 and Chapter 4 of this thesis. Yet, we also hope that some of the derivations might be useful beyond the context of this thesis.

2.1 All roads lead to E_0

Arimoto [14] introduced an alternative description of $E_0(\rho, Q, W)/\rho$ using the concept of Rényi's entropy functions. This description, which is also mentioned in [9] and [15], gives an interpretation to $E_0(\rho, Q, W)/\rho$ as a general measure of information. We start by exploring this connection.

The *Rényi's entropy function of order* α of a discrete random variable $X \sim P(x)$ is defined in [16] as

$$H_\alpha(X) := \frac{\alpha}{1-\alpha} \log \left(\sum_{x \in \mathcal{X}} P(x)^\alpha \right)^{\frac{1}{\alpha}},$$

for $\alpha \geq 0$, $\alpha \neq 1$. This definition is extended to the *Rényi's conditional entropy function of order* α of a discrete random variable X given Y with joint distribution $P(x)W(y|x)$ in [14] as[1]

$$\begin{aligned} H_\alpha(X \mid Y) &:= \frac{\alpha}{1-\alpha} \log \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P(x)^\alpha W(y|x)^\alpha \right)^{\frac{1}{\alpha}} \\ &= H_\alpha(X) + \frac{\alpha}{1-\alpha} \log \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P_{\text{tilt}}^\alpha(x) W(y \mid x)^\alpha \right)^{\frac{1}{\alpha}}, \end{aligned} \tag{2.2}$$

where $P_{\text{tilt}}^\alpha(x) := \dfrac{P(x)^\alpha}{\sum\limits_x P(x)^\alpha}$ is called the *tilted probability distribution*. Note that

$$\begin{aligned} \lim_{\alpha \to 1} H_\alpha(X) = H(X) = H(P) := \sum_{x \in \mathcal{X}} -P(x) \log P(x), \\ \lim_{\alpha \to 1} H_\alpha(X|Y) = H(X|Y) = H(X,Y) - H(Y), \end{aligned}$$

where $H(X,Y)$ denotes the joint entropy of the pair (X,Y). So, Shannon entropy is a special case of Rényi entropy.

Letting $\alpha = \dfrac{1}{1+\rho}$, we get

$$H_{\frac{1}{1+\rho}}(X) = \frac{1}{\rho} \log \left(\sum_{x \in \mathcal{X}} P(x)^{\frac{1}{1+\rho}} \right)^{\frac{1}{1+\rho}},$$

$$H_{\frac{1}{1+\rho}}(X \mid Y) = H_{\frac{1}{1+\rho}}(X) + \frac{1}{\rho} \log \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P_{\text{tilt}}^\alpha(x) W(y \mid x)^{\frac{1}{1+\rho}} \right)^{1+\rho}.$$

Taking $Q = P_{\text{tilt}}^\alpha$ in the definition of $E_0(\rho, Q, W)$ in (2.1), we deduce

$$\frac{E_0(\rho, Q, W)}{\rho} = H_{\frac{1}{1+\rho}}(X) - H_{\frac{1}{1+\rho}}(X \mid Y). \tag{2.3}$$

[1] Although different definitions are proposed in the literature for a possible extension of Rényi's entropy function to a quantity similar to the conditional entropy function, as one suitable for the context of this thesis, we adopt the definition in (2.2).

This quantity is called the *mutual information of order* $\alpha = \frac{1}{1+\rho}$ in [14], and the following properties are observed:

(i) For a fixed Q, $E_0(\rho, Q, W)/\rho$ is a decreasing function in ρ.

(ii) The mutual information $I(Q;W)$ is given by

$$\lim_{\rho \to 0} \frac{E_0(\rho, Q, W)}{\rho} = I(Q;W). \tag{2.4}$$

In fact, it is the slope of the E_0 curve at $\rho = 0$ [5, Figure 5.6.2], i.e.,

$$I(Q;W) = \frac{\partial}{\partial \rho} E_0(\rho, Q, W)\Big|_{\rho=0}.$$

Finally, in analogue to the definition of the channel capacity, the maximization of (2.3) over all $Q(x)$ yields the *capacity of order* $\alpha = \frac{1}{1+\rho}$, for $\alpha \in [1/2, \infty)$.

2.1.1 Error/Guessing Exponents

In the context of channel coding, E_0 appears as a useful system parameter in various error exponent problems and also in sequential decoding. Alongside with the block-length N and the rate R, E_0 emerges in these coding problems in the exponent of the derived bounds to the expected performance, establishing a trade-off between the various performance measures. Below, we illustrate some of these applications:

A.1 Originally, E_0 was defined by Gallager in [17, Theorem 1]. The theorem states that for a DMC W and a fixed rate $R > 0$, the average block decoding error probability $P_{\text{e, avg}}$ using maximum likelihood decoding over the ensemble of random block codes with codewords of length N and distribution Q on them can be upper bounded by

$$P_{\text{e, avg}} \leq \exp_2 \left\{ -N \left[E_0(\rho, Q, W) - \rho R \right] \right\},$$

for any $\rho \in [0, 1]$. Here, the exponent establishes the compromise between $P_{\text{e, avg}}$, N, and R. The tightest exponent in the above upper bound is called the *random coding exponent* and is given by [5, Equation 5.6.16]

$$E_r(R, W) := \max_{\rho \in [0,1]} \max_{Q} \left[E_0(\rho, Q, W) - \rho R \right], \tag{2.5}$$

for $R \geq 0$.

The properties of $E_0(\rho, Q, W)$ with respect to the variable ρ are summarized in [5, Theorem 5.6.3]. For $\rho \geq 0$, $E_0(\rho, Q, W)$ is a positive concave increasing

function in ρ. As a result, for a fixed input distribution Q, the maximization in

$$\max_{\rho\in[0,1]} [E_0(\rho, Q, W) - \rho R]$$

can be described in terms of the following *parametric equations*:

$$R(\rho, Q, W) = \frac{\partial}{\partial\rho} E_0(\rho, Q, W), \tag{2.6}$$

$$E_r(\rho, Q, W) = E_0(\rho, Q, W) - \rho\frac{\partial}{\partial\rho} E_0(\rho, Q, W), \tag{2.7}$$

for R in the range $R(1, Q, W) \leq R \leq R(0, Q, W)$. The *noisy channel coding theorem* [5, Theorem 5.6.4] shows that for any DMC W, $E_r(R, W)$ is a *positive* convex decreasing function of R, for $R < C(W)$.

A.2 Arimoto [18] showed that for any code with block-length N, the block decoding error probability P_{e} under maximum likelihood decoding is lower bounded by

$$P_{\text{e}} \geq 1 - \exp_2\left\{-N \min_Q [E_0(\rho, Q, W) - \rho R]\right\},$$

for any $\rho \in (-1, 0]$. The *strong converse exponent* is defined as [18]

$$E_{sc}(R, W) := \max_{\rho\in(-1,0]} \min_Q [E_0(\rho, Q, W) - \rho R]. \tag{2.8}$$

It is shown in [18, Theorem 2] that if $R > C(W)$, $E_{sc}(R, W)$ is a *positive* increasing convex function in R.

A.3 Suppose that instead of the maximum likelihood decoder a *list decoder* is used in the decoding procedure to produce the list of the L most likely codewords for a given output sequence. In this case, the ensemble average probability of list decoding error $P_{\text{e, avg},L}$ can be upper bounded by [5, Problem 5.20]

$$P_{\text{e, avg},L} \leq \exp_2\left\{-N \max_{\rho\in[0,L]} \max_Q [E_0(\rho, Q, W) - \rho R]\right\}.$$

A.4 Assume an input sequence X is transmitted through a channel W and the output sequence Y is received. A *sequential decoder* can be described as a device which, based on the received value, keeps guessing which particular input was transmitted until the correct decision is made. The number of guesses made during this procedure and the number of computations performed by the decoder are parallel quantities. They both depend on the order in which the guesses are made. This order, in turn, is determined by a function $G(x \mid y)$ called the *guessing function*. Thus, the computational complexity of sequential decoding can be expressed in terms of the random variable $G(X \mid Y)$. In [19],

Massey observes that the average number of guesses is minimized by guessing the value of the random variable X given Y in a decreasing order of the conditional probabilities. In [9], Arıkan considers sequential decoders which have arbitrary guessing functions and derives the following lower bound to the moments of computational complexity of sequential decoding:

$$\mathbb{E}[G(X \mid Y)^\rho] \geq (1 + NR)^{-\rho} \exp_2\{N\left[\rho R - E_0(\rho, W)\right]\},$$

for any $\rho \geq 0$. Here, a trade-off between the rate, the block-length, and the computational complexity is established. The critical value $\max_Q E_0(\rho, Q, W)/\rho$, called the *cutoff rate for the ρ-th moment*, imposes a limit on the rate R. Above this limit, as the block-length N is increased, the ρ-th moment of computation tends to infinity. In particular, above the *cutoff rate* of the channel given by $\max_Q E_0(1, Q, W)$, the sequential decoder needs to perform infinitely many computations, so complexity is unbounded. Conversely, when $R < \max_Q E_0(\rho, Q, W)/\rho$, there is a "tree code" for which the sequential decoder will make a bounded number of computations per decoded bit.

2.1.2 The Uniform Input Distribution

Definition 2.1. A DMC W is a *symmetric channel* if for some *permutation* π on the output alphabet $\mathcal{Y}$ satisfying $\pi = \pi^{-1}$, we have $W(y|1) = W(\pi(y)|0)$ for all $y \in \mathcal{Y}$.

It is well known that the uniform input distribution P_{unif} maximizes (2.5) for any $\rho \in [0, 1]$ when the channel W is symmetric [5, Chapter 5]. The random coding exponent of a symmetric channel is then given by

$$E_r(R, W) = \max_{\rho \in [0,1]} [E_0(\rho, P_{\text{unif}}, W) - \rho R]. \tag{2.9}$$

Moreover, the right hand side of (2.9) is always a lower bound to the random coding exponent of any asymmetric channel W. Thus, for an ensemble of block codes using the maximizing input distribution Q, regardless of channel symmetry, the bound

$$P_{\text{e, avg}} \leq \exp_2\left\{-N\left[E_0(\rho, P_{\text{unif}}, W) - \rho R\right]\right\}$$

holds for any $\rho \in [0, 1]$. Similarly, the uniform input distribution minimizes (2.8) for $\rho \in (-1, 0]$ when the channel is symmetric. The strong converse exponent of a symmetric channel is then given by

$$E_{sc}(R, W) = \max_{\rho \in (-1,0]} [E_0(\rho, P_{\text{unif}}, W) - \rho R]. \tag{2.10}$$

However, when the channel is not symmetric, we cannot relax the bound in (2.8) using (2.10).

Throughout this thesis, we will be concerned with B-DMCs and fix the input distribution to the uniform probability assignment. Beside being optimal for symmetric channels, this choice of the distribution will effectively make the problems more tractable. Yet, this is not the only justification. It turns out that no loss is incurred by this restriction in the today's practically relevant coding systems which are all based on linear coding schemes. To support this claim, we refer to an exercise in [5, Problem 6.3 (b)]. The problem shows that all binary linear codebooks can be generated by drawing each letter of the codewords independently from the uniform distribution. Thus for any B-DMC W, $E_0(\rho, P_{\text{unif}}, W)/\rho$ is the practically relevant information measure for analyzing the performance of practical codes.

2.2 E_0 and E_0' of B-DMCs

From now on, we assume $\mathcal{X} = \mathbb{F}_2$ and fix Q to P_{unif}. Then, (2.1) becomes

$$E_0(\rho, W) := E_0(\rho, P_{\text{unif}}, W) = -\log \sum_{y \in \mathcal{Y}} \left[\frac{1}{2} W(y \mid 0)^{\frac{1}{1+\rho}} + \frac{1}{2} W(y \mid 1)^{\frac{1}{1+\rho}} \right]^{1+\rho}. \tag{2.11}$$

The *symmetric capacity* of the channel — the mutual information evaluated at the uniform input distribution — and the *symmetric cutoff rate* will be denoted by $I(W)$ and $R_0(W) := E_0(1, W)$, respectively.

The proofs of the extremality results we will carry in the following two chapters will rely neither on the 'raw definition' of $E_0(\rho, W)$ in (2.11), nor on the interpretation in terms of Renyi's entropy functions in (2.3). Instead, we will make use of a description of $E_0(\rho, W)$ introduced in [20] which is more suitable for deriving extremal bounds.

For a given B-DMC $W : \mathbb{F}_2 \to \mathcal{Y}$, [20] shows that there exists a random variable Z taking values in the $[0, 1]$ interval such that

$$E_0(\rho, W) = -\log \mathbb{E}\left[g(\rho, Z)\right], \tag{2.12}$$

for any $\rho > -1$, where the function $g(\rho, z)$ is defined as

$$g(\rho, z) := \left(\frac{1}{2}(1+z)^{\frac{1}{1+\rho}} + \frac{1}{2}(1-z)^{\frac{1}{1+\rho}} \right)^{1+\rho}, \tag{2.13}$$

for $\rho > -1$ and $z \in [-1, 1]$.

To see this, define

$$q_W(y) := \frac{W(y \mid 0) + W(y \mid 1)}{2}, \tag{2.14}$$

for $y \in \mathcal{Y}$, and

$$\Delta_W(y) := \frac{W(y \mid 0) - W(y \mid 1)}{W(y \mid 0) + W(y \mid 1)}, \tag{2.15}$$

so that $W(y \mid 0) = q_W(y)\,(1 + \Delta_W(y))$ and $W(y \mid 1) = q_W(y)\,(1 - \Delta_W(y))$. Then, one can manipulate (2.11) to find that $Z := |\Delta_W(Y)|$ with $Y \sim q_W(y)$ satisfies (2.12)[2].

The next lemma gives the first and the second order properties of $g(\rho, z)$ with respect to the variable z. The proof is carried in Appendix 2.A.

Lemma 2.2. *The function $g(\rho, z)$ defined in* (2.13) *is a concave non-increasing function in $z \in [0, 1]$ for $\rho \in [0, \infty)$ and a convex non-decreasing function in $z \in [0, 1]$ for $\rho \in (-1, 0]$. As $g(\rho, z)$ is symmetric around $z = 0$, these properties also determine the function's behavior for $z \in [-1, 0]$.*

We denote by $g^{-1}(\rho, t)$ the inverse of the function $g(\rho, z)$ with respect to its second argument. The variable t always takes values from a subset of the interval $[0, 2]$. More specifically, $t \in [2^{-\rho}, 1]$ when $\rho \geq 0$ and $t \in [1, 2^{-\rho}]$ when $\rho \in (-1, 0]$. For shorthand notation, we denote the range of possible values by $t \in [2^{-\rho}, 1] \cup [1, 2^{-\rho}]$, for $\rho > -1$.

We note that by using (2.12), the function $R(\rho, W) := E_0' = \dfrac{\partial}{\partial \rho} E_0(\rho, W)$ which appears in the parametric form of the random coding exponent we stated in (2.6) can be written as

$$R(\rho, W) = \frac{-\partial \mathbb{E}\left[g(\rho, Z)\right] / \partial \rho}{\mathbb{E}\left[g(\rho, Z)\right] \ln 2} = \frac{\mathbb{E}\left[-\partial g(\rho, Z)/\partial \rho\right]}{\mathbb{E}\left[g(\rho, Z)\right] \ln 2}, \tag{2.16}$$

where the exchange of the partial derivative and the expectation operators in the second equality follows by the dominated convergence theorem.

2.2.1 Fun Facts About E_0 and E_0' of BECs and BSCs

In this section, we explain some simple facts related to the E_0 curves of BECs and BSCs. We will be using some of these facts multiple times throughout the next two chapters.

Consider first the representation in (2.12). It is not difficult to see that the BECs and the BSCs are special cases of this representation.

Fact 1. [20] The random variable Z_{BEC} of a BEC is $\{0, 1\}$ valued and satisfies $\mathbb{P}[Z_{BEC} = 0] = \epsilon$, where $\epsilon \in [0, 1]$ is the erasure probability of the channel. The

[2]Do not confuse the random variable Z with the channel Bhattacharyya parameter denoted as $Z(W)$! For the definition of the latter, see (4.8).

random variable Z_{BSC} of a BSC is a constant given by $z_{BSC} = |1 - 2p|$, where $p \in [0, 1]$ is the crossover probability of the channel.

It is well known that the set of BECs and the set of BSCs with crossover probabilities in $[0, 0.5]$ are ordered in terms of their channel capacities: if the chances of an erasure to happen at the output of a BEC model, or similarly of a bit flip at the output of a BSC model is increased, the transmission capacities shall decrease, see for instance the textbook [5]. Intuitively, we expect this graceful degradation to order as well other measures of channel quality. For that purpose, we start by computing the E_0 and E_0' parameters of a BEC and a BSC as a function of the erasure probability and the crossover probability of the channels. Let BEC be a BEC with erasure probability $\epsilon \in [0, 1]$. Then, one can easily derive that

$$E_0(\rho, BEC) = -\log\left(2^{-\rho}(1-\epsilon) + \epsilon\right), \tag{2.17}$$

and

$$R(\rho, BEC) = \frac{\partial}{\partial\rho} E_0(\rho, BEC) = \frac{2^{-\rho}(1-\epsilon)}{2^{-\rho}(1-\epsilon) + \epsilon}. \tag{2.18}$$

Let BSC be a BSC with crossover probability $p \in [0, 0.5]$. In this case, we are saved from the trouble by [5, Example 1 p.146] which has the derivation of the E_0 parameter of a BSC in Equation (5.6.40) and its rate parameter in Equation (5.6.41). Rewriting these equations, we get

$$E_0(\rho, BSC) = \rho - (1+\rho)\log\left(p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right), \tag{2.19}$$

and

$$R(\rho, BSC) = 1 - h_2(\delta), \tag{2.20}$$

where $\delta = \dfrac{p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}}$.

Now, we show that these parameters are monotone functions in the erasure and crossover probabilities of the channels.

Lemma 2.3.

(i) For any $\rho \geq 0$, $E_0(\rho, BEC)$ $[E_0(\rho, BSC)]$ is decreasing in ϵ $[p]$.

(ii) For any $\rho \in (-1, 0]$, $E_0(\rho, BEC)$ $[E_0(\rho, BSC)]$ is increasing in ϵ $[p]$.

(iii) For any $\rho > -1$, $R(\rho, BEC)$ $[R(\rho, BSC)]$ is decreasing in ϵ $[p]$.

Proof. Taking the first derivative of (2.17) with respect to ϵ, we get

$$\frac{\partial}{\partial\epsilon} E_0(\rho, BEC) = -\frac{1 - 2^{-\rho}}{(2^{-\rho}(1-\epsilon) + \epsilon)\ln 2}.$$

One can check that

$$\frac{\partial}{\partial \epsilon} E_0(\rho, BEC) \begin{cases} > 0, & \text{for } \rho \in (-1, 0) \\ = 0, & \text{for } \rho = 0 \\ < 0 & \text{for } \rho > 0 \end{cases}.$$

As $E_0(0, W) = 0$, the E_0 curves of all BECs will be ordered such that while for $\rho > 0$ the E_0 curves of BECs with smaller erasure probabilities will be larger, the opposite will be true for $\rho \in (-1, 0)$.

Next, we show an ordering also holds for the rate parameters of BECs. Taking the first derivative of (2.18) with respect to ϵ, we get

$$\frac{\partial}{\partial \epsilon} R(\rho, BEC) = -\frac{2^\rho}{(1 + (-1 + 2^\rho)\epsilon)^2} < 0.$$

Hence, the rate parameters will be decreasing with the erasure probability of the channel for any $\rho > -1$. This completes the proof for the BEC.

Let us prove the claims for the set of BSCs. First, we note that the term inside the logarithm in (2.19) satisfies

$$\frac{\partial}{\partial p}\left(p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right) = \frac{p^{-\frac{\rho}{1+\rho}} - (1-p)^{-\frac{\rho}{1+\rho}}}{1+\rho} = \begin{cases} < 0, & \text{for } \rho \in (-1, 0) \\ = 1, & \text{for } \rho = 0 \\ > 0 & \text{for } \rho > 0 \end{cases},$$

for $p \in [0, 0.5]$. Hence, we also have

$$\frac{\partial}{\partial p} E_0(\rho, BSC) \begin{cases} > 0, & \text{for } \rho \in (-1, 0) \\ = 0, & \text{for } \rho = 0 \\ < 0 & \text{for } \rho > 0 \end{cases},$$

which proves the ordering for $E_0(\rho, BSC)$. To prove the claim for $R(\rho, BSC)$, we simply note that in (2.20), for $p \in [0, 0.5]$, we have $\delta \in [0, 0.5]$ increasing in p and the binary entropy function $h_2(\delta)$ increasing in $\delta \in [0, 0.5]$. Thus, as claimed

$$\frac{\partial}{\partial p} R(\rho, BSC) < 0. \qquad \square$$

By this lemma, the second fact is in order:

Fact 2. For any $\rho > -1$, the class of BECs and the class of BSCs ($p \in [0, 0.5]$) are strictly ordered in their $E_0(\rho, W)$ parameters, except at $\rho = 0$ where $E_0(0, W) = 0$, and in their $R(\rho, W)$ parameters.

The ordering we have just discussed is not peculiar to BECs and BSCs and can be generalized to more general classes of channels. Lemma 2.7 in Appendix 2.B shows that the E_0 parameters of any two stochastically degraded DMCs are as well ordered for any choice of the input distribution. Lemma 2.3 will, however, be sufficient for our purposes as the derivations we carry later will not need results of such a generality.

Next, we argue the validity of an assumption we will encounter in the hypothesis of the main theorem.

Lemma 2.4. *For any given B-DMC W and any fixed $\rho > -1$, there exist a BEC BEC and a BSC BSC such that*

$$E_0(\rho, W) = E_0(\rho, BEC) = E_0(\rho, BSC). \tag{2.21}$$

The erasure probability of the channel BEC and the crossover probability of the channel BSC depend both on the channel W and the parameter ρ.

Proof. Observe that, by (2.12), the equality of the E_0 functions in (2.21) is equivalent to the equality of

$$\mathbb{E}\left[g(\rho, Z)\right] = \mathbb{E}\left[g(\rho, Z_{BEC})\right] = g(\rho, z_{BSC}), \tag{2.22}$$

where Z, Z_{BEC} and z_{BSC} correspond to the 'Z' random variables of the channel W, the channel BEC, and the channel BSC, respectively. Therefore, to show that there exist a BSC and a BEC satisfying (2.21), it is sufficient to show that there exist Z_{BEC} and z_{BSC} random variables satisfying (2.22).

By the monotonicity results stated in Lemma 2.2, we know that

$$\begin{aligned} g(\rho, z) &\in [2^{-\rho}, 1], \quad \text{for } \rho \geq 0, \\ g(\rho, z) &\in [1, 2^{-\rho}], \quad \text{for } \rho \in (-1, 0], \end{aligned}$$

for $z \in [0, 1]$. As a result,

$$\begin{aligned} E[g(\rho, Z)] &\in [2^{-\rho}, 1], \quad \text{for } \rho \geq 0, \\ E[g(\rho, Z)] &\in [1, 2^{-\rho}], \quad \text{for } \rho \in (-1, 0]. \end{aligned}$$

Moreover, g being continuous in z for fixed values of ρ implies that every intermediate value of the corresponding bounded interval will be taken by the function $g(\rho, z)$ for $z \in [0, 1]$, i.e., we can always find a $z^* \in [0, 1]$ such that $E[g(\rho, Z)] = g(\rho, z^*)$. Since, as indicated in Fact 1, the random variable Z_{BSC} is a constant z_{BSC}, the BSC defined in (2.22) will be a BSC such that $z_{BSC} = z^*$. From this the crossover probability of the channel can be inferred.

To find a BEC which satisfies (2.22), we will use the BSC we have just defined with parameter z^*. Note that the extreme values of the bounded interval from which $g(\rho, z)$ takes values are given by $2^{-\rho} = g(\rho, 0)$ and $1 = g(\rho, 1)$. Moreover, the function g being continuous in $z \in [0, 1]$ for fixed values of ρ, we can weight these two values with a probability distribution p_0 and $1 - p_0$ such that

$$g(\rho, z^*) = p_0 g(\rho, 0) + (1 - p_0) g(\rho, 1).$$

As by Fact 1, the random variable Z_{BEC} is $\{0, 1\}$ valued, the BEC defined in (2.22) will be a BEC with erasure probability given by $P(Z_{BEC} = 0) = p_0$. □

Upon this lemma, another property of BECs and BSCs is due:

Fact 3. The set of BECs and the set of BSCs both sweep all the possible values the E_0 parameters of B-DMCs can take at any $\rho > -1$.

Suppose now the E_0 curves of a BEC and a BSC intersect at a particular $\rho^* > -1$ such that $\rho^* \neq 0$. We would like to know if there are any other $\rho > -1$, apart from the trivial $\rho = 0$, such that the E_0 curves of these two channels intersect again.

Lemma 2.5. *Suppose a BSC BSC and a BEC BEC satisfy*

$$E_0(\rho^*, BEC) = E_0(\rho^*, BSC), \tag{2.23}$$

for some $\rho^* > -1$ such that $\rho^* \neq 0$. Then, if $\rho^* \leq 1$, there is only one other intersection point between the E_0 curves of the channels at $\rho = 0$. If $\rho^* > 1$, the only intersection point in the interval $(-1, 1]$ is once more at $\rho = 0$, and for the rest either the E_0 curves of the channels are tangent to each others at ρ^*, i.e.,

$$R(\rho^*, BEC) = R(\rho^*, BSC) \tag{2.24}$$

is satisfied, or there exists a different $\rho' > 1$ such that

$$E_0(\rho', BEC) = E_0(\rho', BSC). \tag{2.25}$$

Proof. Let the erasure probability of the channel BEC be ϵ and the channel BSC be such that $z_{BSC} = z$. By (2.17) and (2.19), the condition for equality in (2.23) translates into

$$g(\rho^*, z) = 2^{-\rho^*}(1 - \epsilon) + \epsilon.$$

Let the function $\tilde{g}(\rho, z)$ be defined as

$$\tilde{g}(\rho, z) = \frac{g(\rho, z) - 2^{-\rho}}{1 - 2^{-\rho}}.$$

Observe that

$$\tilde{g}(\rho^*, z) = \epsilon,$$

and in order for (2.25) to hold, we are looking for another ρ' such that $\tilde{g}(\rho', z) = \epsilon$. To find the answer, we need to study the monotonicity properties of the function $\tilde{g}(\rho, z)$ with respect to ρ. Indeed, a straightforward but tedious analysis shows that the first derivative of $\tilde{g}(\rho, z)$ with respect to ρ changes sign only once at $\rho_{\max}(z) \geq 3$ for every fixed value of z such that $\tilde{g}(\rho, z)$ is increasing for $\rho \in (0, \rho_{\max}(z))$ and decreasing for $\rho > \rho_{\max}(z)$ with $\lim_{\rho\to\infty} \tilde{g}(\rho, z) = \tilde{g}(1, z)$. Consequently, if $\rho^* \in (-1, 0) \cup (0, 1]$, no other ρ' can satisfy (2.25). On the other hand, if $\rho^* > 1$, but $\rho^* \neq \rho_{\max}(z)$, then the two curves intersect twice. Finally, if $\rho^* = \rho_{\max}(z)$, not only no other ρ' can satisfy (2.25), but also

$$\tilde{g}(\rho^*, z) = \tilde{g}(\rho_{\max}(z), z) \geq \tilde{g}(\rho, z)$$

holds for all $\rho > -1$. In this case, the E_0 curves of the channels will be tangent to each other, so (2.24) will hold as well. □

The previous lemma says that if the E_0 curves of a BEC and a BSC intersect somewhere between the interval $(-1, 0) \cup (0, 1]$, they cannot intersect a second time, except trivially at 0, and if otherwise they intersect in the interval $(1, \infty)$, either the two curves are tangent to each other or they intersect twice in that interval, and the only intersection point in the interval $(-1, 1]$ is again at 0. The significance of this lemma will become clear later when we interpret the extremality results. The lemma will help us to understand why some intervals of $\rho > -1$ are more interesting in the context of the extremality results presented in Theorem 3.1.

Appendix

The first part of this appendix proves Lemma 2.2. In the second part, Lemma 2.7 shows that the E_0 function of stochastically degraded channels are ordered.

2.A Proof of Lemma 2.2

Proof. Taking the first derivative of (2.13) with respect to z, we get

$$\begin{aligned}\frac{\partial g(\rho, z)}{\partial z} &= \left(\frac{1}{2}(1+z)^{\frac{1}{1+\rho}} + \frac{1}{2}(1-z)^{\frac{1}{1+\rho}}\right)^{\rho} \left(\frac{1}{2}(1+z)^{\frac{-\rho}{1+\rho}} - \frac{1}{2}(1-z)^{\frac{-\rho}{1+\rho}}\right) \\ &= \underbrace{\left(\frac{1}{2}\right)^{1+\rho} \left(1 + \left(\frac{1-z}{1+z}\right)^{\frac{1}{1+\rho}}\right)^{\rho}}_{\geq 0} \left(1 - \left(\frac{1-z}{1+z}\right)^{\frac{-\rho}{1+\rho}}\right). \qquad (2.26)\end{aligned}$$

As we have $(1-z)/(1+z) \leq 1$, for $\forall z \in [0,1]$, the monotonicity claims follow by noting that when $\rho \in [0,\infty)$:

$$\frac{\rho}{1+\rho} \geq 0 \quad \Rightarrow \quad \left(1 - \left(\frac{1-z}{1+z}\right)^{\frac{-\rho}{1+\rho}}\right) \leq 0 \quad \Rightarrow \quad \frac{\partial g(\rho,z)}{\partial z} \leq 0,$$

and when $\rho \in (-1,0]$:

$$\frac{\rho}{1+\rho} \leq 0 \quad \Rightarrow \quad \left(1 - \left(\frac{1-z}{1+z}\right)^{\frac{-\rho}{1+\rho}}\right) \geq 0 \quad \Rightarrow \quad \frac{\partial g(\rho,z)}{\partial z} \geq 0.$$

Taking the second derivative with respect to z, we get

$$\frac{\partial^2 g(\rho,z)}{\partial z^2} = -\frac{\rho}{1+\rho} \underbrace{\left(1-z^2\right)^{\frac{1}{1+\rho}-2} \left(\frac{1}{2}(1+z)^{\frac{1}{1+\rho}} + \frac{1}{2}(1-z)^{\frac{1}{1+\rho}}\right)^{-1+\rho}}_{\geq 0}.$$

The convexity claims follow once again by inspecting the sign of $\dfrac{\rho}{1+\rho}$ in different intervals, i.e., when $\rho \in [0,\infty)$:

$$\frac{\rho}{1+\rho} \geq 0 \quad \Rightarrow \quad \frac{\partial^2 g(\rho,z)}{\partial z^2} \leq 0,$$

and when $\rho \in (-1,0]$:

$$\frac{\rho}{1+\rho} \leq 0 \quad \Rightarrow \quad \frac{\partial^2 g(\rho,z)}{\partial z^2} \geq 0. \qquad \square$$

2.B Stochastic Degradation Ordering

We first introduce the definition of stochastic degradation and then prove the ordering lemma.

Definition 2.6. A DMC $W\colon \mathcal{X} \to \mathcal{Y}_1$ is *stochastically degraded* with respect to another DMC $V\colon \mathcal{X} \to \mathcal{Y}_2$ if there exists a channel $P\colon \mathcal{Y}_1 \to \mathcal{Y}_2$ such that

$$V(z|x) = \sum_{y \in \mathcal{Y}_1} W(y|x) P(z|y), \quad \text{for all } z \in \mathcal{Y}_2. \tag{2.27}$$

Lemma 2.7. *Let W and V be two DMCs such that V is stochatically degraded with respect to W. Then, for any input distribution Q,*

$$E_0(\rho,Q,V) \leq E_0(\rho,Q,W), \quad \textit{for } \rho \geq 0,$$
$$E_0(\rho,Q,V) \geq E_0(\rho,Q,W), \quad \textit{for } \rho \in (-1,0].$$

Proof. Suppose that V is stochastically degraded with respect to W and Q is fixed. We define $E_0(\rho, Q, W) = -\log F(\rho, Q, W)$, for $\rho > -1$, with

$$\begin{aligned} F(\rho, Q, W) &:= \sum_{y\in\mathcal{Y}_1} \left[\sum_{x\in\mathcal{X}} Q(x)W(y|x)^{1/r}\right]^r \\ &= \sum_{y\in\mathcal{Y}_1}\sum_{z\in\mathcal{Y}_2} P(z|y) \left[\sum_{x\in\mathcal{X}} Q(x)W(y|x)^{1/r}\right]^r \\ &= \sum_{y\in\mathcal{Y}_1}\sum_{z\in\mathcal{Y}_2} \left[\sum_{x\in\mathcal{X}} Q(x)(P(z|y)W(y|x))^{1/r}\right]^r, \end{aligned}$$

for $r = 1 + \rho$. Similarly, we write $E_0(\rho, Q, V) = -\log F(\rho, V)$ with

$$F(\rho, Q, V) = \sum_{z\in\mathcal{Y}_2} \left[\sum_{x\in\mathcal{X}} Q(x) \left(\sum_{y\in\mathcal{Y}_1} P(z|y)W(y|x)\right)^{1/r}\right]^r.$$

If for every fixed $z = z_0$, we have

$$\begin{aligned} &\sum_{y\in\mathcal{Y}_1} \left[\sum_{x\in\mathcal{X}} Q(x) \left(P(z_0|y)W(y|x)\right)^{1/r}\right]^r \\ &\qquad\qquad \le \left[\sum_{x\in\mathcal{X}} Q(x) \left(\sum_{y\in\mathcal{Y}_1} Q(z_0|y)W(y|x)\right)^{1/r}\right]^r, \quad (2.28) \end{aligned}$$

for $r \ge 1$ (thus $\rho \ge 0$), and the reverse inequality for $r \in (0, 1]$ (thus $\rho \in (-1, 0]$), we will come through with the desired orderings. Now, if we let

$$a(x, y) = Q(x)^r P(z_0|y)W(y|x) \ge 0,$$

for $r > 0$, (2.28) is equivalent to

$$\sum_{y\in\mathcal{Y}_1} \left[\sum_{x\in\mathcal{X}} a(x, y)^{1/r}\right]^r \le \left[\sum_{x\in\mathcal{X}} \left(\sum_{y\in\mathcal{Y}_1} a(x, y)\right)^{1/r}\right]^r,$$

for $r \ge 1$. But this is true by *Minkowski's integral inequality*, see [5, Problem 4.15(g)]. Similarly, the reverse inequality holds for $r \in (0, 1]$. This concludes the proof. □

Chapter 3

Extremality for Gallager's Reliability Function E_0

Do not get confused by the fancy word: *Extremality* refers here to the condition of being *extremal*, meaning that it is a condition of or relating to *extrema*, i.e., maximal or minimal values [21]. We thus ensure we will be safely doing information theory, upper and lower bounding the possible range of our measure of information. Nonetheless, our bounds will not be simply upper and lower, but in addition, they will be attained by the extremal channels. The goal of this chapter is to characterize the extremality of the $E_0(\rho)$ curves of the BEC and the BSC among all the $E_0(\rho)$ curves that can be generated under the uniform input distribution by the class of B-DMCs whose $E_0(\rho)$ curves pass through a given point (ρ_0, e_0), for some $\rho_0 > -1$.

What's Coming, Doc?

In Theorem 3.1, we will prove that when $\rho_0 \in (-1, 1]$, these two channels remain extremal along the $E_0(\rho)$ curves for any $\rho > -1$. We will also prove that when $\rho_0 > 1$, while these two channels are extremal along the $E_0(\rho)$ curves for any $\rho \in (-1, 1]$, no extremality beyond $\rho > 1$ can be formulated in general[1]. We will also show in the theorem that a certain extremality property holds even when the quantities appearing in the parametric form of the random coding error exponent, i.e., E_0 and E_0', are evaluated at different values of the parameter ρ. In particular, we will prove that among all channels with a given value of $E_0(\rho_1)$, for any $\rho_1 \in [0, 1]$, the BEC and the BSC distinguish themselves as follows: they have, respectively, the largest and the smallest value of $E_0'(\rho_2)$ for any $\rho_2 \geq \rho_1$. As the random coding exponent is obtained by tracing the map $\rho \to (E_0'(\rho), E_0(\rho) - \rho E_0'(\rho))$, for

[1] As a particular case, we will show that when $\rho_0 \in [1, 3]$, the $E_0(\rho)$ curves of the BEC and the BSC are still extremal for $\rho \in (-1, 3]$.

$\rho \in [0, 1]$, among the simple corollaries of this will be the conclusion that of all the symmetric channels with the same capacity, the BEC and the BSC have the largest and the smallest value of the random coding exponents $E_r(R)$, a result reported in [22]. Finally, we will discuss how these properties yield in straightforward fashion extremal properties for the Rényi entropies.

The extremal results for E_0 and E_0' in the region where $\rho > -1$ can be applied to obtain upper and lower bounds to various error exponents and the guessing exponent. We introduced some of these exponents in the previous chapter. For a concise list of the definitions of other error exponents involving the E_0 function, we refer to [23], a recent study which also examined the extremality of $E_0(\rho)$ for $\rho > -1$, but only for the special class of symmetric B-DMCs of the same capacity.

3.1 The Extremality Theorem

The main result of this chapter is stated in the following theorem. The proof of the theorem is given in Section 3.3.

Theorem 3.1. *Given any fixed value of $\rho_1 > -1$, suppose a B-DMC W, a binary symmetric channel BSC, and a binary erasure channel BEC satisfy*

$$E_0(\rho_1, BSC) \overset{(a)}{\leq} E_0(\rho_1, W) \overset{(a')}{\leq} E_0(\rho_1, BEC), \tag{3.1}$$

for $\rho_1 \neq 0$, or

$$\lim_{\rho \to 0} \frac{E_0(\rho, BSC)}{\rho} \overset{(a_0)}{\leq} \lim_{\rho \to 0} \frac{E_0(\rho, W)}{\rho} \overset{(a_0')}{\leq} \lim_{\rho \to 0} \frac{E_0(\rho, BEC)}{\rho}, \tag{3.2}$$

for $\rho_1 = 0$.

(Part 1) If $\rho_1 \in [0, 3]$, then

$$R(\rho_2, BSC) \overset{(b)}{\leq} R(\rho_2, W) \overset{(b')}{\leq} R(\rho_2, BEC), \tag{3.3}$$

$$E_0(\rho_2, BSC) \overset{(c)}{\leq} E_0(\rho_2, W) \overset{(c')}{\leq} E_0(\rho_2, BEC), \tag{3.4}$$

for any $\rho_2 \in [\rho_1, 3]$.

(Part 2) If $\rho_1 \in (-1, 0]$, then

$$R(\rho_2, BEC) \overset{(d)}{\leq} R(\rho_2, W) \overset{(d')}{\leq} R(\rho_2, BSC), \tag{3.5}$$

$$E_0(\rho_2, BSC) \overset{(e)}{\leq} E_0(\rho_2, W) \overset{(e')}{\leq} E_0(\rho_2, BEC), \tag{3.6}$$

for any $\rho_2 \in (-1, \rho_1]$,

(Part 3) If $\rho_1 \in (-1, 0]$, *then*

$$E_0(\rho_2, BSC) \overset{(f)}{\leq} E_0(\rho_2, W) \overset{(f')}{\leq} E_0(\rho_2, BEC), \tag{3.7}$$

for any $\rho_2 \geq 0$.

If $\rho_1 \in [0, 1]$, *then*

$$E_0(\rho_2, BSC) \overset{(g)}{\leq} E_0(\rho_2, W) \overset{(g')}{\leq} E_0(\rho_2, BEC), \tag{3.8}$$

for any $\rho_2 \geq \rho_1$.

If $\rho_1 > 1$, *then*

$$E_0(\rho_2, BEC) \overset{(h)}{\leq} E_0(\rho_2, W) \overset{(h')}{\leq} E_0(\rho_2, BSC), \tag{3.9}$$

for any $\rho_2 \in [0, 1]$.

If $\rho_1 > 1$, *then*

$$E_0(\rho_2, BSC) \overset{(i)}{\leq} E_0(\rho_2, W) \overset{(i')}{\leq} E_0(\rho_2, BEC), \tag{3.10}$$

for any $\rho_2 \in (-1, 0]$.

Moreover, the extremalities hold with strict inequalities, except for $\rho_2 = 0$, *whenever* (a) *and* (a') *in* (3.1) *are strict for* $\rho_1 \neq 0$, *or* (a_0) *and* (a_0') *in* (3.2) *are strict for* $\rho_1 = 0$.

Remark 3.2. In Theorem 3.1, the inequalities (a) and (a_0) imply the inequalities (b), (c), (d), (e), (f), (g), (h), and (i). Similarly, the inequalities (a') and (a_0') imply the inequalities (b') through (i').

Remark 3.3. The value "3" that appears in the intervals in Part 1 of the theorem is a conservative estimate. The reader who follows the proof of Lemma 3.6, stated in Section 3.3 and proved in Appendix 3.A, will notice that this "3" may be replaced by a $\rho^*(W)$ that depends on the channel W. In the proof of Lemma 3.6, it is shown that $\rho^*(W) \geq 3$ for any W, but the lower bound is not necessarily tight. We chose the value 3 so as to not further complicate the statement of the theorem.

For the special case where $\rho_1 = \rho_2 = \rho$, for $\rho \in [0, 1]$, we recover in the next corollary, a result obtained in [20].

Corollary 3.4 ([20])**.** *Given a symmetric B-DMC* W, *for any fixed value of* $\rho \in [0, 1]$, *find a binary symmetric channel* BSC, *and a binary erasure channel* BEC *through*

the equality

$$R(\rho, W) = R(\rho, BEC) = R(\rho, BSC). \tag{3.11}$$

Then,

$$\begin{aligned} E_0(\rho, BEC) \leq E_0(\rho, W) \leq E_0(\rho, BSC), \\ E_r(\rho, BEC) \leq E_r(\rho, W) \leq E_r(\rho, BSC). \end{aligned} \tag{3.12}$$

Proof. Since $E_r(\rho, W) = E_0(\rho, W) - \rho R(\rho, W)$, it suffices to prove the first set of inequalities in view of (3.11). Taking $\rho_1 = \rho_2 = \rho$, (3.12) holds by Theorem 3.1. To see this, observe that had the channels on the contrary satisfied

$$E_0(\rho, BSC) < E_0(\rho, W) < E_0(\rho, BEC),$$

the results in Part 1 of the theorem would imply

$$R(\rho, BSC) < R(\rho, W) < R(\rho, BEC),$$

contradicting the assumption (3.11) of the corollary. □

Another particular case of Theorem 3.1 when $\rho_1 = 0$ recovers the result in [22]: amongst all symmetric B-DMCs of the same capacity, the BEC and the BSC are extremal with respect to the random coding exponent.

Corollary 3.5 ([22, Theorem 2.3])**.** *Given a symmetric B-DMC W of capacity $I(W)$, we define a binary symmetric channel BSC, and a binary erasure channel BEC of the same capacity through the equality*

$$I(W) = I(BEC) = I(BSC).$$

Then, the random coding error exponent of the channels satisfy

$$E_r(R, BSC) \leq E_r(R, W) \leq E_r(R, BEC).$$

Proof. Recall from Chapter 2 that the equality of the capacities is equivalent to

$$\lim_{\rho \to 0} \frac{E_0(\rho, W)}{\rho} = \lim_{\rho \to 0} \frac{E_0(\rho, BEC)}{\rho} = \lim_{\rho \to 0} \frac{E_0(\rho, BSC)}{\rho},$$

and for a symmetric channel, the random coding exponent is given by

$$E_r(R, W) = \max_{\rho \in [0,1]} [E_0(\rho, W) - \rho R].$$

But in this case, we know by Part 1 of Theorem 3.1 that we have

$$E_0(\rho_2, BSC) \leq E_0(\rho_2, W) \leq E_0(\rho_2, BEC),$$

for any $\rho_2 \in [0, 1]$. This, in turn, implies the inequality for the random coding exponent. □

Finally, note that in [23] the above result of [22] was extended to the region where $\rho > -1$. Namely, amongst all symmetric B-DMCs of the same capacity, the BEC and the BSC are extremal with

$$E_0(\rho, BSC) \leq E_0(\rho, W) \leq E_0(\rho, BEC),$$

for all $\rho > -1$. In particular, [23, Theorem 1] can also be recovered from Theorem 3.1.

3.2 A Graphical Interpretation

In this section, we provide a graphical interpretation of Theorem 3.1 and the corollaries through Figures 1 to 5. Suppose that the E_0 curves of a given B-DMC, a BEC, and a BSC pass through a given point (ρ_0, e_0), for some $\rho_0 > -1$.

By the results stated in (3.6) and (3.7), we know that when $\rho_0 \in (-1, 0)$, then these curves do not intersect again except at $\rho = 0$, and the BEC and BSC always remain extremal even though their extremal behaviour get reversed after the intersection points. Figure 3.1 illustrates this relation.

A special case where the E_0 curves of the BEC and the BSC remain extremal for the entire $\rho > -1$ region, and with no reversal, corresponds to channels of the same capacity; as discussed after Corollary 3.5, Theorem 3.1 shows that the E_0 curves of these channels are upper bounded by the BEC's curve and lower bounded by the BSC's one. Figure 3.2 illustrates this relation.

Another case where the E_0 curves of the BEC and the BSC exhibit extremality for the entire region $\rho > -1$ is when $\rho_0 \in (0, 1]$; one can infer from (3.4), (3.7), and (3.8) that the BEC and the BSC will be E_0 extremal, once again with the extremalities reversed after the intersections. Figure 3.3 illustrates this relation.

Now, we consider the case when $\rho_0 > 1$. Part 3 shows that the curves only intersect at $\rho = 0$ in the interval $\rho \in (-1, 1]$, and the BEC and the BSC are extremal in the intervals $(-1, 0)$ and $(0, 1]$ with reversed extremalities. Although Part 1 provides a partial result, it is not clear what happens in the interval $\rho > 1$. It turns out that the BEC and the BSC are no longer extremal for $\rho > 1$ in general. We will

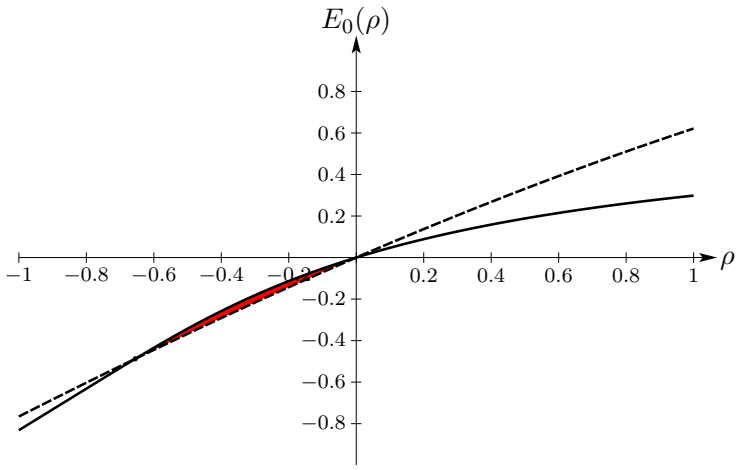

Figure 3.1: Extremality of $E_0(\rho)$ when the channels intersect at $\rho_0 \in (-1, 0)$. Dashed line: BEC(0.3) & Solid line: BSC(0.1102).

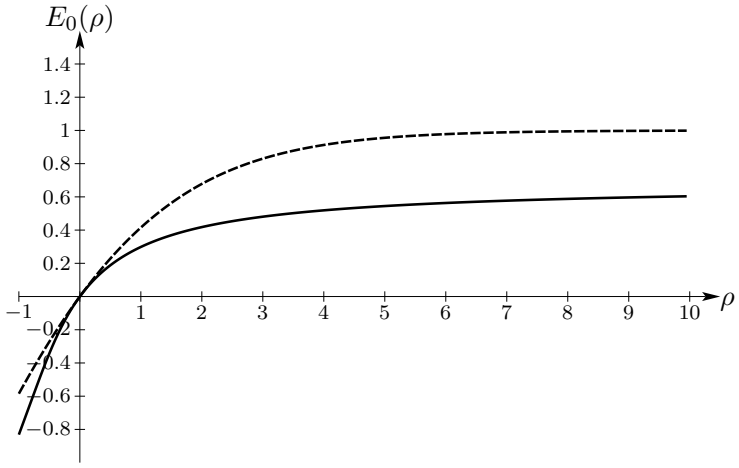

Figure 3.2: Extremality of $E_0(\rho)$ when the channels have equal capacity 0.5. Dashed line: BEC(0.5) & Solid line: BSC(0.1102).

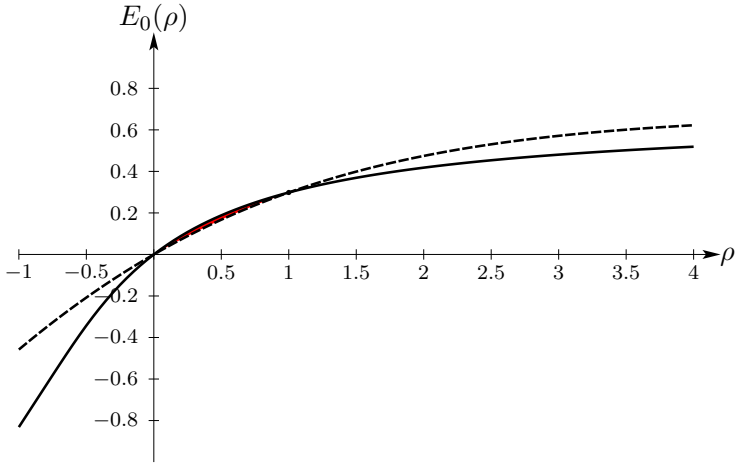

Figure 3.3: Extremality of $E_0(\rho)$ when the channels have equal cutoff rate. Dashed line: BEC(0.626278) & Solid line: BSC(0.1102).

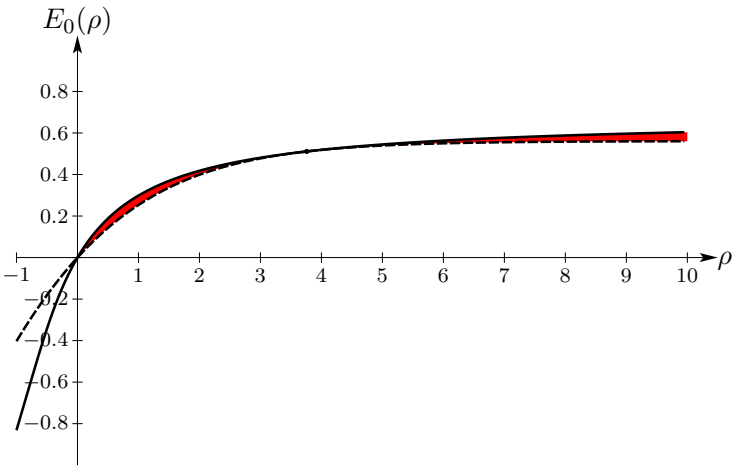

Figure 3.4: Extremality of $E_0(\rho)$ when the channels have equal $E_0(\rho^*)$ and equal rate at $\rho^* > 1$. Dashed line: BEC(0.6777) & Solid line: BSC(0.1102).

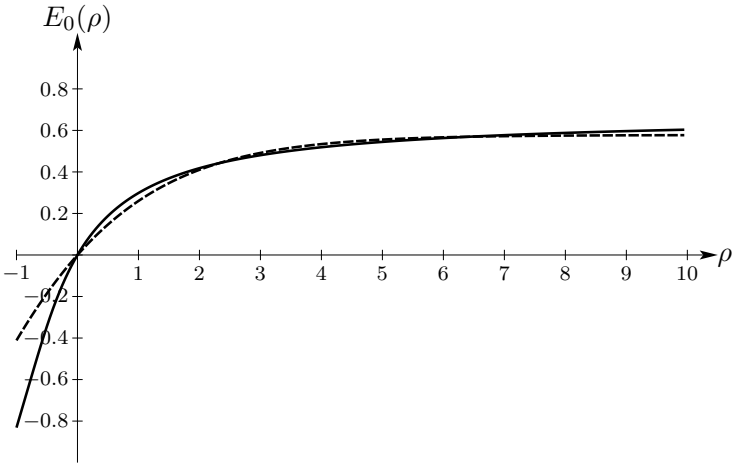

Figure 3.5: Extremality of $E_0(\rho)$ when the channels intersect at $\rho_0 > 1$. Dashed line: BEC(0.67) & Solid line: BSC(0.1102).

show this result by studying the intersection points of the E_0 curves of a given BSC with different BECs using Lemma 2.5.

Suppose a BEC BEC and a BSC BSC satisfy

$$E_0(\rho^*, BEC) = E_0(\rho^*, BSC), \tag{3.13}$$
$$R(\rho^*, BEC) = R(\rho^*, BSC), \tag{3.14}$$

for a particular $\rho^* > 1$. By Lemma 2.5, this corresponds to the case where the E_0 curves of these two channels are tangent at $\rho^* > 1$ and do not intersect at any other point except $\rho = 0$. Moreover, by Theorem 3.1, the capacities of the channels are

such that $I(BEC) \leq I(BSC)$. Figure 3.4 illustrates this relation.

Suppose the erasure probability of the BEC channel is increased. By the ordering we discussed in Fact 2, it is not difficult to see that the E_0 curves of the BSC and the new BEC will not intersect at any point other than $\rho = 0$. Instead, suppose that the erasure probability of the channel is decreased, but the capacity of the new BEC is still smaller than the capacity of the BSC. In this case, as long as the cutoff rate of the BSC is larger than the cutoff rate of the BEC, the BSC and the new BECs will intersect twice after $\rho = 0$, first in the interval $(1, \rho^*)$, then after ρ^*. Figure 3.5 illustrates this relation. Once the cutoff rate of the BEC becomes larger than that of the BSC, we are back at the situation where the intersection point falls in the interval $[0, 1]$, and we recover the general extremality result we have already discussed. Then, we can keep decreasing the erasure probability until the BEC and the BSC have the same capacity to recover another special case. Finally, decreasing more the erasure probability, until there is no other intersection anywhere except at $\rho = 0$, will cause the E_0 curves of the BSC and the new BECs to intersect in the interval $(-1, 0)$. In this latter case, the BSC and the BECs will once more be E_0 extremal for the entire $\rho > -1$ region.

The analysis above shows us that most of the BECs and the BSCs whose E_0 curves intersect in the interval where $\rho > 1$ have two intersection points in that interval. In such a case, the BEC and the BSC are no longer extremal as for a class of B-DMCs which satisfy for all the channels W in the class the equality

$$E_0(\rho_0, W) = E_0(\rho_0, BEC) = E_0(\rho_0, BSC),$$

for any fixed $\rho_0 > 1$, we do not expect the E_0 curves of the channels in this class to intersect again at the point where the E_0 curves of the BEC and the BSC intersect a second time, i.e, where $E_0(\rho', BEC) = E_0(\rho', BSC)$ holds for $\rho' > 1$ such that $\rho' \neq \rho_0$.

3.3 Proof of the Theorem

The proof of Theorem 3.1 rests on the next two convexity lemmas. The lemmas are proved in Appendix 3.A and 3.B, respectively.

Lemma 3.6. *For fixed values of $\rho_1, \rho_2 \in \mathbb{R} \setminus \{-1\}$, we define the function $\tilde{f}_{\rho_1,\rho_2}(t)$ by*

$$\tilde{f}_{\rho_1,\rho_2}(t) := \frac{\partial}{\partial \rho_2} g(\rho_2, g^{-1}(\rho_1, t)),$$

for $t \in [2^{-\rho}, 1] \cup [1, 2^{-\rho}]$. Let $\tilde{f}_\rho(t)$ denote the function when $\rho_1 = \rho_2 = \rho$. Then, $\tilde{f}_\rho(t)$ is a concave function in t when $\rho \in (0, 3]$, convex when $\rho \in (-\infty, -1) \cup (-1, 0]$. Moreover, the function $\tilde{f}_{\rho_1,\rho_2}(t)$ is concave when $\rho_1, \rho_2 \in [0, 1]$ such that $\rho_2 \geq \rho_1$.

Lemma 3.7. *For fixed values of $\rho_1, \rho_2 \in \mathbb{R} \setminus \{-1\}$, the function $f_{\rho_1,\rho_2}(t)$ defined as*

$$f_{\rho_1,\rho_2}(t) := g(\rho_2, g^{-1}(\rho_1, t)),$$

for $t \in [2^{-\rho}, 1] \cup [1, 2^{-\rho}]$, is concave in t when $\rho_1 \in (-1, 0]$ and $\rho_2 \geq 0$, when $\rho_1 \in [0, 1]$ and $\rho_2 \geq \rho_1$, and when $\rho_1 > 1$ and $\rho_2 \in (-1, 0)$, and the function is convex when $\rho_1 > 1$ and $\rho_2 \in (0, 1]$.

Before proving the theorem's statement in its most general form, we will prove two particular cases of the theorem in the next two lemmas assuming $\rho_1 = \rho_2 = \rho$.

Lemma 3.8. *Given any fixed value of $\rho \in (0, 3)$, suppose a B-DMC W, a binary symmetric channel BSC, and a binary erasure channel BEC satisfy the equality*

$$E_0(\rho, BSC) \leq E_0(\rho, W) \leq E_0(\rho, BEC). \tag{3.15}$$

Then, the following holds:

$$R(\rho, BSC) \leq R(\rho, W) \leq R(\rho, BEC), \tag{3.16}$$

where the inequalities are strict if the inequalities in (3.15) *are strict.*

Proof. Let us define another binary erasure channel BEC^* and another binary symmetric channel BSC^* through the following equality:

$$E_0(\rho, BSC^*) = E_0(\rho, W) = E_0(\rho, BEC^*). \tag{3.17}$$

Observe that by (2.12), the equality condition in (3.17) is equivalent to the equality of

$$\mathbb{E}\left[g(\rho, Z)\right] = \mathbb{E}\left[g(\rho, Z_{BEC^*})\right] = g(\rho, z_{BSC^*}). \tag{3.18}$$

Hence, the denominator in

$$R(\rho, W) = \frac{\partial}{\partial \rho} E_0(\rho, W) = \frac{\mathbb{E}\left[-\partial g(\rho, Z)/\partial \rho\right]}{\mathbb{E}\left[g(\rho, Z)\right] \ln 2} \tag{3.19}$$

is the same for the three channels. Then, the proof can be completed using the concavity of the function $\tilde{f}_\rho(t)$ in t for $\rho \in (0, 3]$, which is shown in Lemma 3.6, and the special structure of the Z random variables of a BEC and a BSC. To see this, let us define the random variable $A = g(\rho, Z) \in [2^{-\rho}, 1]$. Then, we note that $\tilde{f}_\rho(A) = \partial g(\rho, Z)/\partial \rho$, and $E[A]$ gives (3.18). So,

$$R(\rho, W) = \frac{\mathbb{E}\left[\tilde{f}_\rho(A)\right]}{E[A] \ln 2}, \quad R(\rho, BSC^*) = \frac{\tilde{f}_\rho(E[A])}{E[A] \ln 2}.$$

To derive the expression for $R(\rho, BEC^*)$, recall by Fact 1 that $Z_{BEC^*} \in \{0, 1\}$. Using $E[A] = \mathbb{E}\left[g(\rho, Z_{BEC^*})\right]$, we get

$$P(Z_{BEC^*} = 0) = \frac{E[A] - 1}{2^{-\rho} - 1}.$$

Hence,

$$R(\rho, BEC^*) = \frac{\tilde{f}_\rho(2^{-\rho})P(Z_{BEC^*} = 0) + \tilde{f}_\rho(1)P(Z_{BEC^*} = 1)}{E[A] \ln 2}.$$

Now, by the two sides of the Jensen's inequality for concave functions, we have

$$\tilde{f}_\rho(1) + \frac{\tilde{f}_\rho(1) - \tilde{f}_\rho(2^{-\rho})}{1 - 2^{-\rho}} \left(\mathbb{E}\left[A\right] - 1\right) \leq \mathbb{E}\left[\tilde{f}_\rho(A)\right] \leq \tilde{f}_\rho(\mathbb{E}\left[A\right]). \tag{3.20}$$

Dividing all sides by $\mathbb{E}\left[A\right] \ln 2 > 0$ and negating the expressions in (3.20), we get

$$R(\rho, BSC^*) \leq R(\rho, W) \leq R(\rho, BEC^*). \tag{3.21}$$

The final step of the proof is to show (3.21) implies (3.16). For that purpose, recall by Fact 2 that the set of BSCs and the set of BECs are strictly ordered in their E_0 and R parameters for $\rho \in (0, 3]$. As we have

$$E_0(\rho, BSC) \leq E_0(\rho, BSC^*), \tag{3.22}$$
$$E_0(\rho, BEC^*) \leq E_0(\rho, BEC), \tag{3.23}$$

we conclude by Lemma 2.3 that

$$R(\rho, BSC) \leq R(\rho, BSC^*), \tag{3.24}$$
$$R(\rho, BEC^*) \leq R(\rho, BEC) \tag{3.25}$$

holds for $\rho > 0$. From this (3.16) follows. Moreover, if the inequalities in (3.15) are strict than the ones in (3.22) and (3.23), and thus, (3.24) and (3.25) are strict as well. Consequently, the inequalities in (3.16) hold strictly as claimed. □

Remark 3.9. Note that Lemma 3.8 and Corollary 3.4 are of the same flavor. Indeed, one can easily derive one from the other using the degradation argument discussed in Fact 2. So, the result of [20] could also have been used to characterize the behavior of the E_0 curves for the $\rho \in (0, 1]$ interval. However, the proof of the lemma and the proof of the corollary are different as they involve different convexity analysis.

Lemma 3.10. *Given any fixed value of $\rho \in (-1, 0)$, suppose a B-DMC W, a binary symmetric channel BSC, and a binary erasure channel BEC satisfy the condition*

(3.15) *of Lemma 3.8. Then, the following holds:*

$$R(\rho, BEC) \leq R(\rho, W) \leq R(\rho, BSC), \tag{3.26}$$

where the inequalities are strict if the inequalities in (3.15) *are strict.*

Proof. Let BEC^* and BSC^* be as defined in the proof of Lemma 3.8. Once again, the equality condition in (3.17) implies the denominator in (3.19) is the same for the three channels. Then, the inequalities

$$R(\rho, BEC^*) \leq R(\rho, W) \leq R(\rho, BSC^*) \tag{3.27}$$

follow using the convexity of the function $\tilde{f}_\rho(t)$ in t when $\rho \in (-1, 0]$, which is shown in Lemma 3.6, and applying Jensen's inequalities. Finally, as $E_0(\rho, BSC) \leq E_0(\rho, BSC^*)$ and $E_0(\rho, BEC^*) \leq E_0(\rho, BEC)$, we know by Fact 2 that these BSCs and BECs are ordered by degradation. So, we conclude by Lemma 2.3 that we have $R(\rho, BSC^*) \leq R(\rho, BSC)$ and $R(\rho, BEC) \leq R(\rho, BEC^*)$, for $\rho \in (-1, 0]$. From this (3.26) follows. The claim about the strictness of the inequalities can be proved similarly as in the proof of Lemma 3.8. □

Now, we are ready to prove the theorem.

Proof of Theorem 3.1. We will first prove the claims for $\rho_1 \in (-1, 0) \cup (0, \infty)$, leaving the case $\rho_1 = 0$ to the last. In fact, we will show that the results proved for $\rho_1 \in (-1, 0) \cup (0, \infty)$ will immediately extend to $\rho_1 = 0$ by the continuity of E_0 in its arguments.

We start by proving the inequalities (3.3) and (3.4) in Part 1 for the case $\rho_1 \in (0, 3]$. By Lemma 3.8, we know that (3.3) holds for $\rho_2 = \rho_1$. So, we only need to prove the theorem for $\rho_2 \in (0, 3]$ such that $\rho_2 > \rho_1$. By the continuity of $E_0(\rho, BEC)$ and $E_0(\rho, BSC)$ in the channels' erasure and crossover probabilities, respectively, it suffices to show that

$$E_0(\rho_1, BSC) < E_0(\rho_1, W) < E_0(\rho_1, BEC)$$

implies

$$E_0(\rho_2, BSC) < E_0(\rho_2, W) < E_0(\rho_2, BEC). \tag{3.28}$$

Then, Lemma 3.8 will imply

$$R(\rho_2, BSC) < R(\rho_2, W) < R(\rho_2, BEC).$$

We define $\Gamma(\rho) = E_0(\rho, W) - E_0(\rho, BEC)$. Let $\Gamma'(\rho)$ denote the first derivative

of $\Gamma(\rho)$ with respect to ρ. Noting that $R(\rho, W) = \dfrac{\partial}{\partial \rho} E_0(\rho, W)$, the inequality in (3.28) is implied by the following statement:

$$\Gamma(\rho_1) < 0 \quad \text{and by Lemma 3.8} \quad (\Gamma(\rho) < 0 \Rightarrow \Gamma'(\rho) < 0) \quad \Rightarrow \quad \Gamma(\rho_2) < 0.$$

But this is true by elementary considerations on differential equations. Indeed, suppose to the contrary that

$$\Gamma(\rho_1) < 0, \quad \text{and} \quad (\Gamma(\rho) < 0 \Rightarrow \Gamma'(\rho) < 0), \quad \text{but} \quad \Gamma(\rho_2) \geq 0.$$

Then, there exists $\rho_1 < \rho_3 \leq \rho_2$ such that $\Gamma(\rho) < 0$, for $\forall \rho \in [\rho_1, \rho_3)$, and $\Gamma(\rho_3) = 0$. But then there exists $\rho_1 < \rho_4 < \rho_3$ such that

$$\Gamma'(\rho_4) = \frac{\Gamma(\rho_3) - \Gamma(\rho_1)}{\rho_3 - \rho_1} > 0,$$

and $\Gamma(\rho_4) < 0$, contradicting the assumption. The inequality for the BSC can be obtained similarly by letting

$$\Gamma(\rho) = E_0(\rho, BSC) - E_0(\rho, W),$$

and applying the above argument once more.

We continue with the proof of the inequalities in (3.5) and (3.6) in Part 2 for the case $\rho_1 \in (-1, 0)$. The proof follows along the same lines of the previous part. By Lemma 3.10, we know that the inequalities in (3.5) hold for $\rho_2 = \rho_1$. So, we only need to prove the theorem for $\rho_2 < \rho_1$. By the continuity of $E_0(\rho, BEC)$ and $E_0(\rho, BSC)$ in the channels' erasure and crossover probabilities, respectively, it suffices to show that

$$E_0(\rho_1, BSC) < E_0(\rho_1, W) < E_0(\rho_1, BEC)$$

implies

$$E_0(\rho_2, BSC) < E_0(\rho_2, W) < E_0(\rho_2, BEC).$$

Then, Lemma 3.10 will imply

$$R(\rho_2, BEC) < R(\rho_2, W) < R(\rho_2, BSC).$$

We define $\Gamma(\rho) = E_0(\rho, W) - E_0(\rho, BEC)$. Then, the corollary is implied by the following statement:

$$\Gamma(\rho_1) < 0 \quad \text{and by Lemma 3.10} \quad (\Gamma(\rho) < 0 \Rightarrow \Gamma'(\rho) > 0) \quad \Rightarrow \quad \Gamma(\rho_2) < 0.$$

But this is true by an analogous reasoning as before. The inequality for the BSC can be obtained similarly by letting

$$\Gamma(\rho) = E_0(\rho, BSC) - E_0(\rho, W),$$

and applying the above argument once more. This concludes the proof of Part 2.

For Part 3, we will only do the proof of (3.7) for the case $\rho_1 \in (-1, 0)$ and $\rho_2 \geq 0$ as all the other claims can be proved in the same way using the convexity properties of the function $f_{\rho_1,\rho_2}(t)$ discussed in Lemma 3.7.

Let $A = g(\rho_1, Z)$. We know that the condition in (3.1) is equivalent to

$$\mathbb{E}\left[g(\rho_1, Z_{BEC})\right] \leq \mathbb{E}\left[g(\rho_1, Z)\right] \leq g(\rho_1, z_{BSC}).$$

Define the BEC BEC^* and the BSC BSC^* through the equality

$$\mathbb{E}\left[g(\rho_1, Z)\right] = \mathbb{E}\left[g(\rho_1, Z_{BEC^*})\right] = g(\rho_1, z_{BSC^*}).$$

As by Lemma 3.7 we know the function $f_{\rho_1,\rho_2}(t)$ is concave in t when $\rho_1 \in (-1, 0]$ and $\rho_2 \geq 0$, we can apply the two sides of Jensen's inequality to obtain

$$\mathbb{E}\left[f_{\rho_1,\rho_2}(g(\rho_1, Z_{BEC^*}))\right] \leq \mathbb{E}\left[f_{\rho_1,\rho_2}(A)\right] \leq f_{\rho_1,\rho_2}(g(\rho_1, Z_{BSC^*})),$$

which is equivalent to

$$\mathbb{E}\left[g(\rho_2, Z_{BEC^*})\right] \leq \mathbb{E}\left[g(\rho_2, Z)\right] \leq g(\rho_2, z_{BSC^*}).$$

To get the claimed inequalities in (3.7), we simply need to use the ordering argument of Fact 2 for the two BECs and the two BSCs. As we have illustrated this argument before in the proof of Lemma 3.8, we do not repeat it here.

The last step is to prove the theorem for the case $\rho_1 = 0$. We will only present the proof extension for the inequalities (b') and (c') in Part 1 as the same argument can be used to extend all the remaining inequalities. Moreover, once again by the continuity of $E_0(\rho, BEC)$ in the channels' erasure probability, it suffices to show the results assuming (a_0') in (3.2) holds with strict inequality.

So, we assume the given channels W and BEC satisfy $I(W) < I(BEC)$. Then, using the identity in (2.4), we get

$$\lim_{\rho \to 0^+} \frac{E_0(\rho, W) - E_0(\rho, BEC)}{\rho} = I(W) - I(BEC) < 0.$$

(We assumed $\rho \to 0^+$ for simplicity as the above limit for $\rho \to 0$ is well defined).

Hence, for any sufficiently small $\rho > 0$, we have

$$E_0(\rho, W) < E_0(\rho, BEC).$$

Moreover, we already proved that, for all $\rho_2 \in [\rho, 3]$, this implies

$$E_0(\rho_2, W) \leq E_0(\rho_2, BEC).$$

As $\rho > 0$ is arbitrary, we conclude the result should hold for all $\rho_2 \in [0, 3]$.

Now, we can carry the proof as follows. First, we let $\epsilon \in [0, 1]$ be the erasure probability of the BEC BEC_ϵ which satisfies $I(W_\epsilon) = I(W)$. Then, we take a sequence of BECs BEC_{ϵ_n} of erasure probabilities $\epsilon_n \in [0, 1]$ such that the sequence ϵ_n is increasing to ϵ. In this case, we know that

$$I(W) < I(BEC_{\epsilon_n}).$$

By the previous argument, we conclude that for all the channels BEC_{ϵ_n},

$$E_0(\rho_2, W) \leq E_0(\rho_2, BEC_{\epsilon_n})$$

holds for all $\rho_2 \in [0, 3]$. Taking the limit for the sequence ϵ_n, we conclude by continuity that the result also holds for the channel BEC_ϵ, i.e.,

$$E_0(\rho_2, W) \leq E_0(\rho_2, BEC_\epsilon)$$

holds for $\rho_2 \in [0, 3]$. As we have $E_0(\rho_2, BEC_\epsilon) \leq E_0(\rho_2, BEC)$, the inequality (c') in (3.4) is proved. By Lemma 3.8, the inequality (b') follows. □

3.4 Extremality of Rényi Entropies

In this section, we show how the results of Theorem 3.1 can be translated into extremalities for Rényi entropies using the definition given in (2.3).

Observe that the assumption in (3.1) of Theorem 3.1 can be equivalently stated as

$$\frac{E_0(\rho_1, BSC)}{\rho_1} \leq \frac{E_0(\rho_1, W)}{\rho_1} \leq \frac{E_0(\rho_1, BEC)}{\rho_1},$$

for $\rho_1 > 0$, and

$$\frac{E_0(\rho_1, BEC)}{\rho_1} \leq \frac{E_0(\rho_1, W)}{\rho_1} \leq \frac{E_0(\rho_1, BSC)}{\rho_1},$$

for $\rho_1 \in (-1, 0)$. The reversal caused by the sign of ρ will not be a problem for

extending the proof of the previous section to Rényi entropies as we know that, by Lemma 2.3, while for $\rho_1 > 0$ a 'worse' BEC and a 'worse' BSC have smaller E_0 parameters, for $\rho \in (-1, 0)$ the opposite is true. Consequently, all the results obtained for the parameter $E_0(\rho, W)$ can be restated in terms of Rényi entropies via (2.3). For the sake of brevity, we will only restate in the next corollary the result given in (3.8) in Part 3 of the theorem in terms of Rényi entropies.

Corollary 3.11. *Given a binary uniform random variable X, among all jointly distibuted random variables (X, Y) of equal Rényi equivocation $H_\alpha(X \mid Y)$ of order $\alpha \in [1/2, 1]$, the Rényi equivocation of order $\beta \geq 0$ such that $\beta \geq \alpha$ is maximized when X and Y are coupled by a BEC, and minimized when coupled by a BSC. For $\beta \leq \alpha$ values, the maximizing and minimizing distributions are reversed.*

Proof. Recall that $\alpha = 1/(1+\rho)$. So for $\alpha \in [1/2, 1]$, we have $\rho \in [0, 1]$. Moreover, α is decreasing with ρ. Hence, the inequalities for $\beta \leq \alpha$ and for $\beta \geq \alpha$ follow directly from (3.8) in Part 3 of Theorem 3.1 using the definition given in equation (2.3) together with the fact that $H_\alpha(X) = 1$ under the uniform distribution. □

Appendix

The Appendices contain three parts. In the first two of them, we prove Lemma 3.6 and Lemma 3.7, respectively. The final part proves two other lemmas used in these proofs.

3.A Proof of Lemma 3.6

Proof. We begin by introducing some definitions to simplify notations. Let

$$g'(\rho, z) := \frac{\partial g(\rho, z)}{\partial z}.$$

We define

$$\begin{aligned} \lambda(z) &:= \frac{1-z}{1+z}, \\ \alpha(\rho, z) &:= (1 + \lambda(z)^{\frac{1}{1+\rho}})^\rho, \\ \beta(\rho, z) &:= (1 - \lambda(z)^{\frac{-\rho}{1+\rho}}), \end{aligned} \tag{3.29}$$

for $z \in [0, 1]$, $\rho \in \mathbb{R} \setminus \{-1\}$. By (2.26) in Lemma 2.2, we have

$$g'(\rho, z) = \left(\frac{1}{2}\right)^{1+\rho} \alpha(\rho, z)\beta(\rho, z).$$

Taking the first derivative of $\tilde{f}_{\rho_1,\rho_2}(t)$ with respect to t, we obtain

$$\begin{aligned}\frac{\partial \tilde{f}_{\rho_1,\rho_2}(t)}{\partial t} &= \frac{\partial}{\partial t}\frac{\partial}{\partial \rho_2} g(\rho_2, g^{-1}(\rho_1, t)) \\ &= \frac{\partial}{\partial \rho_2}\frac{\partial}{\partial t} g(\rho_2, g^{-1}(\rho_1, t)) \\ &= \frac{\partial}{\partial \rho_2}\frac{g'(\rho_2, g^{-1}(\rho_1, t))}{g'(\rho_1, g^{-1}(\rho_1, t))}.\end{aligned}$$

Let $z = g^{-1}(\rho_1, t)$. As $g(\rho, z)$ is a monotone function in z by Lemma 2.2, so is $z = g^{-1}(\rho, t)$ in t. Hence, we can check the convexity of $\tilde{f}_{\rho_1,\rho_2}(t)$ with respect to t from the monotonicity with respect to z of the following expression:

$$\begin{aligned}&\frac{\partial}{\partial \rho_2}\frac{g'(\rho_2, z)}{g'(\rho_1, z)} \\ =&\frac{\partial}{\partial \rho_2} 2^{\rho_1-\rho_2}\frac{\alpha(\rho_2, z)\beta(\rho_2, z)}{\alpha(\rho_1, z)\beta(\rho_1, z)} \\ =&\frac{2^{-\rho_2}\alpha(\rho_2, z)\beta(\rho_2, z)}{2^{-\rho_1}\alpha(\rho_1, z)\beta(\rho_1, z)}\left(\frac{\partial 2^{-\rho_2}\alpha(\rho_2, z)/\partial \rho_2}{2^{-\rho_2}\alpha(\rho_2, z)} + \frac{\partial \beta(\rho_2, z)/\partial \rho_2}{\beta(\rho_2, z)}\right),\end{aligned} \tag{3.30}$$

where

$$\begin{aligned}&\frac{\partial 2^{-\rho_2}\alpha(\rho_2, z)}{\partial \rho_2} \\ =&\frac{\partial}{\partial \rho_2}\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right)^{\rho_2} \\ =&\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right)^{\rho_2} \times \\ &\qquad\left(\ln\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right) + \rho_2\frac{\frac{1}{2}\lambda(z)^{\frac{1}{1+\rho_2}}\frac{-1}{(1+\rho_2)^2}\ln\lambda(z)}{\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}}\right) \\ =&2^{-\rho_2}\alpha(\rho_2, z)\left(\ln\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right) - \frac{\rho_2\lambda(z)^{\frac{1}{1+\rho_2}}\ln\lambda(z)}{(1+\rho_2)^2\left(1+\lambda(z)^{\frac{1}{1+\rho_2}}\right)}\right),\end{aligned}$$

and

$$\frac{\partial \beta(\rho_2, z)}{\partial \rho_2} = \frac{\partial}{\partial \rho_2}\left(1-\lambda(z)^{\frac{-\rho_2}{1+\rho_2}}\right) = \frac{1}{(1+\rho_2)^2}\lambda(z)^{\frac{-\rho_2}{1+\rho_2}}\ln\lambda(z).$$

Hence, the expression inside the parenthesis in (3.30) equals

$$\ln\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right) - \frac{\rho_2\lambda(z)^{\frac{1}{1+\rho_2}}\ln\lambda(z)}{(1+\rho_2)^2\left(1+\lambda(z)^{\frac{1}{1+\rho_2}}\right)} + \frac{\lambda(z)^{\frac{-\rho_2}{1+\rho_2}}\ln\lambda(z)}{(1+\rho_2)^2\left(1-\lambda(z)^{\frac{-\rho_2}{1+\rho_2}}\right)}$$
$$= \ln\left(\frac{1+\lambda(z)^{\frac{1}{1+\rho_2}}}{2}\right) - \frac{\rho_2\lambda(z)^{\frac{1}{1+\rho_2}}\ln\lambda(z)}{(1+\rho_2)^2\left(1+\lambda(z)^{\frac{1}{1+\rho_2}}\right)} + \frac{\ln\lambda(z)}{(1+\rho_2)^2\left(\lambda(z)^{\frac{\rho_2}{1+\rho_2}}-1\right)}.$$

To simplify derivations we define

$$\Phi(k,\rho_1,\rho_2) := \frac{\left(\frac{1+k^{\frac{1}{1+\rho_2}}}{2}\right)^{\rho_2}\left(1-k^{\frac{-\rho_2}{1+\rho_2}}\right)}{\left(\frac{1+k^{\frac{1}{1+\rho_1}}}{2}\right)^{\rho_1}\left(1-k^{\frac{-\rho_1}{1+\rho_1}}\right)}, \tag{3.31}$$

and

$$\Psi(k,\rho_2) := \ln\left(\frac{1+k^{\frac{1}{1+\rho_2}}}{2}\right) + \frac{\ln k}{(1+\rho_2)^2}\left(-\frac{\rho_2 k^{\frac{1}{1+\rho_2}}}{1+k^{\frac{1}{1+\rho_2}}} + \frac{1}{k^{\frac{\rho_2}{1+\rho_2}}-1}\right)$$
$$= \ln\left(\frac{1+k^{\frac{1}{1+\rho_2}}}{2}\right) + \frac{\left(1+k^{\frac{1}{1+\rho_2}} - \rho_2\left(k-k^{\frac{1}{1+\rho_2}}\right)\right)\ln k}{(1+\rho_2)^2\eta(k,\rho_2)}, \tag{3.32}$$

where

$$\eta(k,\rho_2) := \left(1+k^{\frac{1}{1+\rho_2}}\right)\left(k^{\frac{\rho_2}{1+\rho_2}}-1\right). \tag{3.33}$$

Then, (3.30) equals to the product

$$\frac{\partial}{\partial\rho_2}\frac{g'(\rho_2,z)}{g'(\rho_1,z)} = \Phi(\lambda(z),\rho_1,\rho_2)\Psi(\lambda(z),\rho_2).$$

Let $k=\lambda(z)\in[0,1]$. As $k=\lambda(z)$ is decreasing in z, to check the monotonicity of the above expression with respect to z, we can equivalently check the monotonicity of $\Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)$ with respect to k.

Taking the derivative with respect to k gives

$$\begin{aligned}
&\frac{\partial \Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)}{\partial k} \\
=&\Phi'(k,\rho_1,\rho_2)\Psi(k,\rho_2) + \Phi(k,\rho_1,\rho_2)\Psi'(k,\rho_2) \\
=&\Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)\left(\frac{\partial \ln \Phi(k,\rho_1,\rho_2)}{\partial k} + \frac{\Psi'(k,\rho_2)}{\Psi(k,\rho_2)}\right),
\end{aligned} \tag{3.34}$$

where $\Phi'(k,\rho_1,\rho_2) = \dfrac{\partial \Phi(k,\rho_1,\rho_2)}{\partial k}$ and $\Psi'(k,\rho) = \dfrac{\partial \Psi(k,\rho)}{\partial k}$.

Now, we derive the expressions in (3.34):

$$\begin{aligned}
\ln \Phi(k,\rho_1,\rho_2) = &\rho_2 \ln\left(\frac{1+k^{\frac{1}{1+\rho_2}}}{2}\right) + \ln\left(1-k^{\frac{-\rho_2}{1+\rho_2}}\right) \\
&- \rho_1 \ln\left(\frac{1+k^{\frac{1}{1+\rho_1}}}{2}\right) - \ln\left(1-k^{\frac{-\rho_1}{1+\rho_1}}\right),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln \Phi(k,\rho_1,\rho_2)}{\partial k} =&\frac{\rho_2}{1+\rho_2}\frac{k^{\frac{-\rho_2}{1+\rho_2}}}{1+k^{\frac{1}{1+\rho_2}}} + \frac{\rho_2}{1+\rho_2}\frac{k^{\frac{-\rho_2}{1+\rho_2}-1}}{1-k^{\frac{-\rho_2}{1+\rho_2}}} \\
&- \frac{\rho_1}{1+\rho_1}\frac{k^{\frac{-\rho_1}{1+\rho_1}}}{1+k^{\frac{1}{1+\rho_1}}} - \frac{\rho_1}{1+\rho_1}\frac{k^{\frac{-\rho_1}{1+\rho_1}-1}}{1-k^{\frac{-\rho_1}{1+\rho_1}}} \\
=&\frac{\rho_2}{1+\rho_2}\frac{1+k}{k\left(1+k^{\frac{1}{1+\rho_2}}\right)\left(k^{\frac{\rho_2}{1+\rho_2}}-1\right)} \\
&- \frac{\rho_1}{1+\rho_1}\frac{1+k}{k\left(1+k^{\frac{1}{1+\rho_1}}\right)\left(k^{\frac{\rho_1}{1+\rho_1}}-1\right)} \\
=&F(k,\rho_2) - F(k,\rho_1),
\end{aligned}$$

where

$$F(k,\rho) := \frac{\rho}{1+\rho}\frac{1+k}{k}\frac{1}{\eta(k,\rho)}, \tag{3.35}$$

and

$$
\begin{aligned}
&\Psi'(k,\rho_2)\\
=&\frac{\partial}{\partial k}\left(\ln\left(\frac{1+k^{\frac{1}{1+\rho_2}}}{2}\right)+\frac{\ln k}{(1+\rho_2)^2}\left(-\frac{\rho_2 k^{\frac{1}{1+\rho_2}}}{1+k^{\frac{1}{1+\rho_2}}}+\frac{1}{k^{\frac{\rho_2}{1+\rho_2}}-1}\right)\right)\\
=&\ \frac{k^{-\frac{\rho_2}{1+\rho_2}}}{(1+\rho_2)(1+k^{\frac{1}{1+\rho_2}})}+\frac{1}{(1+\rho_2)^2 k}\left(-\rho_2\frac{k^{\frac{1}{1+\rho_2}}}{1+k^{\frac{1}{1+\rho_2}}}+\frac{1}{k^{\frac{\rho_2}{1+\rho_2}}-1}\right)\\
&+\frac{\ln k}{(1+\rho_2)^2}\left(-\frac{\rho_2 k^{-\frac{\rho_2}{1+\rho_2}}}{(1+\rho_2)\left(1+k^{\frac{1}{1+\rho_2}}\right)^2}-\frac{\rho_2 k^{-\frac{1}{1+\rho_2}}}{(1+\rho_2)\left(k^{\frac{\rho_2}{1+\rho_2}}-1\right)^2}\right)\\
=&\ \frac{k+1}{(1+\rho_2)^2\ \ k\ \ \eta(k,\rho_2)}-\frac{\rho_2\,(k+1)\left(k^{\frac{\rho_2}{1+\rho_2}}+k^{\frac{1}{1+\rho_2}}\right)\ln k}{(1+\rho_2)^3\ \ k\ \ \eta^2(k,\rho_2)}\\
=&\ \frac{k+1}{(1+\rho_2)^2\ \ k\ \ \eta(k,\rho_2)^2}\left(\eta(k,\rho_2)-\left(k^{\frac{\rho_2}{1+\rho_2}}+k^{\frac{1}{1+\rho_2}}\right)\ln k^{\frac{\rho_2}{1+\rho_2}}\right), \qquad (3.36)
\end{aligned}
$$

where $\eta(k,\rho)$ is defined in (3.33).

To summarize the steps so far, we have shown that the second derivative of $\tilde{f}_{\rho_1,\rho_2}(t)$ with respect to t is given by

$$
\begin{aligned}
\frac{\partial^2\tilde{f}_{\rho_1,\rho_2}(t)}{\partial t^2}&=\frac{\partial}{\partial t}\frac{\partial}{\partial\rho_2}\frac{g'(\rho_2,g^{-1}(\rho_1,t))}{g'(\rho_1,g^{-1}(\rho_1,t))}\\
&=\frac{\partial}{\partial z}\left(\frac{\partial}{\partial\rho_2}\frac{g'(\rho_2,z)}{g'(\rho_1,z)}\right)\frac{\partial z}{\partial t}\\
&=\frac{\partial\Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)}{\partial k}\frac{\partial k}{\partial z}\frac{\partial z}{\partial t}
\end{aligned}
$$

where $z=g^{-1}(\rho_1,t)$, $k=\lambda(z)$ with $\lambda(z)$ defined in (3.29), $\Phi(k,\rho_1,\rho_2)$ given by (3.31), and $\Psi(k,\rho_2)$ given by (3.32).

We first prove the claims of the lemma for $\rho_1=\rho_2=\rho$. Coming back to (3.34),

$$
\frac{\partial\Phi(k,\rho,\rho)\Psi(k,\rho)}{\partial k}=\Psi'(k,\rho)
$$

as $\Phi(k,\rho,\rho)=1$ and $\dfrac{\partial\ln\Phi(k,\rho,\rho)}{\partial k}=0$. Hence to prove the convexity claims, we need to investigate the sign of $\Psi'(k,\rho)$ we derived in (3.36). Note that the factor in front of the parenthesis in (3.36) is always positive for $k\in[0,1]$, $\rho_2\in\mathbb{R}\setminus\{-1\}$, and the term inside the parenthesis equals to the function $m(k,\rho_2)$ defined in Lemma 3.12 in Appendix 3.C. So the sign of $\Psi'(k,\rho_2)$ is determined by the sign of $m(k,\rho_2)$.

By Lemma 3.12 , we have

$$\begin{aligned}
&\Psi'(k,\rho_2) \geq 0, \quad \forall \rho_2 < -1\\
&\Psi'(k,\rho_2) \leq 0, \quad \forall \rho_2 \in (-1,0),\\
&\Psi'(k,0) = 0,\\
&\Psi'(k,\rho_2) \leq 0, \quad \forall \rho_2 \in (0,\rho^*(k)),\\
&\Psi'(k,\rho^*(k)) = 0,\\
&\Psi'(k,\rho_2) \geq 0, \quad \forall \rho_2 \geq \rho^*(k).
\end{aligned}$$

where $\rho^*(k) \geq 3$ is a constant which depends on $k \in [0,1]$. As k is decreasing in z, which is, by Lemma 2.2, non-increasing in t when $\rho \geq 0$, we have

$$\frac{\partial^2 \tilde{f}_\rho(t)}{\partial t^2} = \underbrace{\Psi'(k,\rho)}_{\leq 0} \underbrace{\frac{\partial k}{\partial z}}_{<0} \underbrace{\frac{\partial z}{\partial t}}_{\leq 0} \leq 0,$$

for $\rho \in [0,\rho^*(k)]$, and

$$\frac{\partial^2 \tilde{f}_\rho(t)}{\partial t^2} = \underbrace{\Psi'(k,\rho)}_{\geq 0} \underbrace{\frac{\partial k}{\partial z}}_{<0} \underbrace{\frac{\partial z}{\partial t}}_{\leq 0} \geq 0,$$

for $\rho \geq \rho^*(k)$. Hence, the function $\tilde{f}_\rho(t)$ is concave in t when $\rho \in (0,3]$ as claimed. On the other hand, we know by Lemma 2.2 that z is non-decreasing in t when $\rho \in (-1,0)$. Hence, the function $\tilde{f}_\rho(t)$ is convex in t whenever $\rho \in (-1,0)$. Finally, when $\rho < -1$, z is non-increasing in t by Lemma 2.2, so that $\tilde{f}_\rho(t)$ is convex in t.

To prove the last claim of the lemma concerning the case where $\rho_1, \rho_2 \in (0,1]$ such that $\rho_1 < \rho_2$, we need to determine the sign of $\Psi(k,\rho_2)$. Note that, $\Psi'(k,\rho) \leq 0$ for $\rho \in (0,3]$ implies

$$\Psi(k,\rho) \geq \lim_{k\to 1} \Psi(1,\rho) = \frac{2}{(1+\rho)^2} \lim_{k\to 1} \frac{\ln k}{\eta(k,\rho)} = \frac{1}{\rho\,(1+\rho)} \geq 0,$$

since

$$\begin{aligned}
\lim_{k\to 1} \frac{\ln k}{\eta(k,\rho)} &= \frac{0}{0}\\
&= \lim_{k\to 1} \frac{\partial \ln k/\partial k}{\partial \eta(k,\rho)/\partial k} = \lim_{k\to 1} \frac{k+\rho k}{k\left(k+\rho k - k^{\frac{1}{1+\rho}} + \rho k^{\frac{\rho}{1+\rho}}\right)} = \frac{1+\rho}{2\rho}.
\end{aligned}$$

As a result, $\Psi(k,\rho_2) \geq 0$ whenever $\rho_2 \in (0,3]$. Recall that we are interested in the

sign of the following expression

$$\frac{\partial \Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)}{\partial k} = \Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)\left(\frac{\Psi'(k,\rho_2)}{\Psi(k,\rho_2)} + F(k,\rho_2) - F(k,\rho_1)\right).$$

Lemma 3.13 in Appendix 3.C shows that the function $F(k,\rho)$ is decreasing in $\rho \in (0,1]$. Moreover, we have just shown $\dfrac{\Psi'(k,\rho_2)}{\Psi(k,\rho_2)} \leq 0$, for $\rho_2 \in (0,3]$. Consequently, when $\rho_1, \rho_2 \in (0,1]$ such that $\rho_1 \leq \rho_2$, we have

$$\frac{\Psi'(k,\rho_2)}{\Psi(k,\rho_2)} + F(k,\rho_2) - F(k,\rho_1) \leq 0,$$

and the product $\Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)$ is non-increasing in k. As k is decreasing in z, which is, by Lemma 2.2, non-increasing in t when $\rho \geq 0$, the expression in (3.30), is decreasing in z whenever $\rho_1, \rho_2 \in (0,1]$ such that $\rho_1 \leq \rho_2$. In this case,

$$\frac{\partial^2 \tilde{f}_{\rho_1,\rho_2}(t)}{\partial t^2} = \underbrace{\frac{\partial \Phi(k,\rho_1,\rho_2)\Psi(k,\rho_2)}{\partial k}}_{\leq 0}\underbrace{\frac{\partial k}{\partial z}}_{<0}\underbrace{\frac{\partial z}{\partial t}}_{\leq 0} \leq 0$$

holds. Hence, the function $\tilde{f}_{\rho_1,\rho_2}(t)$ is concave in t as claimed. □

3.B Proof of Lemma 3.7

Proof. Taking the first derivative of $f_{\rho_1,\rho_2}(t)$ with respect to t, we obtain

$$\begin{aligned}\frac{\partial f_{\rho_1,\rho_2}(t)}{\partial t} &= \frac{\partial g(\rho_2, g^{-1}(\rho_1,t))}{\partial t} \\ &= \frac{g'(\rho_2, g^{-1}(\rho_1,t))}{g'(\rho_1, g^{-1}(\rho_1,t))}. \end{aligned} \tag{3.37}$$

Let $z = g^{-1}(\rho_1,t)$. As $g(\rho,z)$ is a monotone function in z by Lemma 2.2, so is $z = g^{-1}(\rho,t)$ in t. Hence we can check the convexity of $f_{\rho_1,\rho_2}(t)$ with respect to t, from the monotonicity with respect to z of the expression $g'(\rho_2,z)/g'(\rho_1,z)$.

Taking the derivative with respect to z, we get

$$\begin{aligned}\frac{\partial}{\partial z}\frac{g'(\rho_2,z)}{g'(\rho_1,z)} &= \frac{\partial}{\partial z} 2^{\rho_1-\rho_2}\frac{\alpha(\rho_2,z)\beta(\rho_2,z)}{\alpha(\rho_1,z)\beta(\rho_1,z)} \\ &= \frac{2^{-\rho_2}\alpha(\rho_2,z)\beta(\rho_2,z)}{2^{-\rho_1}\alpha(\rho_1,z)\beta(\rho_1,z)}\left(\ell(\rho_2,z) - \ell(\rho_1,z)\right),\end{aligned}$$

where

$$\ell(\rho, z) := \frac{\partial \alpha(\rho, z)/\partial z}{\alpha(\rho, z)} + \frac{\partial \beta(\rho, z)/\partial z}{\beta(\rho, z)}.$$

One can easily check that $\alpha(\rho, z) \geq 0$, for any $\rho > -1$, and while $\beta(\rho, z) \geq 0$, for $\rho \in (-1, 0)$, we have $\beta(\rho, z) \leq 0$, for $\rho \geq 0$. Moreover, we claim that

$$\ell(\rho_2, z) - \ell(\rho_1, z) \geq 0 \tag{3.38}$$

holds when $\rho_1 \in (-1, 0)$ and $\rho_2 \geq 0$, or when $\rho_1 \in (0, 1]$ and $\rho_2 \geq \rho_1$, and we claim that

$$\ell(\rho_2, z) - \ell(\rho_1, z) \leq 0$$

holds when $\rho_1 > 1$ and $\rho_2 \in (-1, 0)$, or when $\rho_1 > 1$ and $\rho_2 \in (0, 1]$. Therefore, by this claim, if $\rho_1 \in (-1, 0]$ and $\rho_2 \geq 0$, we have

$$\frac{\partial}{\partial z} \frac{g'(\rho_2, z)}{g'(\rho_1, z)} = \underbrace{\frac{2^{-\rho_2}\alpha(\rho_2, z)\beta(\rho_2, z)}{2^{-\rho_1}\alpha(\rho_1, z)\beta(\rho_1, z)}}_{\leq 0} \underbrace{(\ell(\rho_2, z) - \ell(\rho_1, z))}_{\geq 0} \leq 0,$$

and if $\rho_1 \in [0, 1]$ and $\rho_2 \geq \rho_1$, we have

$$\frac{\partial}{\partial z} \frac{g'(\rho_2, z)}{g'(\rho_1, z)} = \underbrace{\frac{2^{-\rho_2}\alpha(\rho_2, z)\beta(\rho_2, z)}{2^{-\rho_1}\alpha(\rho_1, z)\beta(\rho_1, z)}}_{\geq 0} \underbrace{(\ell(\rho_2, z) - \ell(\rho_1, z))}_{\geq 0} \geq 0.$$

On the other hand, if $\rho_1 > 1$ and $\rho_2 \in (-1, 0)$, we have

$$\frac{\partial}{\partial z} \frac{g'(\rho_2, z)}{g'(\rho_1, z)} = \underbrace{\frac{2^{-\rho_2}\alpha(\rho_2, z)\beta(\rho_2, z)}{2^{-\rho_1}\alpha(\rho_1, z)\beta(\rho_1, z)}}_{\leq 0} \underbrace{(\ell(\rho_2, z) - \ell(\rho_1, z))}_{\leq 0} \geq 0,$$

and if $\rho_1 > 1$ and $\rho_2 \in (0, 1]$, we have

$$\frac{\partial}{\partial z} \frac{g'(\rho_2, z)}{g'(\rho_1, z)} = \underbrace{\frac{2^{-\rho_2}\alpha(\rho_2, z)\beta(\rho_2, z)}{2^{-\rho_1}\alpha(\rho_1, z)\beta(\rho_1, z)}}_{\geq 0} \underbrace{(\ell(\rho_2, z) - \ell(\rho_1, z))}_{\leq 0} \leq 0.$$

Recall that we are interested in the sign of the second derivative of f_{ρ_1,ρ_2} with respect to t given by

$$\begin{aligned}\frac{\partial^2 f_{\rho_1,\rho_2}(t)}{\partial t^2} &= \frac{\partial}{\partial t} \frac{\partial g(\rho_2, g^{-1}(\rho_1, t))}{\partial t} \\ &= \frac{\partial}{\partial z} \frac{g'(\rho_2, z)}{g'(\rho_1, z)} \frac{\partial z}{\partial t}.\end{aligned}$$

As by Lemma 2.2, z is non-decreasing in t for $\rho_1 \in (-1, 0)$ and non-increasing for $\rho_1 \geq 0$, the function $f_{\rho_1,\rho_2}(t)$ is concave in t when $\rho_1 \in (-1, 0]$ and $\rho_2 \geq 0$, or when $\rho_1 \in [0, 1]$ and $\rho_2 \geq \rho_1$, or when $\rho_1 > 1$ and $\rho_2 \in (-1, 0)$, and convex when $\rho_1 > 1$ and $\rho_2 \in (0, 1]$.

Now, we prove the claim in (3.38). For that purpose, we show that the function $\ell(\rho.z)$ is non-decreasing in ρ for the interval $\rho \in (-1, 3)$, and $\dfrac{\partial \ell(\rho, z)}{\partial \rho}$ changes sign only once after $\rho \geq 3$. As

$$\lim_{\rho \to 1} \ell(\rho, z) = \lim_{\rho \to \infty} \ell(\rho, z) = \frac{1}{z - z^3}$$

holds, we can then conclude that

$$\begin{aligned} \ell(\rho, z) &\geq \ell(1, z), \quad \text{when } \rho \geq 1, \\ \ell(\rho, z) &\leq \ell(1, z), \quad \text{when } \rho \in (-1, 1]. \end{aligned}$$

The above inequalities ensure $\ell(\rho_2, z) - \ell(\rho_1, z) \geq 0$ when $\rho_1 \in (-1, 0)$ and $\rho_2 \geq 0$, or when $\rho_1 \in [0, 1]$ and $\rho_2 \geq \rho_1$. Similarly, the previous arguments ensure that $\ell(\rho_2, z) - \ell(\rho_1, z) \leq 0$ when $\rho_1 > 1$ and $\rho_2 \in (-1, 1]$.

Note that

$$\ell(\rho, z) = \frac{\partial}{\partial z} \left(\ln \left(2^{-\rho} \alpha(\rho, z) \right) + \ln \left(\beta(\rho, z) \right) \right).$$

Hence,

$$\begin{aligned} \frac{\partial \ell(\rho, z)}{\partial \rho} &= \frac{\partial}{\partial z} \left(\frac{\partial 2^{-\rho} \alpha(\rho, z) / \partial \rho}{2^{-\rho} \alpha(\rho, z)} + \frac{\partial \beta(\rho, z) / \partial \rho}{\beta(\rho, z)} \right) \\ &= \frac{\partial \Psi(k, \rho)}{\partial k} \frac{\partial k}{\partial z} \\ &= \Psi'(k, \rho) \frac{\partial k}{\partial z} \end{aligned}$$

where $k = \lambda(z)$ is defined in (3.29) and $\Psi'(k, \rho)$ is defined in (3.36). Luckily, we have already investigated the sign of $\Psi'(k, \rho)$ in the proof of Lemma 3.6 stated in the previous appendix. Indeed, we have shown that $\Psi'(k, \rho) \leq 0$, for $\rho \in (-1, 3)$, and the function changes sign only once after $\rho \geq 3$. As k is decreasing in z, the sign of $\dfrac{\partial \ell(\rho, z)}{\partial \rho}$ is exactly the opposite of $\Psi'(k, \rho)$. This concludes the proof. □

3.C Lemma 3.12 and Lemma 3.13

Lemma 3.12. *For $k \in [0,1]$, we define*

$$m(k,\rho) := -1 + k - k^{\frac{1}{1+\rho}} + k^{\frac{\rho}{1+\rho}} - \left(k^{\frac{\rho}{1+\rho}} + k^{\frac{1}{1+\rho}}\right) \ln k^{\frac{\rho}{1+\rho}}. \tag{3.39}$$

Then, for $\forall k \in [0,1]$, we have

$$\begin{aligned} &m(k,\rho) \geq 0, \quad \forall \rho < -1, \\ &m(k,\rho) \leq 0, \quad \forall \rho \in (-1,0), \\ &m(k,0) = 0. \end{aligned}$$

Moreover, $\exists\ \rho^*(k) \geq 3$ which depends on k such that:

$$\begin{aligned} &m(k,\rho) \leq 0, \quad \forall \rho \in (-1,\rho^*(k)), \\ &m(k,\rho^*) = 0, \\ &m(k,\rho) \geq 0, \quad \forall \rho \in (\rho^*,\infty). \end{aligned}$$

Proof. We now follow a series of transformations. Let $t = \dfrac{\rho}{1+\rho}$. Then, (3.39) reduces to

$$m\left(k, \frac{t}{1-t}\right) = -1 + k - k^{1-t} + k^t - (k^t + k^{1-t}) \ln k^t.$$

In addition, let $s = -t \ln k$. Then,

$$m\left(k, \frac{-s}{\ln k + s}\right) = -1 + k - ke^s + e^{-s} + s(e^{-s} + ke^s).$$

We first note that the function is zero at $s = 0$. Taking the first derivative with respect to s, we get

$$\begin{aligned} \frac{\partial}{\partial s} m\left(k, \frac{-s}{\ln k + s}\right) &= -ke^s - e^{-s} + e^{-s} + ke^s + s(-e^{-s} + ke^s) \\ &= s(-e^{-s} + ke^s) \\ &= t(k^t - k^{1-t}) \ln k. \end{aligned}$$

Hence the function $m\left(k, \dfrac{-s}{\ln k + s}\right)$ is non-increasing in s for $t \in [0, 1/2]$, and non-decreasing otherwise. Moreover, the derivative of $m\left(k, \dfrac{t}{1-t}\right)$ with respect to t is given by

$$\frac{\partial}{\partial t} m\left(k, \frac{t}{1-t}\right) = \frac{\partial}{\partial s} m\left(k, \frac{-s}{\ln k + s}\right) \frac{\partial s}{\partial t}.$$

As s is non-decreasing in t, we have shown that $m\left(k, \frac{t}{1-t}\right)$ is non-increasing in t for $t \in [0, 1/2]$, and non-decreasing otherwise. Similarly, the derivative of $m(k, \rho)$ with respect to ρ is given by

$$\frac{\partial m\left(k, \rho\right)}{\partial \rho} = \frac{\partial}{\partial t} m\left(k, \frac{t}{1-t}\right) \frac{\partial t}{\partial \rho}.$$

As t is increasing in ρ for the intervals $(-\infty, -1)$, and $(-1, \infty)$, $m(k, \rho)$ will be non-increasing in ρ for $t \in [0, 1/2]$, and non-decreasing otherwise. We simply need to map this result to the claims of the lemma in terms of the intervals defined by ρ.

For the interval $t \in [1, \infty)$, we have $\rho < -1$, and $m(k, \rho)$ is non-decreasing in ρ. Moreover,

$$\lim_{\rho \to -\infty} m(k, \rho) = (-1 + k - 1 + k) - (k+1)\ln k = -2(1-k) + (k+1)\ln k \geq 0,$$

where the sign follows by noting that at $k = 1$ the expression evaluates to 0, and it is non-increasing in k as

$$\frac{\partial}{\partial k}\left(-2(1-k) + (k+1)\ln k\right) = 1 - \frac{1}{k} + \ln \frac{1}{k} \leq 0$$

follows by using the inequality $\ln x \leq x - 1$. This shows $m(k, \rho) \geq 0$ for $\rho < -1$.

For the interval $t \in (-\infty, 0]$, we have $\rho \in (-1, 0]$, and $m(k, \rho)$ is non-decreasing in ρ. As we have $m(k, 0) = 0$, we conclude $m(k, \rho) \leq 0$ for $\rho \in (-1, 0)$.

For the interval $t \in [0, 1/2]$, we have $\rho \in [0, 1]$, and $m(k, \rho)$ is non-increasing in ρ. As we have $m(k, 0) = 0$, we conclude $m(k, \rho) \leq 0$ for $\rho \in (0, 1]$.

For the interval $t \in [1/2, 1]$, we have $\rho \geq 1$, and $m(k, \rho)$ is non-decreasing in ρ. As $m(k, 1) \leq 0$, and

$$\lim_{\rho \to \infty} m(k, \rho) = (-1 + k - 1 + k) - (k+1)\ln k = -2(1-k) + (k+1)\ln k \geq 0,$$

the function will eventually cross zero. Now, we prove that the crossing point ρ^*, i.e., $m(k, \rho^*) = 0$, is such that $\rho^* \geq 3$. For that purpose, we only need to show that $m(k, 3)$ is increasing in k because $m(1, 3) = 0$ holds.

Taking the first derivative with respect to k, we get

$$\frac{\partial m(k, 3)}{\partial k} = \frac{4(-1 + k^{3/4}) - 3/4(1 + 3\sqrt{k})\ln k}{4k^{3/4}} \geq 0,$$

where equality holds if and only if $k = 1$. The sign follows by noting that the

denominator is positive, the numerator is decreasing in k, and is equal to 0 if and only if $k = 1$. Indeed, taking the first derivative with respect to k of the numerator, we get

$$\frac{\partial}{\partial k}\left(4(-1+k^{3/4}) - 3/4(1+3\sqrt{k})\ln k\right) = \frac{-3(2+6\sqrt{k}-8k^{3/4}+3\sqrt{k}\ln k)}{8k} \leq 0,$$

where equality holds if and only if $k = 1$. The sign follows by noting that the denominator is positive, the numerator is increasing in k, and is equal to 0 if and only if $k = 1$. To see this, once more we take the first derivative with respect to k of the numerator. Then, we get

$$\frac{\partial}{\partial k}\left(-3(2+6\sqrt{k}-8k^{3/4}+3\sqrt{k}\ln k)\ln k\right) = \frac{-9(4-4k^{1/4}+\ln k)}{2\sqrt{k}} \geq 0,$$

where equality holds if and only if $k = 1$. The sign follows by noting that the denominator is positive, the numerator is decreasing in k, and is equal to 0 if and only if $k = 1$. To show this, we need to take the first derivative with respect to k of the numerator one last time. Doing so, we get

$$\frac{\partial}{\partial k}\left(-9(4-4k^{1/4}+\ln k)\right) = \frac{9(-1+k^{1/4})}{k} \leq 0,$$

for $k \in [0, 1]$, and where equality holds if and only if $k = 1$. This concludes the proof of the lemma. □

Lemma 3.13. *The function $F(k, \rho)$ defined in* (3.35) *is a decreasing function in $\rho \in [0, 1]$.*

Proof. For convenience, we define the function

$$H(k, \rho) := -\frac{k}{1+k}F(k, \rho).$$

Then,

$$H(k, \rho) = \frac{\rho}{1+\rho}\frac{1}{\left(1+k^{\frac{1}{1+\rho}}\right)\left(1-k^{\frac{\rho}{1+\rho}}\right)} \geq 0, \tag{3.40}$$

where $k \in [0, 1]$. We note that instead of $F(k, \rho)$, we can check the monotonicity of $H(k, \rho)$ with respect to ρ. We now follow a series of transformations. Let $t = \dfrac{\rho}{1+\rho}$, for $t \in [0, \frac{1}{2}]$. Then, (3.40) reduces to

$$H\left(k, \frac{t}{1-t}\right) = \frac{t}{(1-k^t)(1+k^{1-t})}.$$

In addition, let $s = -t \ln k$, for $s \in [0, \frac{1}{2} \ln \frac{1}{k}]$. Then,

$$H\left(k, \frac{-s}{\ln k + s}\right) = \frac{1}{\ln \frac{1}{k}} \ \frac{s}{1 - e^{-s}} \ \frac{1}{1 + ke^{s}}. \tag{3.41}$$

We note that the first fraction in (3.41) can be treated as a constant and we ignore it. We define the variable $a = \frac{1}{k} \geq 1$. For simplicity, we consider the function

$$\frac{1}{H\left(k, \frac{-s}{\ln k + s}\right)} = \underbrace{\frac{\ln a}{a}}_{constant} \ \frac{1 - e^{-s}}{s} \left(a + e^{s}\right).$$

We first show that $\ln\left(\frac{1 - e^{-s}}{s}\left(a + e^{s}\right)\right)$ is a convex function for all $s \geq 0$. Taking the first derivative with respect to s, we obtain

$$\frac{\partial}{\partial s}\left(-\ln s + \ln\left(\frac{1}{1 - e^{-s}}\right) + \ln\left(\frac{e^{s}}{a + e^{s}}\right)\right) = -\frac{1}{s} + \frac{e^{s}}{a + e^{s}} + \frac{1}{e^{s} - 1}. \tag{3.42}$$

Taking the second derivative in s, we get

$$\begin{aligned}
&\frac{\partial^2}{\partial s^2}\left(-\ln s + \ln\left(\frac{1}{1 - e^{-s}}\right) + \ln\left(\frac{e^{s}}{a + e^{s}}\right)\right) \\
=&\frac{1}{s^2} + \frac{ae^{s}}{(a + e^{s})^2} - \frac{e^{s}}{(e^{s} - 1)^2} \\
\geq&\frac{1}{s^2} - \frac{e^{s}}{(e^{s} - 1)^2} \\
=&\frac{1}{s^2} - \left(\frac{1}{e^{\frac{s}{2}} + e^{\frac{-s}{2}}}\right)^2 \\
=&\frac{1}{s^2} - \frac{1}{\left(2 \sinh \frac{s}{2}\right)^2} \geq 0,
\end{aligned}$$

where the non-negativity follows from $\sinh x \geq x$, for $x \geq 0$. We proved that $\ln\left(\frac{1 - e^{-s}}{s}\left(a + e^{s}\right)\right)$ is a convex function for all $s \geq 0$. Therefore the function has only one minimum, and to decide whether the expression is decreasing in $s \in [0, \frac{1}{2} \ln a]$, it is sufficient to evaluate (3.42) at $s = \frac{1}{2} \ln a$:

$$\begin{aligned}
&\frac{\partial}{\partial s}\left(-\ln s + \ln\left(\frac{1}{1 - e^{-s}}\right) + \ln\left(\frac{e^{s}}{a + e^{s}}\right)\right)\Big|_{s = \frac{1}{2} \ln a} \\
=& -\frac{1}{\ln \sqrt{a}} + \frac{\sqrt{a}}{a + \sqrt{a}} + \frac{1}{\sqrt{a} - 1} \\
=& -\frac{1}{\ln \sqrt{a}} + \frac{2\sqrt{a}}{a - 1} \leq 0,
\end{aligned}$$

since for $b = \sqrt{a} \geq 1$, we can show that

$$\frac{b^2 - 1}{2b} - \ln b \geq 0. \tag{3.43}$$

Taking the first derivative of (3.43) with respect to b, we get

$$\frac{\partial}{\partial b}\frac{b^2 - 1}{2b} - \ln b = \frac{1}{2} + \frac{1}{2b^2} - \frac{1}{b} = \frac{(b-1)^2}{2b^2} \geq 0.$$

Therefore, we proved that for each $k \in [0, 1]$ the function $1/H\left(k, \frac{-s}{\ln k + s}\right)$ is decreasing in s. By definition, the variable t is increasing in ρ, and $s = -t \ln k$ is also increasing in t for a given k. As a consequence, the function

$$F(k, \rho) = -\frac{1+k}{k} H(k, \rho) \tag{3.44}$$

is decreasing in ρ. □

Chapter 4

Polarization for E_0

Arıkan's polar codes [2] are constructed by the repeated application of the *polar transform*. From two independent copies of a given binary input channel $W\colon \mathbb{F}_2 \to \mathcal{Y}$, this transform synthesizes two new binary input channels $W^-\colon \mathbb{F}_2 \to \mathcal{Y}^2$ and $W^+\colon \mathbb{F}_2 \to \mathcal{Y}^2 \times \mathbb{F}_2$. The transition probabilities of these channels are given by [2]

$$W^-(y_1y_2|u_1) := \sum_{u_2\in\mathbb{F}_2} \frac{1}{2} W(y_1|u_1 \oplus u_2)W(y_2|u_2), \tag{4.1}$$

$$W^+(y_1y_2u_1|u_2) := \frac{1}{2} W(y_1|u_1 \oplus u_2)W(y_2|u_2). \tag{4.2}$$

Being binary input channels themselves, W^- and W^+ are in turn candidates for the polar transform. One may thus apply the polar transform to these new channels to obtain the channels $W^{--} := (W^-)^-$, $W^{-+} := (W^-)^+$, $W^{+-} := (W^+)^-$, and $W^{++} := (W^+)^+$. More generally, the repeated application will yield at stage n, a set of 2^n channels

$$\left\{W^{s^n} :\ s^n \in \{+,-\}^n\right\}. \tag{4.3}$$

In analyzing the properties of these channels, it is useful to introduce an auxiliary stochastic process. Let $(\Omega, \mathscr{F}, P)$ denote a probability space. Assume the random sequence $B_1, \ldots, B_n$ defined on this space is drawn i.i.d. according to a Bernoulli distribution with probabilities equal to $1/2$. Let $\mathscr{F}_n$ be the σ-algebra generated by this Bernoulli sequence. The *channel polarization process* W_n is defined in [24] by $W_0 = W$ and

$$W_{n+1} := \begin{cases} W_n^- & \text{if } B_{n+1} = 1 \\ W_n^+ & \text{if } B_{n+1} = 0 \end{cases}, \quad \text{for } n \geq 0. \tag{4.4}$$

In this way, W_n is uniformly distributed over the set of 2^n channels. Each path

realization W^{s^n} of this process corresponds exactly to one of the synthetic channels $W_N^{(i)}$, for $i = 1, \ldots, N$, introduced earlier in the Introduction.

Due to the exponential growth of the output alphabets of the synthetic channels, tracking the evolution of the channel polarization process directly is not so easy. As a result, the properties of polar codes are analyzed by examining additional random processes that follow the evolution of information measures as the underlying communication channel undergoes the sequence of polar transformations. One such measure of information is the *symmetric capacity process* $I_n(W) := I(W_n)$.

Let us now retell the famous polarization 'tale' of this most famous information theoretical quantity: Arıkan proves in [2, Propositions 8 and 10] that the process $I_n(W)$ is a bounded martingale on the interval $[0, 1]$, converges *almost surely* (a.s.) to a random variable I_∞ such that $\mathbb{E}[I_\infty] = I_0$, where I_∞ takes values a.s. in $\{0, 1\}$. These prove that the recursive application of the polar transform leads to channel polarization, see Definition 1.1.

Due to the recursive construction procedure, the properties of the single step polar transform are essential in proving the convergence result related to I_n and furthermore in shaping the theory of channel polarization and polar codes. Below is a list of three fundamental properties of the mapping $(W, W) \to (W^-, W^+)$ which will help us understand better how the polar transform works and how polarization happens.

$\mathcal{P}1$. ***Polarization Property***: By [2, Proposition 4],

$$I(W^-) \leq I(W) \leq I(W^+). \tag{4.5}$$

$\mathcal{P}2$. ***Conservation Property***: By the same proposition,

$$I(W^-) + I(W^+) = 2I(W). \tag{4.6}$$

$\mathcal{P}3$. ***Extremality Property***: It is well known that (proved as a corollary to extremes of information combining [25]) among all channels W with a given symmetric capacity $I(W)$, the BEC and the BSC polarize most and least in the sense of having the largest and the smallest differences between $I(W^+)$ and $I(W^-)$. So,

$$I(BSC^+) - I(BSC^-) \leq I(W^+) - I(W^-) \leq I(BEC^+) - I(BEC^-). \tag{4.7}$$

Now, let us quickly retrace the proof of the polarization of I_n.

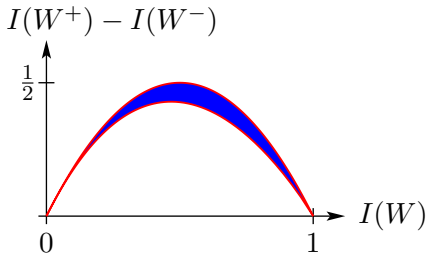

Figure 4.1: $I(W^+) - I(W^-) < \xi$ implies that $I(W) \notin (\gamma, 1-\gamma)$.

$\mathcal{L}1$. ***'Conservation Law'***: The process is a bounded martingale on the interval $[0,1]$ as $I_n \in [0,1]$ and

$$\begin{aligned}&\mathbb{E}[I_n \mid \mathscr{F}_{n-1}]\\ =&\mathbb{E}[I_n \mid B_1, \ldots, B_{n-1}] = \mathbb{E}[I_n \mid W_{n-1}]\\ =&\frac{1}{2}I(W_{n-1}^-) + \frac{1}{2}I(W_{n-1}^+)\\ =&I(W_{n-1}) = I_{n-1}\end{aligned}$$

holds by (4.6). By general results on bounded martingales, I_n converges a.s. to I_∞ such that $\mathbb{E}[I_\infty] = I_0$. It remains to identify the convergence points.

$\mathcal{L}2$. ***'Convergence Law'***: We plotted in Figure 4.1, the range of feasible $I(W)$ versus $I(W^+) - I(W^-)$ pairs using (4.7). From the figure, we see that polarization does strictly happen, i.e., we have strict inequalities in (4.5), as long as the value of $I(W)$ is not at the extremes of its boundaries. So, the synthesized channels keep getting polarized until they become either perfect or completely noisy. Thus, $I_\infty = \{0,1\}$ a.s.

As a consequence, the fraction of the perfect channels must tend to $I(W)$ with the recursive application of the polar transform. By Definition 1.1, channel polarization is attained.

The proof is recollected, but there is more to the story; these nice properties of the transform are not just limited to the symmetric capacity parameter. It was shown in [2, Proposition 5] that the channel *Bhattacharyya distance*,

$$Z(W) := \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}, \tag{4.8}$$

likewise satisfies $Z(W^+) \leq Z(W) \leq Z(W^-)$ and

$$Z(W^+) = Z(BEC^+) = Z(W)^2, \tag{4.9}$$

$$Z(W) \leq Z(W^-) \leq Z(BEC^-) = 2Z(W) - Z(W)^2, \tag{4.10}$$

among all the channels W with a given Bhattacharyya parameter $Z(W)$. Via the cutoff rate relationship

$$R_0(W) = -\log\left(\frac{1+Z(W)}{2}\right), \tag{4.11}$$

it can thus also be inferred that $R_0(W^-) \leq R_0(W) \leq R_0(W^+)$ and the BEC is extremal among all the channels W with a given symmetric cutoff rate $R_0(W)$. Even earlier, in a paper that predates polar coding [4], Arıkan had already shown that the method of channel combining and splitting via this transform improves the symmetric cutoff rate $R_0(W)$, that is[1],

$$R_0(W^-) + R_0(W^+) \geq 2R_0(W). \tag{4.12}$$

Very well, could these be just a coincidence? Recall from Section 2.2 that both the symmetric capacity and the symmetric cutoff rate are quantities that can be obtained as special cases of the information measure $E_0(\rho, W)/\rho$ by

$$R_0(W) = E_0(1, W),$$

and

$$I(W) = \lim_{\rho \to 0} E_0(\rho, W)/\rho,$$

where 'Gallager's E_0' [5, p. 138] evaluated for the uniform input distribution is given by (2.11). Thus, a natural question to ask is if the polar transform improves E_0, which would make (4.12) a special case (and also show that the left hand side of (4.6) is at least as large as the right hand side). Likewise, it is natural to ask whether extremality results similar to (4.7) and the pair (4.9) and (4.10) also hold for this more general channel parameter.

What's Coming, Doc?

In view of the above observations, we will inquire in this chapter how $E_0(\rho, W)$ and consequently $E_0(\rho, W)/\rho$ is affected by the polar transform and its recursive application. During this inquiry, derivations will lead to slightly more general results pertaining to a more general *polar transform* denoted by $\langle W_1, W_2 \rangle^{\pm}$ that synthesizes two channels from two independent (but not necessarily identical) binary input channels $W_1 \colon \mathbb{F}_2 \to \mathcal{Y}_1$ and $W_2 \colon \mathbb{F}_2 \to \mathcal{Y}_2$. Given two such channels, $W_{1,2}^- := \langle W_1, W_2 \rangle^- \colon \mathbb{F}_2 \to \mathcal{Y}_1 \times \mathcal{Y}_2$ and $W_{1,2}^+ := \langle W_1, W_2 \rangle^+ \colon \mathbb{F}_2 \to \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathbb{F}_2$ denote

[1] This conclusion may be derived as a consequence of (4.9) and (4.10), via the relationship (4.11), and the concavity and monotonicity of $\log$.

the synthetic channels with transition probabilities given by

$$W_{1,2}^{-}(y_1y_2|u_1) := \sum_{u_2\in\mathbb{F}_2} \frac{1}{2} W_1(y_1|u_1\oplus u_2)W_2(y_2|u_2), \tag{4.13}$$

$$W_{1,2}^{+}(y_1y_2u_1|u_2) := \frac{1}{2} W_1(y_1|u_1\oplus u_2)W_2(y_2|u_2). \tag{4.14}$$

In this chapter, we will prove that the three properties we have just mentioned can be extended to E_0, and we will discuss a number of implications that follows. More specifically, we will report the following conclusions:

- In Theorem 4.6 and Theorem 4.7, we will prove that channel combining and splitting via Arıkan's polar transform improves Gallager's reliability function E_0 for binary input channels. In addition, we will show that the improvement in E_0 translates to an improvement in the complexity–error probability trade-off and also to a particular chain rule for Rényi's entropies [16].
- In Theorem 4.8, we will prove that amongst all B-DMCs W_1 and W_2 (not necessarily symmetric) with given values of $E_0(\rho, W_1)$ and $E_0(\rho, W_2)$ at a given $\rho \geq 0$, the BECs and the BSCs polarize most and least in the sense of having the largest and the smallest differences between $E_0(\rho, \langle W_1, W_2\rangle^{+})$ and $E_0(\rho, \langle W_1, W_2\rangle^{-})$, for any $\rho \in [0,1]\cup[2,\infty]$. On the other hand, for any $\rho \in [1,2]$, we will show that the BSCs maximize and the BECs minimize the E_0 values obtained after applying either $\langle W_1, W_2\rangle^{+}$ or $\langle W_1, W_2\rangle^{-}$. The theorem will also reveal that besides the special values $\rho = 0, 1$, the value $\rho = 2$ further exhibits an interesting property.
- Using the theorems, Proposition 4.11 will show that the process $E_0(\rho, W_n)/\rho$ is a bounded submartingale converging a.s. to the extremes of the bounded interval $[0,1]$.

4.1 Polarization Property of E_0

The next two lemmas derive suitable expressions for the E_0 parameters of $W_{1,2}^{-}$ and $W_{1,2}^{+}$. The expressions will be similar to the representation given in (2.12).

Lemma 4.1. *Given $W_1\colon \mathbb{F}_2 \to \mathcal{Y}_1$ and $W_2\colon \mathbb{F}_2 \to \mathcal{Y}_2$, let Z_1 and Z_2 be independent random variables taking values in the interval $[0,1]$ such that $E_0(\rho, W_1) = -\log \mathbb{E}\left[g(\rho, Z_1)\right]$ and $E_0(\rho, W_2) = -\log \mathbb{E}\left[g(\rho, Z_2)\right]$ hold for a fixed $\rho \geq 0$ as defined in* (2.12)*. Then,*

$$E_0(\rho, W_{1,2}^{-}) = -\log \mathbb{E}\left[g(\rho, Z_1Z_2)\right], \tag{4.15}$$

where $g(\rho, z)$ given by (2.13)*.*

Proof. Let $y_1 \in \mathcal{Y}_1$ and $y_2 \in \mathcal{Y}_2$. Using (4.13) and the definitions in (2.14) and (2.15), we can write

$$\begin{aligned}
E_0(\rho, W_{1,2}^-) &= -\log \sum_{y_1,y_2} \left[\frac{1}{2} W_{1,2}^-(y_1, y_2 \mid 0)^{\frac{1}{1+\rho}} + \frac{1}{2} W_{1,2}^-(y_1, y_2 \mid 1)^{\frac{1}{1+\rho}}\right]^{1+\rho} \\
&= -\log \sum_{y_1 y_2} \frac{1}{2} q_{W_1}(y_1)\, q_{W_2}(y_2) \times \\
&\qquad \left[\frac{1}{2}\Big((1+\Delta_{W_1}(y_1))\,(1+\Delta_{W_2}(y_2)) \right. \right. \\
&\qquad\qquad \left. \left. + (1-\Delta_{W_1}(y_1))\,(1-\Delta_{W_2}(y_2))\Big)^{\frac{1}{1+\rho}} \right. \\
&\qquad + \frac{1}{2}\Big((1-\Delta_{W_1}(y_1))\,(1+\Delta_{W_2}(y_2)) \\
&\qquad\qquad \left. + (1+\Delta_{W_1}(y_1))\,(1-\Delta_{W_2}(y_2))\Big)^{\frac{1}{1+\rho}}\right]^{1+\rho} \\
&= -\log \sum_{y_1 y_2} q_{W_1}(y_1) q_{W_2}(y_2) g(\rho, \Delta_{W_1}(y_1)\Delta_{W_2}(y_2)),
\end{aligned} \tag{4.16}$$

We define $Z_1 = |\Delta_{W_1}(Y_1)|$ and $Z_2 = |\Delta_{W_2}(Y_2)|$, where Y_1 and Y_2 are independent random variables with distributions q_{W_1} and q_{W_2}, respectively. From this construction, the lemma follows. □

Lemma 4.2. *Given B-DMCs W_1 and W_2, let Z_1 and Z_2 be as in Lemma 4.1. Then,*

$$E_0(\rho, W_{1,2}^+) = -\log \mathbb{E}[h(\rho, Z_1, Z_2)], \tag{4.17}$$

where

$$h(\rho, z_1, z_2) := \frac{1+z_1 z_2}{2} g\Big(\rho, \frac{z_1+z_2}{1+z_1 z_2}\Big) + \frac{1-z_1 z_2}{2} g\Big(\rho, \frac{z_1-z_2}{1-z_1 z_2}\Big) \tag{4.18}$$

is defined for $\rho \geq 0$, $z_1, z_2 \in [-1, 1]$.

Proof. Let $y_1 \in \mathcal{Y}_1$, $y_2 \in \mathcal{Y}_2$, and $u \in \mathbb{F}_2$. Using (4.14), we can write

$$\begin{aligned}
&E_0(\rho, W_{1,2}^+) \\
&= -\log \sum_{y_1,y_2,u} \left[\frac{1}{2} W_{1,2}^+(y_1, y_2, u \mid 0)^{\frac{1}{1+\rho}} + \frac{1}{2} W_{1,2}^+(y_1, y_2, u \mid 1)^{\frac{1}{1+\rho}}\right]^{1+\rho} \\
&= -\log \sum_{y_1,y_2,u} \left[\frac{1}{2}\left(\frac{1}{2} W_1(y_1 \mid u) W_2(y_2 \mid 0)\right)^{\frac{1}{1+\rho}} \right. \\
&\qquad\qquad \left. + \frac{1}{2}\left(\frac{1}{2} W_1(y_1 \mid u \oplus 1) W_2(y_2 \mid 1)\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}.
\end{aligned}$$

Using (2.14) and (2.15), we get

$$
\begin{aligned}
&E_0(\rho, W_{1,2}^{+}) \\
&= -\log \sum_{y_1 y_2} \frac{1}{2} q_{W_1}(y_1) q_{W_2}(y_2) \times \\
&\qquad \Bigg(\Bigg[\frac{1}{2} \big((1 + \Delta_{W_1}(y_1))\,(1 + \Delta_{W_2}(y_2)) \big)^{\frac{1}{1+\rho}} \\
&\qquad\qquad + \frac{1}{2} \big((1 - \Delta_{W_1}(y_1))\,(1 - \Delta_{W_2}(y_2)) \big)^{\frac{1}{1+\rho}} \Bigg]^{1+\rho} \\
&\qquad + \Bigg[\frac{1}{2} \big((1 - \Delta_{W_1}(y_1))\,(1 + \Delta_{W_2}(y_2)) \big)^{\frac{1}{1+\rho}} \\
&\qquad\qquad + \frac{1}{2} \big((1 + \Delta_{W_1}(y_1))\,(1 - \Delta_{W_2}(y_2)) \big)^{\frac{1}{1+\rho}} \Bigg]^{1+\rho} \Bigg) \\
&= -\log \Bigg(\sum_{y_1 y_2} \frac{1}{2} q_{W_1}(y_1) q_{W_2}(y_2) \big(1 + \Delta_{W_1}(y_1) \Delta_{W_2}(y_2) \big) \times \\
&\qquad \Bigg[\frac{1}{2} \left(1 + \frac{\Delta_{W_1}(y_1) + \Delta_{W_2}(y_2)}{1 + \Delta_{W_1}(y_1)\Delta_{W_2}(y_2)} \right)^{\frac{1}{1+\rho}} \\
&\qquad\qquad + \frac{1}{2} \left(1 - \frac{\Delta_{W_1}(y_1) + \Delta_{W_2}(y_2)}{1 + \Delta_{W_1}(y_1)\Delta_{W_2}(y_2)} \right)^{\frac{1}{1+\rho}} \Bigg]^{1+\rho} \\
&\qquad + \sum_{y_1 y_2} \frac{1}{2} q_{W_1}(y_1) q_{W_2}(y_2) \big(1 - \Delta_{W_1}(y_1) \Delta_{W_2}(y_2) \big) \times \\
&\qquad \Bigg[\frac{1}{2} \left(1 - \frac{\Delta_{W_1}(y_1) - \Delta_{W_2}(y_2)}{1 - \Delta_{W_1}(y_1)\Delta_{W_2}(y_2)} \right)^{\frac{1}{1+\rho}} \\
&\qquad\qquad + \frac{1}{2} \left(1 + \frac{\Delta_{W_1}(y_1) - \Delta_{W_2}(y_2)}{1 - \Delta_{W_1}(y_1)\Delta_{W_2}(y_2)} \right)^{\frac{1}{1+\rho}} \Bigg]^{1+\rho} \Bigg) \\
&= -\log \sum_{y_1 y_2} \frac{1}{2} q_{W_1}(y_1) q_{W_2}(y_2) h(\rho, \Delta_{W_1}(y_1), \Delta_{W_2}(y_2))
\end{aligned}
\tag{4.19}
$$

where $h(\rho, z_1, z_2)$ is given by (4.18).

Now, we define $Z_1 = |\Delta_{W_1}(Y_1)|$ and $Z_2 = |\Delta_{W_2}(Y_2)|$, where Y_1 and Y_2 are independent random variables with distributions q_{W_1} and q_{W_2}, respectively. We should check whether with this construction the right hand side of (4.17) is equivalent to the equation in (4.19). For that purpose, we note that $\Delta \in [-1, 1]$ and the function $g(\rho, z)$ is symmetric about $z = 0$. So, when $\Delta_{W_1}(y_1)$ and $\Delta_{W_2}(y_2)$ are of the same

sign, we have

$$\begin{aligned}
&\big(1+Z_1Z_2\big)g\Big(\rho,\frac{Z_1+Z_2}{1+Z_1Z_2}\Big)\\
&\qquad=\big(1+\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)\big)g\Big(\rho,\frac{\Delta_{W_1}(y_1)+\Delta_{W_2}(y_2)}{1+\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)}\Big),\\
&\big(1-Z_1Z_2\big)g\Big(\rho,\frac{Z_1-Z_2}{1-Z_1Z_2}\Big)\\
&\qquad=\big(1-\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)\big)g\Big(\rho,\frac{\Delta_{W_1}(y_1)-\Delta_{W_2}(y_2)}{1-\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)}\Big).
\end{aligned}$$

On the other hand, when $\Delta_W(y_1)$ and $\Delta_W(y_2)$ are of opposite signs, we have

$$\begin{aligned}
&\big(1-Z_1Z_2\big)g\Big(\rho,\frac{Z_1-Z_2}{1-Z_1Z_2}\Big)\\
&\qquad=\big(1+\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)\big)g\Big(\rho,\frac{\Delta_{W_1}(y_1)+\Delta_{W_2}(y_2)}{1+\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)}\Big),\\
&\big(1+Z_1Z_2\big)g\Big(\rho,\frac{Z_1+Z_2}{1+Z_1Z_2}\Big)\\
&\qquad=\big(1-\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)\big)g\Big(\rho,\frac{\Delta_{W_1}(y_1)-\Delta_{W_2}(y_2)}{1-\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)}\Big).
\end{aligned}$$

Since we are interested in the sum of these two parts, we can see that the construction we propose is still equivalent to (4.19). This concludes the proof. □

Remark 4.3. The equations (4.16) and (4.19) derived in the proofs of Lemma 4.1 and Lemma 4.2, respectively, can also be derived by using the results given in [26] for the evolution of $\Delta_W(y)$ under check and variable node operations. The current proofs are kept in the presentation for being self contained.

Remark 4.4. It may be of interest to note that in general $W_{1,2}^{\pm}$ is not the same as $W_{2,1}^{\pm}$, so the order of W_1 and W_2 does matter. However the two channels have the same E_0, (and consequently the same symmetric cutoff rate, same symmetric mutual information, same Bhattacharyya parameter, etc.) as interchanging the random variables Z_1 and Z_2 does not change the values in (4.15) and (4.17). So, we have $E_0(\rho, W_{1,2}^{\pm}) = E_0(\rho, W_{2,1}^{\pm})$.

The following lemma puts in order the E_0 parameters of the channels W_1, W_2, $W_{1,2}^-$, and $W_{1,2}^+$

Lemma 4.5 (Polarization Property)**.** *The channels W_1, W_2, $W_{1,2}^-$, and $W_{1,2}^+$ satisfy for any $\rho \geq 0$ the following ordering:*

$$\begin{aligned}
&E_0(\rho, W_{1,2}^-) \leq E_0(\rho, W_1) \leq E_0(\rho, W_{1,2}^+),\\
&E_0(\rho, W_{1,2}^-) \leq E_0(\rho, W_2) \leq E_0(\rho, W_{1,2}^+).
\end{aligned} \tag{4.20}$$

Proof of Lemma 4.5. We only show the inequalities in (4.20) for the channel W_1. The proof for the channel W_2 follows by Remark 4.4. By (2.12), Lemma 4.1, and Lemma 4.2, the inequalities in (4.20) are equivalent to

$$\mathbb{E}\left[\frac{1+Z_1Z_2}{2}g\Big(\rho,\frac{Z_1+Z_2}{1+Z_1Z_2}\Big)+\frac{1-Z_1Z_2}{2}g\Big(\rho,\frac{Z_1-Z_2}{1-Z_1Z_2}\Big)\right] \leq \mathbb{E}\left[g(\rho,Z_1)\right], \quad (4.21)$$

$$\mathbb{E}\left[g(\rho,Z_1)\right] \leq \mathbb{E}\left[g(\rho,Z_1Z_2)\right]. \quad (4.22)$$

By Lemma 2.2, the function $g(\rho,z)$ is non-increasing in the variable $z\in[0,1]$ when $\rho\geq 0$. Hence, the inequality in (4.22) holds. On the other side, note that for any realizations z_1 and z_2, the factors $(1+z_1z_2)/2$ and $(1-z_1z_2)/2$ form a distribution. As we also know by Lemma 2.2 that the function $g(\rho,z)$ is concave in $z\in[-1,1]$, we can apply Jensen's inequality to obtain

$$\frac{1+z_1z_2}{2}g\Big(\rho,\frac{z_1+z_2}{1+z_1z_2}\Big)+\frac{1-z_1z_2}{2}g\Big(\rho,\frac{z_1-z_2}{1-z_1z_2}\Big) \leq g\Big(\rho,\frac{z_1+z_2}{2}+\frac{z_1-z_2}{2}\Big)=g(\rho,z_1). \quad (4.23)$$

Taking the expectations of both sides, we get the inequality in (4.21). □

The previous lemma shows that our information measure evolves as expected under the polar transform and proves that the polarization property extends to E_0.

4.2 Polar Transform Improves E_0

In this section, we will show the following theorem.

Theorem 4.6. *For any binary input channel W and any $\rho\geq 0$,*

$$E_0(\rho,W^-)+E_0(\rho,W^+)\geq 2E_0(\rho,W).$$

The inequality in Theorem 4.6 holds with equality if and only if the channel W is perfect, or the channel W is completely noisy, or $\rho=0$.

Theorem 4.6 will be obtained as a corollary to a slightly more general result pertaining to the more general polar transform that synthesizes the channels $W_{1,2}^-$ and $W_{1,2}^+$ defined in (4.13) and (4.14), respectively.

Theorem 4.7 (Gain Property)**.** *For any two binary input channels W_1 and W_2 and any $\rho\geq 0$,*

$$E_0(\rho,W_{1,2}^-)+E_0(\rho,W_{1,2}^+)\geq E_0(\rho,W_1)+E_0(\rho,W_2). \quad (4.24)$$

Theorem 4.6 trivially follows from Theorem 4.7 by setting $W_1 = W_2 = W$.

The apparent 'creation' of E_0 by the polar transform does not violate any 'conservation' theorem. While mutual information cannot be improved by processing of the input or output of the channel, E_0 is not a conserved quantity and may be created out of thin air by processing. Indeed, any good coding method implicitly relies on the possibility to create E_0 by processing.

4.2.1 Proof of Theorem 4.7

Proof of Theorem 4.7. By the observations we made in (2.12), Lemma 4.1, and Lemma 4.2, we know that

$$\begin{aligned}
E_0(\rho, W_1) &= -\log \mathbb{E}\big[g(\rho, Z_1)\big], \\
E_0(\rho, W_2) &= -\log \mathbb{E}\big[g(\rho, Z_2)\big], \\
E_0(\rho, W_{1,2}^-) &= -\log \mathbb{E}\big[g(\rho, Z_1 Z_2)\big], \\
E_0(\rho, W_{1,2}^+) &= -\log \mathbb{E}\big[h(\rho, Z_1, Z_2)\big],
\end{aligned}$$

where Z_1, Z_2 are *independent* random variables taking values in the interval $[0, 1]$. By these identities, showing (4.24) is equivalent to showing

$$\mathbb{E}\big[g(\rho, Z_1)\big]\mathbb{E}\big[g(\rho, Z_2)\big] \geq \mathbb{E}\big[g(\rho, Z_1 Z_2)\big]\mathbb{E}\big[h(\rho, Z_1, Z_2)\big].$$

The proof is carried in two steps. We first claim that the following inequality is satisfied:

$$g(\rho, z_1)g(\rho, z_2) \geq g(\rho, z_1 z_2)h(\rho, z_1, z_2), \tag{4.25}$$

for any $z_1, z_2 \in [0, 1]$ and $\rho \geq 0$. From (4.25) and noting the independence of Z_1 and Z_2 we see that

$$\begin{aligned}
\mathbb{E}\big[g(\rho, Z_1)\big]\mathbb{E}\big[g(\rho, Z_2)\big] &= \mathbb{E}\big[g(\rho, Z_1)g(\rho, Z_2)\big] \\
&\geq \mathbb{E}\big[g(\rho, Z_1 Z_2)h(\rho, Z_1, Z_2)\big].
\end{aligned}$$

Lemma 4.14 in Appendix 4.A shows that the function $g(\rho, z_1 z_2)$ is non-increasing in z_1 and z_2 separately for any $\rho \geq 0$. Similarly, Lemma 4.15 in Appendix 4.A shows that the function $h(\rho, z_1, z_2)$ is also non-increasing in z_1 and z_2 separately for any $\rho \geq 0$. These monotonicity properties are useful as they imply, via Lemma 4.16 in Appendix 4.A, that the random variables $g(\rho, Z_1 Z_2)$ and $h(\rho, Z_1, Z_2)$ are positively

correlated. As a result

$$\begin{aligned}\mathbb{E}\big[g(\rho, Z_1)\big]\mathbb{E}\big[g(\rho, Z_2)\big] &\geq \mathbb{E}\big[g(\rho, Z_1Z_2)h(\rho, Z_1, Z_2)\big] \\ &\geq \mathbb{E}\big[g(\rho, Z_1Z_2)\big]\mathbb{E}\big[h(\rho, Z_1, Z_2)\big],\end{aligned}$$

concluding the proof of the relation in (4.24).

Now, we prove the inequality claimed in (4.25). For that purpose, we first apply the following changes of variables

$$\begin{aligned}t &= \operatorname{arctanh} z_1, \quad w = \operatorname{arctanh} z_2, \\ k &= \operatorname{arctanh}(z_1 z_2), \quad s = \frac{1}{1+\rho},\end{aligned}$$

where $s \in (0, 1]$ and $t, w, k \in [0, \infty)$. Using these, we obtain

$$g(\rho, z_1) = g\Big(\frac{1-s}{s}, \tanh(t)\Big) = \frac{\cosh(st)^{1/s}}{\cosh(t)}, \tag{4.26}$$

$$g(\rho, z_2) = g\Big(\frac{1-s}{s}, \tanh(w)\Big) = \frac{\cosh(sw)^{1/s}}{\cosh(w)}, \tag{4.27}$$

$$g(\rho, z_1z_2) = g\Big(\frac{1-s}{s}, \tanh(k)\Big) = \frac{\cosh(sk)^{1/s}}{\cosh(k)}, \tag{4.28}$$

and

$$\begin{aligned}h(\rho, z_1, z_2) &= h\Big(\frac{1-s}{s}, \tanh(t), \tanh(w)\Big) \\ &= \frac{\cosh(s(t+w))^{1/s} + \cosh(s(t-w))^{1/s}}{2\cosh(t)\cosh(w)}.\end{aligned} \tag{4.29}$$

We further define

$$a = t + w, \quad b = t - w,$$

so that $t = (a+b)/2$, $w = (a-b)/2$, and $a \geq |b|$. Then, the variable k is given by

$$k = \frac{1}{2}\ln\left(\frac{\cosh(a)}{\cosh(b)}\right),$$

and the expression in (4.28) becomes

$$g(\rho, z_1z_2) = \frac{\left(\dfrac{\cosh(a)^s + \cosh(b)^s}{2}\right)^{1/s}}{\dfrac{\cosh(a) + \cosh(b)}{2}}. \tag{4.30}$$

With (4.26) and (4.27) at hand, a bit of algebra reveals that the left hand side of

(4.25) is given by

$$\frac{\left(\frac{\cosh(sa)+\cosh(sb)}{2}\right)^{1/s}}{\cosh(t)\cosh(w)}.$$

Similarly, using equations (4.29) and (4.30), the right hand side of (4.25) is given by

$$\frac{\left(\frac{\cosh(a)^s+\cosh(b)^s}{2}\right)^{1/s}}{\frac{\cosh(a)+\cosh(b)}{2}} \times \frac{\cosh(sa)^{1/s}+\cosh(sb)^{1/s}}{2\cosh(t)\cosh(w)}.$$

Therefore, we see that the inequality (4.25) is equivalent to

$$\frac{\left(1+\left(\frac{\cosh(sb)}{\cosh(sa)}\right)\right)^{1/s}}{1+\frac{\cosh(sb)^{1/s}}{\cosh(sa)^{1/s}}} \geq \frac{\left(1+\left(\frac{\cosh(b)}{\cosh(a)}\right)^s\right)^{1/s}}{1+\frac{\cosh(b)}{\cosh(a)}}.$$

Let $u = \left(\frac{\cosh(sb)}{\cosh(sa)}\right)^{1/s}$ and $v = \frac{\cosh(b)}{\cosh(a)}$. Then, by Lemma 4.14 in Appendix 4.A, whenever $a \geq |b|$, we have $1 \geq u \geq v \geq 0$ since

$$f_s(b) = \frac{\cosh(s|b|)^{1/s}}{\cosh(|b|)} \geq \frac{\cosh(sa)^{1/s}}{\cosh(a)} = f_s(a).$$

As a result, we have reduced the inequality (4.25) to the following form:

$$F_s(u) \geq F_s(v) \quad \text{when} \quad 1 \geq u \geq v \geq 0,$$

where

$$F_s(u) = \frac{(1+u^s)^{1/s}}{1+u}.$$

But, we know this is true by Lemma 4.13 in Appendix 4.A. Hence, inequality (4.25) holds as claimed. □

4.3 Extremal Channels of E_0 for the Polar Transform

In this section, we study the extremality of the BEC and the BSC for Gallager's reliability function E_0 of binary input discrete memoryless channels evaluated under the uniform input distribution from the aspect of channel polarization. The next theorem shows that amongst all binary discrete memoryless channels of a given $E_0(\rho)$ value, for a fixed $\rho \geq 0$, the BEC and the BSC are extremal in the evolution of E_0 under the polar transform.

Theorem 4.8 (Extremality Property)**.** *Given two B-DMCs W_1 and W_2, and given any fixed value of $\rho \geq 0$, we define two BSCs BSC_1 and BSC_2, and two BECs BEC_1 and BEC_2 through the equalities*

$$E_0(\rho, W_1) = E_0(\rho, BEC_1) = E_0(\rho, BSC_1), \tag{4.31}$$

$$E_0(\rho, W_2) = E_0(\rho, BEC_2) = E_0(\rho, BSC_2). \tag{4.32}$$

Then, for the $W_{1,2}^-$ polar transformation

$$E_0(\rho, BEC_{1,2}^-) \leq E_0(\rho, W_{1,2}^-) \leq E_0(\rho, BSC_{1,2}^-) \tag{4.33}$$

holds for any $\rho \geq 0$. For the $W_{1,2}^+$ polar transformation

$$E_0(\rho, BSC_{1,2}^+) \leq E_0(\rho, W_{1,2}^+) \leq E_0(\rho, BEC_{1,2}^+) \tag{4.34}$$

holds for any $\rho \in [0,1] \cup [2,\infty]$, and

$$E_0(\rho, BEC_{1,2}^+) \leq E_0(\rho, W_{1,2}^+) \leq E_0(\rho, BSC_{1,2}^+) \tag{4.35}$$

holds for any $\rho \in [1,2]$.

The difficulty in finding minimal and maximal values to the E_0 parameters obtained after applying the polar transform arises from the infinite size of the search space of channels W_1 and W_2 of a given E_0 value. The Lemmas 4.1 and 4.2 will be used to simplify our task.

4.3.1 Proof of Theorem 4.8

The next two lemmas study the first and second order properties for two functions related to the alternative representations of the E_0 parameters of the channels $W_{1,2}^-$ and $W_{1,2}^+$ given in (4.15) and (4.17), respectively.

Lemma 4.9. *For any $z \in [0,1]$ and $\rho \geq 0$, the function $F_{z,\rho}(t)\colon [2^{-\rho}, 1] \to [g(\rho, z), 1]$ defined as*

$$F_{z,\rho}(t) := g(\rho, z g^{-1}(\rho, t)), \tag{4.36}$$

where $g^{-1}(\rho, t)$ denotes the inverse of the function g with respect to its second argument, is convex with respect to t.

Lemma 4.10. *For any $z \in [0,1]$ and $\rho \geq 0$, the function $H_{z,\rho}(t)\colon [2^{-\rho}, 1] \to [2^{-\rho}, g(\rho, z)]$ defined as*

$$H_{z,\rho}(t) := h(\rho, g^{-1}(\rho, t), z) \tag{4.37}$$

is concave with respect to t when $\rho \in [0,1] \cup [2,\infty]$, and convex when $\rho \in [1,2]$.

The proof of Lemma 4.9 and Lemma 4.10 are carried out in Appendix 4.B and Appendix 4.C, respectively. The convexity results stated in Lemma 4.9 and Lemma 4.10 shall constitute key steps in the subsequent proof of Theorem 4.8.

Proof. We start proving the result given in (4.33) for the minus transformation. This proof relies on Lemma 4.1 and the convexity result stated in Lemma 4.9.

From the representation given in (2.12) and Lemma 4.1, we know that

$$\begin{aligned}\exp_2\{-E_0(\rho, W_1)\} &= \mathbb{E}[g(\rho, Z_1)],\\ \exp_2\{-E_0(\rho, W_2)\} &= \mathbb{E}[g(\rho, Z_2)],\\ \exp_2\{-E_0(\rho, W_{1,2}^-)\} &= \mathbb{E}[g(\rho, Z_1 Z_2)],\end{aligned}$$

where Z_1 and Z_2 are independent random variables taking values in $[0, 1]$. We also know that for BSCs $Z_{BSC_1} = z_{BSC_1}$, $Z_{BSC_2} = z_{BSC_2}$. Hence,

$$\begin{aligned}\exp_2\{-E_0(\rho, BSC_1)\} &= g(\rho, z_{BSC_1}),\\ \exp_2\{-E_0(\rho, BSC_2)\} &= g(\rho, z_{BSC_2}),\\ \exp_2\{-E_0(\rho, BSC_{1,2}^-)\} &= g(\rho, z_{BSC_1} z_{BSC_2}).\end{aligned}$$

By $E_0(\rho, W_1) = E_0(\rho, BSC_1)$, $E_0(\rho, W_2) = E_0(\rho, BSC_2)$, we have

$$\begin{aligned}\mathbb{E}\left[g(\rho, Z_1)\right] &= g(\rho, z_{BSC_1}),\\ \mathbb{E}\left[g(\rho, Z_2)\right] &= g(\rho, z_{BSC_2}).\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\exp_2\{-E_0(\rho, W_{1,2}^-)\} &= \mathbb{E}_{Z_1}[\mathbb{E}_{Z_2}[F_{z_1,\rho}\left(g(\rho, Z_2)\right) \mid Z_1 = z_1]]\\ &\geq \mathbb{E}_{Z_1}[F_{Z_1,\rho}\left(\mathbb{E}_{Z_2}[g(\rho, Z_2)]\right)]\\ &= \mathbb{E}_{Z_1}[F_{Z_1,\rho}\left(g(\rho, z_{BSC_2})\right)]\\ &\overset{(1)}{=} \mathbb{E}_{Z_1}\left[F_{z_{BSC_2},\rho}\left(g(\rho, Z_1)\right)\right]\\ &\geq F_{z_{BSC_2},\rho}\left(\mathbb{E}_{Z_1}[g(\rho, Z_1)]\right)\\ &= F_{z_{BSC_2},\rho}\left(g(\rho, z_{BSC_1})\right)\\ &= \exp_2\{-E_0(\rho, BSC_{1,2}^-)\},\end{aligned}$$

where we used Jensen's inequality twice, and where $\overset{(1)}{=}$ follows by the fact that $F_{Z_1,\rho}\left(g(\rho, z_{BSC_2})\right) = F_{z_{BSC_2},\rho}\left(g(\rho, Z_1)\right)$ holds. This proves the upper bound in (4.33).

Let ϵ_1 and ϵ_2 be the erasure probabilities of the channels BEC_1 and BEC_2, respectively. Then, we know that both Z_{BEC_1} and Z_{BEC_2} are $\{0, 1\}$ valued, $P(Z_{BEC_1} =$

$0) = \epsilon_1$, and $P(Z_{BEC_2} = 0) = \epsilon_2$. Therefore,

$$\begin{aligned} \exp_2\{-E_0(\rho, BEC_1)\} &= \epsilon_1(1-2^{-\rho}) + 2^{-\rho}, \\ \exp_2\{-E_0(\rho, BEC_2)\} &= \epsilon_2(1-2^{-\rho}) + 2^{-\rho}. \end{aligned} \tag{4.38}$$

Moreover, the channel $BEC_{1,2}^-$ is a BEC with erasure probability $\epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$. Hence we get

$$\exp_2\{-E_0(\rho, BEC_{1,2}^-)\} = [\epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2]\,(1-2^{-\rho}) + 2^{-\rho}.$$

By $E_0(\rho, W_1) = E_0(\rho, BEC_1)$, $E_0(\rho, W_2) = E_0(\rho, BEC_2)$, we have

$$\begin{aligned} \mathbb{E}\left[g(\rho, Z_1)\right] &= \mathbb{E}\left[g(\rho, Z_{BEC_1})\right] = \epsilon_1(1-2^{-\rho}) + 2^{-\rho}, \\ \mathbb{E}\left[g(\rho, Z_2)\right] &= \mathbb{E}\left[g(\rho, Z_{BEC_2})\right] = \epsilon_2(1-2^{-\rho}) + 2^{-\rho}. \end{aligned}$$

Due to convexity, we also know the following inequality holds:

$$F_{z,\rho}(t) \le 1 + \frac{g(\rho, z) - 1}{2^{-\rho} - 1}(t-1),$$

for $2^{-\rho} \le t \le 1$. As a result,

$$\begin{aligned} &\exp_2\{-E_0(\rho, W_{1,2}^-)\} = \mathbb{E}_{Z_1}[\mathbb{E}_{Z_2}[F_{z_1,\rho}\left(g(\rho, Z_2)\right) \mid Z_1 = z_1]] \\ &\le \mathbb{E}_{Z_1}\left[1 + \frac{g(\rho, Z_1) - 1}{2^{-\rho} - 1}(\mathbb{E}_{Z_2}[g(\rho, Z_2)] - 1)\right] \\ &= 1 + \frac{\mathbb{E}_{Z_1}[g(\rho, Z_1)] - 1}{2^{-\rho} - 1}(\mathbb{E}_{Z_2}[g(\rho, Z_2)] - 1) \\ &= 1 + \frac{[\epsilon_1(1-2^{-\rho}) + 2^{-\rho} - 1]\,[\epsilon_2(1-2^{-\rho}) + 2^{-\rho} - 1]}{2^{-\rho} - 1} \\ &= 1 - \epsilon_1\epsilon_2(1-2^{-\rho}) + (\epsilon_1 + \epsilon_2)\,(1-2^{-\rho}) + 2^{-\rho} - 1 \\ &= [\epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2]\,(1-2^{-\rho}) + 2^{-\rho} \\ &= \exp_2\{-E_0(\rho, BEC_{1,2}^-)\}. \end{aligned}$$

This proves the lower bound in (4.33) and concludes the proof for the minus transformation.

The proof of the theorem for the plus transformation can be completed following steps similar to the minus case. The proof relies on Lemma 4.2 and the convexity result stated in Lemma 4.10. Here, we briefly sketch the proof. By Lemma 4.2, we have

$$\mathbb{E}\left[h(\rho, Z_1, Z_2)\right] = \exp_2\{-E_0(\rho, W_{1,2}^+)\}.$$

Observe that the function $h(\rho, z_1, z_2)$ defined in (4.18) is symmetric in the variables

z_1 and z_2, i.e., $h(\rho, z_1, z_2) = h(\rho, z_2, z_1)$. We define the random variables

$$A_1 = g(\rho, Z_1) \quad \text{and} \quad A_2 = g(\rho, Z_2).$$

Using the concavity of the function $H_{z,\rho}(t)$ with respect to t for fixed values of $\rho \in [0,1] \cup [2,\infty]$ and $z \in [0,1]$, we obtain the inequalities in (4.34)

$$\begin{aligned}\exp_2\{-E_0(\rho, W_{1,2}^+)\} &= \mathbb{E}\left[H_{g^{-1}(\rho,A_2),\rho}(A_1)\right] \\ &\leq h(\rho, z_{BSC_1}, z_{BSC_2}) = \exp_2\{-E_0(\rho, BSC_{1,2}^+)\},\end{aligned}$$

$$\begin{aligned}\exp_2\{-E_0(\rho, W_{1.2}^+)\} &= \mathbb{E}\left[H_{g^{-1}(\rho,A_2),\rho}(A_1)\right] \\ &\geq \epsilon_1\epsilon_2\left(1-2^{-\rho}\right) + 2^{-\rho} = \exp_2\{-E_0(\rho, BEC_{1,2}^+)\},\end{aligned}$$

as the channel $BEC_{1,2}^+$ is a BEC with erasure probability $\epsilon_1\epsilon_2$. Similarly, the convexity of the function $H_{z,\rho}(t)$ with respect to t for $\rho \in [1,2]$ leads to the reverse inequalities in (4.35). □

4.4 Discussion

In the following four subsections, we state a number of corollaries to the results of the previous sections. The first is a restatement of Theorem 4.6 and a consequence of Theorem 4.8 in the language of martingales. The second is an implication of Theorem 4.6 on complexity, and the third a special case of Rényi chain rules we get via 4.7. The final subsection discusses some special values of the parameter ρ.

4.4.1 Gain & Convergence Law for E_0

In Theorem 4.6, we showed that the polar transform improves E_0. This implies that the process $E_0(\rho, W_n)$ associated to the channel polarization process W_n, for $n \geq 0$, is a submartingale. After a normalization by the value of $\rho \geq 0$, one can easily see that the process $E_0(\rho, W_n)/\rho$ is as well a submartingale and takes values from the bounded interval $[0,1]$. By general results on bounded martingales, the process converges a.s., see for instance [27]. Now, applying the extremality results of Theorem 4.8 reveals the convergence points of this process are the extremes of the bounded interval.

Proposition 4.11 (Gain & Convergence Law)**.** *The process $E_0(\rho, W_n)/\rho$, for any $\rho \geq 0$ and for $n \geq 0$, is a bounded submartingale which converges a.s. to $\{0,1\}$.*

The formal proof for identifying the convergence points of the process can be carried by bounding, for any $\rho \geq 0$, the difference between the E_0 parameters of

the channel W^+ and the channel W, and the one between the channel W and the channel W^- using the inequalities given in Theorem 4.8.

4.4.2 Improving the Reliability–Complexity Trade-off

Beside its usefulness as an argument in channel polarization related proofs, an interesting interpretation of the inequality in Theorem 4.6 is given in [4]. Arıkan discusses the concept of the reliability–complexity exponent under maximum likelihood decoding of a code drawn from a random code ensemble. This concept was introduced in [28] and in [29, Section 6.6]. He suggests the general method of channel combining and splitting can be used to improve this trade-off. Here we explore this idea.

A maximum likelihood decoder for a randomly constructed code needs to compute the likelihoods for all codewords, and this incurs a complexity of $\chi \simeq 2^{NR}$. At the same time such a code has error probability $P_{\text{e, avg}} \simeq 2^{-NE_r(R,W)}$, where $E_r(R,W)$ is the random coding exponent [5]. Consequently, the complexity χ and the error probability $P_{\text{e, avg}}$ are algebraically related, $P_{\text{e, avg}} \simeq \chi^{-E_r(R,W)/R}$, and the quantity $E_r(R,W)/R$ is defined as the '*reliability–complexity exponent*'.

For a given rate R and B-DMC W, consider the particular ρ value, say ρ^*, which maximizes the random coding exponent [5, p. 139]

$$E_r(R,W) = \max_{\rho \in [0,1]} \left[E_0(\rho, W) - \rho R \right].$$

For that particular ρ^*, we have

$$\begin{aligned} 2E_r(R,W) &= 2E_0(\rho^*, W) - 2\rho^* R \\ &\leq E_0(\rho^*, W^+) + E_0(\rho^*, W^-) - \rho^* 2R. \end{aligned}$$

Now, if the rate $2R$ is split into two parts R^+ and R^- proportional to $E_0(\rho^*, W^+)$ and $E_0(\rho^*, W^-)$, respectively, i.e., they satisfy $2R = R^+ + R^-$ and

$$\frac{R^+}{E_0(\rho^*, W^+)} = \frac{R^-}{E_0(\rho^*, W^-)},$$

then the reliability–complexity trade-off $E_r(R,W)/R$ of random codes will satisfy

$$\begin{aligned} \frac{E_r(R,W)}{R} = \frac{E_0(\rho^*, W)}{R} - \rho^* &\leq \frac{E_0(\rho^*, W^-) + E_0(\rho^*, W^+)}{2R} - \rho^* \\ &= \frac{E_0(\rho^*, W^-) + E_0(\rho^*, W^+)}{R^- + R^+} - \rho^* \\ &= \frac{E_0(\rho^*, W^-)}{R^-} - \rho^* \leq \frac{E_r(R^-, W^-)}{R^-}, \end{aligned}$$

and similarly,

$$\frac{E_r(R,W)}{R} \leq \frac{E_r(R^+,W^+)}{R^+}.$$

Therefore, both of the synthesized channels will have a better reliability–complexity exponent function than the original channel. In that respect, the inequality in Theorem 4.6 implies the particular polar transform combined with a successive cancellation decoder does improve the reliability–complexity exponent of random codes.

4.4.3 Chain Rule for Rényi's Entropies

If P is the uniform input distribution on the set $\mathbb{F}_2$, then P^{α}_{tilt} is also the uniform distribution on $\mathbb{F}_2$. Therefore, $E_0(\rho,W)/\rho$ can be defined in terms of Rényi's entropy functions as:

$$\frac{E_0(\rho,W)}{\rho} = H_{\frac{1}{1+\rho}}(X) - H_{\frac{1}{1+\rho}}(X \mid Y).$$

Moreover, we get

$$H_{\frac{1}{1+\rho}}(X) = \frac{1}{\rho}\log\left(\sum_{x\in\mathbb{F}_2} P(x)^{\frac{1}{1+\rho}}\right)^{\frac{1}{1+\rho}} = \log|\mathbb{F}_2| = \log 2,$$

Using the definitions in (4.13) and (4.14), $E_0(\rho,W^-_{1,2})$ and $E_0(\rho,W^+_{1,2})$ can be expressed by

$$\frac{E_0(\rho,W^-_{1,2})}{\rho} = \log 2 - H_{\frac{1}{1+\rho}}(U_1 \mid Y_1Y_2),$$
$$\frac{E_0(\rho,W^+_{1,2})}{\rho} = \log 2 - H_{\frac{1}{1+\rho}}(U_2 \mid Y_1Y_2U_1),$$

where Y_1 is the output of the channel W_1 with input $X_1 = U_1 \oplus U_2$, and Y_2 is the output of the channel W_2 with input $X_2 = U_2$. In addition,

$$\frac{E_0(\rho,W_1)}{\rho} + \frac{E_0(\rho,W_2)}{\rho} = 2\log 2 - H_{\frac{1}{1+\rho}}(X_1X_2 \mid Y_1Y_2).$$

Since the mapping between (X_1,X_2) and (U_1,U_2) is one-to-one, we have

$$H_{\frac{1}{1+\rho}}(X_1X_2 \mid Y_1Y_2) = H_{\frac{1}{1+\rho}}(U_1U_2 \mid Y_1Y_2).$$

From these expressions, we deduce that the relationship between the E_0 parameters of the channels W_1, W_2, $W^-_{1,2}$, and $W^+_{1,2}$ derived in Theorem 4.7 implies a certain

'*chain rule inequality*' holds for the polar transform, i.e.,

$$H_{\frac{1}{1+\rho}}(U_1U_2 \mid Y_1Y_2) \geq H_{\frac{1}{1+\rho}}(U_1 \mid Y_1Y_2) + H_{\frac{1}{1+\rho}}(U_2 \mid Y_1Y_2U_1),$$

for $\rho \geq 0$ whenever U_1, U_2 are i.i.d. uniform on $\mathbb{F}_2$. Equivalently, we can conclude that whenever (i) (X_1, Y_1) and (X_2, Y_2) are independent, and (ii) X_1 and X_2 are uniformly distributed on $\mathbb{F}_2$, then, with $U_1 = X_1 \oplus X_2$ and $U_2 = X_2$,

$$H_\alpha(U_1U_2|Y_1Y_2) \geq H_\alpha(U_1|Y_1Y_2) + H_\alpha(U_2|Y_1Y_2U_1),$$

for any $\alpha \leq 1$.

4.4.4 Special Cases

In Theorem 4.8, we showed that amongst all B-DMCs W_1 and W_2 of fixed $E_0(\rho, W_1)$ and $E_0(\rho, W_2)$ values for a given value of $\rho \geq 0$, the channel $BEC^-_{1,2}$ results in a lower bound to the value of $E_0(\rho, W^-_{1,2})$ and the channel $BSC^-_{1,2}$ in an upper bound to the value of $E_0(\rho, W^-_{1,2})$. A similar extremal property holds for the plus transformation except for the difference that the result breaks into two parts depending on the value of the parameter ρ: While the channel $BEC^+_{1,2}$ upper bounds and the channel $BSC^+_{1,2}$ lower bounds the value of $E_0(\rho, W^+_{1,2})$ when $\rho \in [0, 1] \cup [2, \infty]$, these roles are reversed when $\rho \in [1, 2]$. Using these results, we identify in this section some special cases of the ρ values in order to recover known and discover new results.

Symmetric Capacity

We know that the symmetric capacity process $I_n(W)$, for $n \geq 0$, is a bounded martingale which converges a.s. to $\{0, 1\}$. Here, we show that Theorem 4.8 can be used to identify these convergence points.

Corollary 4.12. *Under the assumptions of Theorem 4.8 for $W_1 = W_2 = W$, we have for $\rho \in [0, 1]$ the following inequalities:*

$$\begin{aligned} E_0(\rho, BSC^+) - E_0(\rho, BSC^-) \leq E_0(\rho, W^+) - E_0(\rho, W^-) \\ \leq E_0(\rho, BEC^+) - E_0(\rho, BEC^-). \end{aligned}$$

Corollary 4.12 shows that among channels W with a given value of $E_0(\rho, W)$ for a given $\rho \in [0, 1]$, the BEC and the BSC are the most and the least polarizing under Arıkan's polar transformations in the sense that their polar transforms W^+ and W^- have the largest and the smallest differences in their E_0 values. Dividing all sides of the inequality above by ρ and taking the limit as $\rho \to 0$, we recover via (2.4) the extremality property stated in (4.7)

Note that the preservation property of the symmetric capacities holds regardless of whether the combined channels are identical or not, as it is a consequence of the chain rule for mutual information. Namely, the channels satisfy

$$I(W_1) + I(W_2) = I(W_{1,2}^-) + I(W_{1,2}^+).$$

We will also make this observation in Chapter 9 where we will extend the idea of channel polarization over non-stationary B-DMCs.

Cutoff Rate, Bhatthacharyya Parameter

In this case, Theorem 4.8 implies the channels having their cutoff rates equal to $E_0(1, W_1)$ and $E_0(1, W_2)$ satisfy

$$\begin{aligned} E_0(1, BEC_{1,2}^-) \leq E_0(1, W_{1,2}^-) \leq E_0(1, BSC_{1,2}^-), \\ E_0(1, BSC_{1,2}^+) = E_0(1, W_{1,2}^+) = E_0(1, BEC_{1,2}^+). \end{aligned}$$

Hence, by (4.11), the extremalities for the Bhattacharyya parameter are also obtained. Indeed, one can show that $Z(W_{1,2}^+) = Z(W_1)Z(W_2)$ holds. So, another result of [2, Proposition 5] is recovered by letting $W_1 = W_2 = W$, i.e., $Z(W^+) = Z(W)^2$.

Parameter at $\rho = 2$

A previously unknown result is found by taking $\rho = 2$ in the theorem. As in the case $\rho = 1$, we observe that the E_0 parameters of the channels $W_{1,2}^+$, $BEC_{1,2}^+$, and $BSC_{1,2}^+$ are equal to each other. Moreover, if we define

$$Z(\rho, W) := \frac{2^\rho \exp_2\{-E_0(\rho, W)\} - 1}{2^\rho - 1}, \tag{4.39}$$

for $\rho \geq 0$, then $Z(2, W_{1,2}^+) = Z(2, W_1)Z(2, W_2)$ holds by Theorem 4.8. To see this, simply note that for a BEC BEC with erasure probability ϵ, we have by (4.38) the relation $Z(\rho, BEC) = \epsilon$, for any $\rho \geq 0$. So, by letting $W_1 = W_2 = W$, we get $Z(2, W^+) = Z(2, W)^2$.

As a last comment, we observe that the $\rho = 1, 2$ values share a common property: One can recover the value of the parameter $E_0(\rho, W_{1,2}^+)$ from the values of $E_0(\rho, W_1)$ and $E_0(\rho, W_2)$ when $\rho = 1, 2$, without necessarily knowing the particular channels W_1 and W_2, by using the relation in (4.39) and the fact that $Z(\rho, W_{1,2}^+) = Z(\rho, W_1)Z(\rho, W_2)$ holds for $\rho = 1, 2$.

Appendix

In the first part of this appendix, 4 lemmas will be stated and proved: Lemmas 4.13, 4.14, 4.15, and 4.16. In the second part Lemma 4.9 will be proved, and in the final part Lemma 4.10 will be proved.

4.A Lemmas 4.13, 4.14, 4.15, and 4.16

Lemma 4.13. *For $s \in (0,1]$, define the function $F_s \colon [0,1] \to [1, 2^{\frac{1-s}{s}}]$ as*

$$F_s(x) := \frac{(1+x^s)^{\frac{1}{s}}}{1+x}.$$

Then, F_s is a non-decreasing function.

Proof. Taking the derivative of $F_s(x)$ with respect to x, we have

$$\frac{\partial}{\partial x} F_s(x) = \frac{(1+x^s)^{\frac{1}{s}-1}(x^s - x)}{x(1+x)^2} \geq 0,$$

since $(x^s - x) \geq 0$ for $x \in [0,1]$ and $s \in (0,1]$. □

Lemma 4.14. *For $s \in (0,1]$, define the function $f_s \colon [0,\infty) \to [2^{\frac{s-1}{s}}, 1]$ as*

$$f_s(k) := \frac{\cosh(ks)^{\frac{1}{s}}}{\cosh(k)}.$$

Then, f_s is a non-increasing function. Moreover, this implies the function $g(\rho, z)$ defined in (2.13) *is non-increasing in the variable $z \in [0,1]$ for any fixed $\rho \geq 0$.*

Proof. We can equivalently show that $\log(f_s(k))$ is non-increasing in k. Taking the first derivative gives

$$\frac{\partial}{\partial k}\left(\frac{1}{s}\ln(\cosh(ks)) - \ln(\cosh(k))\right) = \tanh(ks) - \tanh(k) \leq 0$$

as $\tanh(\cdot)$ is increasing in its argument. To prove the second monotonicity relation, we let $k = \operatorname{arctanh} z$ and $s = \frac{1}{1+\rho}$. Then,

$$g(\rho, z) = f_{\frac{1}{1+\rho}}(\operatorname{arctanh} z).$$

Since $\operatorname{arctanh}$ is a monotone increasing function, it follows that the function $g(\rho, z)$ is non-increasing in z. □

Lemma 4.15. *The function* $h(\rho, z_1, z_2)\colon [0,\infty) \times [0,1] \times [0,1] \to [2^{-\rho}, 1]$ *defined in* (4.18)*, is non-increasing in the variables* z_1 *and* z_2 *separately for any* $\rho \geq 0$*.*

Proof. By the symmetry of h with respect to z_1, z_2, i.e., $h(\rho, z_1, z_2) = h(\rho, z_2, z_1)$, it suffices to show the claim for z_1 alone. In the expression below, we will suppress ρ in all function arguments, and denote $g'(u) = \dfrac{\partial}{\partial u} g(\rho, u)$. Taking the derivative of h with respect to z_1, we get

$$\begin{aligned}
\frac{\partial}{\partial z_1} h(z_1, z_2) &= \frac{1}{2} z_2\, g\Big(\frac{z_1 + z_2}{1 + z_1 z_2}\Big) + \frac{1 - {z_2}^2}{2(1 + z_1 z_2)}\, g'\Big(\frac{z_1 + z_2}{1 + z_1 z_2}\Big) \\
&\quad - \frac{1}{2} z_2\, g\Big(\frac{z_1 - z_2}{1 - z_1 z_2}\Big) + \frac{1 - {z_2}^2}{2(1 - z_1 z_2)}\, g'\Big(\frac{z_1 - z_2}{1 - z_1 z_2}\Big) \\
&= \frac{1}{2} z_2 \Big[g\left(\frac{z_1 + z_2}{1 + z_1 z_2}\right) - g\left(\frac{z_1 - z_2}{1 - z_1 z_2}\right) \Big] \\
&\quad + \frac{1 - {z_2}^2}{2(1 + z_1 z_2)}\, g'\Big(\frac{z_1 + z_2}{1 + z_1 z_2}\Big) + \frac{1 - {z_2}^2}{2(1 - z_1 z_2)}\, g'\Big(\frac{z_1 - z_2}{1 - z_1 z_2}\Big).
\end{aligned}$$

The last two terms with $g'(\cdot)$ are negative by Lemma 4.14, so it suffices to show that

$$g\left(\frac{z_1 + z_2}{1 + z_1 z_2}\right) \leq g\left(\frac{z_1 - z_2}{1 - z_1 z_2}\right).$$

To that end, observe that, for any $z_1, z_2 \in [0,1]$ we have

$$\frac{z_1 + z_2}{1 + z_1 z_2} \geq \frac{|z_1 - z_2|}{1 - z_1 z_2},$$

and by Lemma 4.14 and the symmetry of g around $z = 0$, the inequality required follows. □

Lemma 4.16. *Suppose* $f\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ *and* $g\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ *are two functions that satisfy*

$$\big[f(x,y) - f(x',y)\big]\big[g(x,y') - g(x',y')\big] \geq 0,$$

and

$$\big[f(x,y) - f(x,y')\big]\big[g(x,y) - g(x,y')\big] \geq 0,$$

for every x, x', y, y'*. Then, for any independent random variables* X, Y *the random variables* $f(X,Y)$ *and* $g(X,Y)$ *are positively correlated.*

Note that if $\mathcal{X}$ and $\mathcal{Y}$ are ordered sets and f and g are monotone (in the same sense) in their arguments then they satisfy the requirements of the lemma.

Proof. Let (X', Y') be an independent copy of (X, Y). By the first premise of the

lemma

$$\big[f(X,Y)-f(X',Y)\big]\big[g(X,Y')-g(X',Y')\big]\geq 0.$$

Taking expectations, we get

$$\mathbb{E}[f(X,Y)g(X,Y')]+\mathbb{E}[f(X',Y)g(X',Y')] \\ \geq \mathbb{E}[f(X,Y)g(X',Y')]+\mathbb{E}[f(X',Y)g(X,Y')],$$

equivalently (as $(X,Y,Y')\sim(X',Y,Y')$ and $(X,Y,X',Y')\sim(X',Y,X,Y')$),

$$\mathbb{E}[f(X,Y)g(X,Y')]\geq\mathbb{E}[f(X,Y)]\mathbb{E}[g(X,Y)]. \tag{4.40}$$

By the second premise of the lemma

$$\big[f(X,Y)-f(X,Y')\big]\big[g(X,Y)-g(X,Y')\big]\geq 0.$$

Taking expectations, we get

$$\mathbb{E}[f(X,Y)g(X,Y)]+\mathbb{E}[f(X,Y')g(X,Y')] \\ \geq \mathbb{E}[f(X,Y)g(X,Y')]+\mathbb{E}[f(X,Y')g(X,Y)],$$

which is equivalent to (as $(X,Y)\sim(X,Y')$ and $(X,Y,Y')\sim(X,Y',Y)$)

$$\mathbb{E}[f(X,Y)g(X,Y)]\geq\mathbb{E}[f(X,Y)g(X,Y')]. \tag{4.41}$$

Putting together (4.40) and (4.41) concludes the proof. □

4.B Proof of Lemma 4.9

Proof. We prove that the function $F_{z,\rho}(t)=g(\rho,zg^{-1}(\rho,t))$ defined in (4.36) is convex with respect to the variable t for fixed values of $\rho\geq 0$ and $z\in[0,1]$. Taking the first derivative with respect to t, we obtain

$$\frac{\partial F_{z,\rho}(t)}{\partial t}=\frac{\partial}{\partial t}g(\rho,zg^{-1}(\rho,t))=\frac{g'(\rho,zg^{-1}(\rho_1,t))}{g'(\rho_1,g^{-1}(\rho_1,t))}z.$$

We define $u=g^{-1}(\rho,t)$. As, by Lemma 2.2, $g(\rho,u)$ is a non-increasing function in $u\in[0,1]$ for $\rho\geq 0$, so is $g^{-1}(\rho,t)$ in t. So, we can check the convexity of $F_{z,\rho}(t)$ with respect to t, from the monotonicity with respect to u of the following expression:

$$z\frac{g'(\rho,zu)}{g'(\rho,u)}. \tag{4.42}$$

To simplify notation, we use

$$\begin{aligned}\lambda(u) &:= \frac{1-u}{1+u},\\ \alpha(\rho,u) &:= (1+\lambda(u)^{\frac{1}{1+\rho}})^{\rho} \geq 0,\\ \beta(\rho,u) &:= (1-\lambda(u)^{\frac{-\rho}{1+\rho}}) \leq 0.\end{aligned}$$

Then, by equation (2.26), we have

$$\begin{aligned}\frac{\partial g(\rho,u)}{\partial u} &= \left(\frac{1}{2}\right)^{1+\rho}\alpha(\rho,u)\beta(\rho,u),\\ \frac{\partial g(\rho,zu)}{\partial u} &= \left(\frac{1}{2}\right)^{1+\rho} z\alpha(\rho,zu)\beta(\rho,zu),\end{aligned}$$

and (4.42) is given by

$$z\frac{g'(\rho_2,zu)}{g'(\rho_1,u)} = z\frac{\alpha(\rho,zu)\beta(\rho,zu)}{\alpha(\rho,u)\beta(\rho,u)}. \tag{4.43}$$

Now taking the derivative of (4.43) with respect to u, we get

$$\begin{aligned}\frac{\partial}{\partial u}\left(z\frac{\alpha(\rho,zu)\beta(\rho,zu)}{\alpha(\rho,u)\beta(\rho,u)}\right) &= z\underbrace{\frac{\alpha(\rho,zu)\beta(\rho,zu)}{\alpha(\rho,u)\beta(\rho,u)}}_{\geq 0}\times\\ &\left(\frac{\partial\alpha(\rho,zu)/\partial u}{\alpha(\rho,zu)} + \frac{\partial\beta(\rho,zu)/\partial u}{\beta(\rho,zu)} - \frac{\partial\alpha(\rho,u)/\partial u}{\alpha(\rho,u)} - \frac{\partial\beta(\rho,u)/\partial u}{\beta(\rho,u)}\right).\end{aligned} \tag{4.44}$$

We can see that the sign of the expression inside the parenthesis in (4.44) will determine the monotonicity in u of the expression in (4.43). At this point, we note that

$$\frac{\partial\alpha(\rho,u)/\partial u}{\alpha(\rho,u)} + \frac{\partial\beta(\rho,u)/\partial u}{\beta(\rho,u)} = \left(\frac{\partial\alpha(\rho,zu)/\partial u}{\alpha(\rho,zu)} + \frac{\partial\beta(\rho,zu)/\partial u}{\beta(\rho,zu)}\right)\Bigg|_{z=1}. \tag{4.45}$$

Moreover, we claim that the expression inside the parenthesis in the right hand side of (4.45) is non-decreasing in $z \in [0,1]$. As a consequence, the expression in (4.43) is non-increasing in $u \in [0,1]$, which implies $F_{z,\rho}(t)$ is a concave function in $u = g^{-1}(\rho,t)$. Since u is non-increasing in t, we have

$$\frac{\partial^2 F_{z,\rho}(t)}{\partial t^2} = \underbrace{\frac{\partial}{\partial u}\left(z\frac{g'(\rho_2,zu)}{g'(\rho_1,u)}\right)}_{\leq 0}\underbrace{\frac{\partial u}{\partial t}}_{\leq 0} \geq 0.$$

This proves that $F_{z,\rho}(t)$ is a convex function in t.

In the rest of the appendix, we prove our claim that the expression inside the parenthesis in the right hand side of (4.45) is non-decreasing in z. We have,

$$\frac{\partial \alpha(\rho, zu)}{\partial u} = \frac{\rho}{1+\rho} z\lambda'(zu)\lambda(zu)^{\frac{-\rho}{1+\rho}}(1+\lambda(zu)^{\frac{1}{1+\rho}})^{\rho-1}, \tag{4.46}$$

$$\frac{\partial \beta(\rho, zu)}{\partial u} = \frac{\rho}{1+\rho} z\lambda'(zu)\lambda(zu)^{\frac{-\rho}{1+\rho}-1}, \tag{4.47}$$

where $\lambda'(u) = \dfrac{\partial \lambda(u)}{\partial u} = \dfrac{-2}{(1+u)^2}$. Hence,

$$\begin{aligned}
&\frac{\partial \alpha(\rho, zu)/\partial u}{\alpha(\rho, zu)} + \frac{\partial \beta(\rho, zu)/\partial u}{\beta(\rho, zu)} \\
=&\frac{\rho}{1+\rho}\lambda(zu)^{\frac{-\rho}{1+\rho}-1} z\lambda'(zu)\left(\frac{\lambda(zu)}{1+\lambda(zu)^{\frac{1}{1+\rho}}} + \frac{1}{1-\lambda(zu)^{\frac{-\rho}{1+\rho}}}\right) \\
=&\frac{\rho}{1+\rho}\lambda(zu)^{\frac{-\rho}{1+\rho}-1} z\lambda'(zu)\left(\frac{\lambda(zu) - \lambda(zu)^{\frac{1}{1+\rho}} + 1 + \lambda(zu)^{\frac{1}{1+\rho}}}{(1+\lambda(zu)^{\frac{1}{1+\rho}})(1-\lambda(zu)^{\frac{-\rho}{1+\rho}})}\right) \\
=&\frac{\rho}{1+\rho}\lambda(zu)^{\frac{-\rho}{1+\rho}-1} z\lambda'(zu)(1+\lambda(zu))(1+\lambda(zu)^{\frac{1}{1+\rho}})^{-1}(1-\lambda(zu)^{\frac{-\rho}{1+\rho}})^{-1} \\
=&\frac{\rho}{1+\rho} z\lambda'(zu)(1+\lambda(zu)^{-1})(1+\lambda(zu)^{\frac{1}{1+\rho}})^{-1}(\lambda(zu)^{\frac{\rho}{1+\rho}} - 1)^{-1} \\
=&\frac{\rho}{1+\rho}\frac{-4z}{(1+zu)^2(1-zu)}\left(1 + \left(\frac{1-zu}{1+zu}\right)^{\frac{1}{1+\rho}}\right)^{-1} \times \\
&\qquad\qquad \left(-1 + \left(\frac{1-zu}{1+zu}\right)^{\frac{\rho}{1+\rho}}\right)^{-1}.
\end{aligned}$$

By re-arranging, we get

$$\begin{aligned}
&\frac{\partial \alpha(\rho, zu)/\partial u}{\alpha(\rho, zu)} + \frac{\partial \beta(\rho, zu)/\partial u}{\beta(\rho, zu)} \\
&= \frac{4\rho}{1+\rho}\left(\underbrace{\frac{1-z^2u^2}{z}\left((1+zu)^{\frac{\rho}{1+\rho}} - (1-zu)^{\frac{\rho}{1+\rho}}\right)}_{\text{Part 2}} \times \right. \\
&\qquad\qquad \left. \underbrace{\left((1+zu)^{\frac{1}{1+\rho}} + (1-zu)^{\frac{1}{1+\rho}}\right)}_{\text{Part 1}}\right)^{-1}.
\end{aligned}$$

We will consider the expressions labeled as Part 1 and Part 2 separately. Note that both are positive valued. Moreover, we will show that both are non-increasing in

$z \in [0,1]$. Then, this will prove our claim as

$$\frac{\partial}{\partial z}\Bigg(\frac{(1-z^2u^2)}{z}\left((1+zu)^{\frac{1}{1+\rho}}+(1-zu)^{\frac{1}{1+\rho}}\right)\times \\ \left((1+uz)^{\frac{\rho}{1+\rho}}-(1-uz)^{\frac{\rho}{1+\rho}}\right)\Bigg)\leq 0$$

implies

$$\frac{\partial}{\partial z}\left(\frac{\partial\alpha(\rho,zu)/\partial u}{\alpha(\rho,zu)}+\frac{\partial\beta(\rho,zu)/\partial u}{\beta(\rho,zu)}\right)\geq 0,$$

for $\rho \geq 0$.

For Part 2, we have

$$\begin{aligned}
&\frac{\partial}{\partial z}\left(\frac{(1-u^2z^2)}{z}\left((1+uz)^{\frac{\rho}{1+\rho}}-(1-uz)^{\frac{\rho}{1+\rho}}\right)\right)\\
=&\frac{1}{z^2}\frac{\rho uz\left(1-u^2z^2\right)\left((1+uz)^{\frac{\rho}{1+\rho}-1}+(1-uz)^{\frac{\rho}{1+\rho}-1}\right)}{1+\rho}\\
&+\frac{1}{z^2}\left(1+u^2z^2\right)\left(-(1+uz)^{\frac{\rho}{1+\rho}}+(1-uz)^{\frac{\rho}{1+\rho}}\right)\\
=&\frac{1}{z^2}\Bigg((1+uz)^{\frac{\rho}{1+\rho}}\left(\frac{\rho}{1+\rho}uz\left(1-uz\right)-\left(1+u^2z^2\right)\right)\\
&+(1-uz)^{\frac{\rho}{1+\rho}}\left(\frac{\rho}{1+\rho}uz\left(1+uz\right)+\left(1+u^2z^2\right)\right)\Bigg).
\end{aligned}$$

Letting $r=\frac{\rho}{1+\rho}\in[0,1)$ and $x=uz\in[0,1]$, Part 2 equals

$$\begin{aligned}
=&\frac{1}{z^2}\left(-(1+x)^r\left((r+1)x^2-rx+1\right)+(1-x)^r\left((r+1)x^2+rx+1\right)\right)\\
=&\frac{1}{z^2}\left(-f_1(x,r)+f_2(x,r)\right),
\end{aligned}$$

where

$$\begin{aligned}
f_1(x,r)&:=(1+x)^r\left((r+1)x^2-rx+1\right),\\
f_2(x,r)&:=(1-x)^r\left((r+1)x^2+rx+1\right).
\end{aligned}$$

We will show that $f_1(x,r)\geq f_2(x,r)$ holds for $x\in[0,1]$, and for $r\in[0,1)$. Since $f_1(x,r), f_2(x,r)\geq 0$, this is equivalent to showing that $\ln\dfrac{f_1(x,r)}{f_2(x,r)}\geq 0$ holds.

We have

$$\ln \frac{f_1(x,r)}{f_2(x,r)} = r \ln \frac{1+x}{1-x} + \ln\left((r+1)\,x^2 - rx + 1\right) \\ - \ln\left((r+1)\,x^2 + rx + 1\right).$$

We immediately observe that the above sum equals to 0 when $r = 0$. Now, we will show that

$$\frac{\partial}{\partial r} \ln \frac{f_1(x,r)}{f_2(x,r)} \geq 0.$$

Hence, this will prove our claim that $f_1(x,r) \geq f_2(x,r)$ holds.

Taking the first derivative with respect to r, we have

$$\frac{\partial}{\partial r} \ln \frac{f_1(x,r)}{f_2(x,r)} = \ln \frac{1+x}{1-x} - \frac{2x\,(1+x^2)}{(1+(r+1)x^2)^2 - (rx)^2}.$$

So, we will be done if

$$\ln \frac{1+x}{1-x} \geq 2x\left(1+x^2\right) \max_{r\in[0,1)} \frac{1}{(1+(r+1)x^2)^2 - (rx)^2}.$$

One can easily check that the expression in the denominator $(1+(r+1)x^2)^2 - (rx)^2$ is non-decreasing in $r \in [0,1)$, hence the reciprocal is non-increasing in r. As a result, the maximum is attained at $r = 0$. Therefore, we only have to prove that

$$\ln \frac{1+x}{1-x} \geq \frac{2x\,(1+x^2)}{(1+x^2)^2} = \frac{2x}{(1+x^2)}$$

holds. But, we have

$$\ln \frac{1+x}{1-x} = 2x\left(1 + \frac{1}{3}x^2 + \frac{1}{5}x^4 + \frac{1}{7}x^6 + \ldots\right) \geq 2x \geq \frac{2x}{(1+x^2)}.$$

So, $-f_1(x,r) + f_2(x,r) \leq 0$ holds for $r \in [0,1)$ and $x \in [0,1]$. Consequently, Part 2 is decreasing in z. Now, Part 1 is also decreasing in z as

$$\frac{\partial}{\partial z}\left((1+zu)^{\frac{1}{1+\rho}} + (1-zu)^{\frac{1}{1+\rho}}\right) = \frac{u\left((1+uz)^{\frac{-\rho}{1+\rho}} - (1-uz)^{\frac{-\rho}{1+\rho}}\right)}{1+\rho} \leq 0.$$

This proves our claim that the right hand side of (4.45) is non-decreasing in z and concludes the proof. □

4.C Proof of Lemma 4.10

Proof. We prove that the function $H_{z,\rho}(t) = h(\rho, g^{-1}(\rho, t), z)$ defined in (4.37) is concave with respect to the variable t when $\rho \in [0,1] \cup [2,\infty]$, and convex otherwise when $\rho \in [1,2]$, for any fixed $z \in [0,1]$.

Taking the first derivative with respect to t, we get

$$\frac{\partial}{\partial\, t} H_{z,\rho}(t) = \frac{h'(\rho, g^{-1}(\rho,t), z)}{g'(\rho, g^{-1}(\rho,t))}.$$

As we did in Appendix B, we define $u = g^{-1}(\rho, t)$. Since $g(\rho, u)$ is a non-increasing function in u by Lemma 2.2, so is $g^{-1}(\rho, t)$ in t. Hence we can check the concavity of $H_{z,\rho}(t)$ with respect to the variable t, by verifying that $h'(\rho, u, z)/g'(\rho, u)$ is non-decreasing in u. So, we check that

$$\frac{\partial}{\partial u}\left(\frac{h'(\rho,u,z)}{g'(\rho,u)}\right) = \frac{h''(\rho,u,z)g'(\rho,u) - h'(\rho,u,z)g''(\rho,u)}{g'(\rho,u)^2} \geq 0.$$

Since the denominator is always positive, we only need to show that

$$h''(\rho,u,z)g'(\rho,u) - h'(\rho,u,z)g''(\rho,u) \geq 0.$$

Moreover, we observe that $h(\rho, u, 0) = g(\rho, u)$. So, we can equivalently show that the following relation holds:

$$\frac{h''(\rho,u,z)}{h'(\rho,u,z)} \geq \frac{h''(\rho,u,0)}{h'(\rho,u,0)}. \tag{4.48}$$

We first apply the transformations

$$u = \tanh(k) \quad z = \tanh(w)$$

where $k, w \in [0,\infty)$. For shorthand notation, let

$$\widetilde{h}(\rho, k, w) := h(\rho, \tanh(k), \tanh(w)).$$

Using these, we obtain

$$\widetilde{h}(\rho,k,w) = \frac{\cosh\left(\frac{1}{1+\rho}(k+w)\right)^{1+\rho} + \cosh\left(\frac{1}{1+\rho}(k-w)\right)^{1+\rho}}{2\cosh(k)\cosh(w)}.$$

Then, we get

$$\frac{\frac{\partial \widetilde{h}(\rho,k,w)}{\partial k^2}}{\frac{\partial \widetilde{h}(\rho,k,w)}{\partial k}} = -2\tanh(k) + \frac{\rho\cosh(k)}{1+\rho}\times$$
$$\left(\begin{array}{c} \cosh\left(\frac{1}{1+\rho}(k+w)\right)^{\rho-1} + \cosh\left(\frac{1}{1+\rho}(k-w)\right)^{\rho-1} \\ \hline \cosh\left(\frac{1}{1+\rho}(k+w)\right)^{\rho}\sinh\left(\frac{\rho}{1+\rho}k - \frac{1}{1+\rho}w\right) \\ \qquad + \cosh\left(\frac{1}{1+\rho}(k-w)\right)^{\rho}\sinh\left(\frac{\rho}{1+\rho}k + \frac{1}{1+\rho}w\right) \end{array}\right). \tag{4.49}$$

We note that the additive term $-2\tanh(k)$, and the non-negative multiplicative factor $\frac{\rho}{1+\rho}\cosh(k)$ do not depend on w. Hence, we only need to show that the term inside the parenthesis is smallest when evaluated at $w = 0$. For this purpose, we define the transformations

$$a = \frac{k+w}{1+\rho}, \quad b = \frac{k-w}{1+\rho}$$

such that $k = (1+\rho)\dfrac{a+b}{2}$, and $w = (1+\rho)\dfrac{a-b}{2}$. The condition $k, w \geq 0$ is equivalent to $a \geq |b|$. Using these transformations, the reciprocal of the term inside parenthesis in equation (4.49) becomes

$$R(\rho,a,b) := \frac{\cosh(b)^{1-\rho}\cosh(a)\sinh\left(\frac{a+b}{2}\rho - \frac{a-b}{2}\right)}{\cosh(a)^{1-\rho} + \cosh(b)^{1-\rho}} + \frac{\cosh(a)^{1-\rho}\cosh(b)\sinh\left(\frac{a+b}{2}\rho + \frac{a-b}{2}\right)}{\cosh(a)^{1-\rho} + \cosh(b)^{1-\rho}}.$$

Therefore, the inequality given in (4.48) will hold iff

$$R(\rho,a,b) \leq R\left(\rho, \frac{a+b}{2}, \frac{a+b}{2}\right) = \cosh\left(\frac{a+b}{2}\right)\sinh\left(\frac{a+b}{2}\rho\right). \tag{4.50}$$

We define

$$\begin{aligned} f(\rho,a,b) :=& \cosh\left(\frac{a+b}{2}\right)\sinh\left(\rho\frac{a+b}{2}\right)\left(\cosh(a)^{1-\rho} + \cosh(b)^{1-\rho}\right) \\ &- \cosh(a)^{1-\rho}\cosh(b)\sinh\left(\rho\frac{a+b}{2} + \frac{a-b}{2}\right) \\ &- \cosh(b)^{1-\rho}\cosh(a)\sinh\left(\rho\frac{a+b}{2} - \frac{a-b}{2}\right). \end{aligned}$$

We note that $f(\rho,a,b) \geq 0$ is equivalent to the inequality in (4.50), which in turn is

equivalent to the inequality in (4.48).

After simplifications, the function reduces to:

$$f(\rho, a, b) = \sinh\left(\frac{a-b}{2}\right) J(\rho, a, b),$$

where

$$J(\rho, a, b) := \cosh(b)^{1-\rho} \cosh\left(a - \rho\frac{a+b}{2}\right) - \cosh(a)^{1-\rho} \cosh\left(b - \rho\frac{a+b}{2}\right).$$

Since for $a \geq |b|$, we have $\sinh\left(\frac{a-b}{2}\right) \geq 0$, we only need to show that $J(\rho, a, b) \geq 0$.

We introduce the variables κ and ω using $a = \kappa + \omega$, and $b = \kappa - \omega$ where $\kappa, \omega \in [0, \infty)$. Then, we get

$$\begin{aligned} J(\rho, \kappa+\omega, \kappa-\omega) = \cosh(\kappa-\omega)^{1-\rho} \cosh(\kappa - \rho\kappa + \omega) \\ - \cosh(\kappa - \rho\kappa - \omega) \cosh(\kappa+\omega)^{1-\rho}. \end{aligned}$$

We note that $J(\rho, \kappa+\omega, \kappa-\omega)\Big|_{\kappa=0} = 0$. Moreover, $J(\rho, \kappa+\omega, \kappa-\omega)$ is non-decreasing in the variable κ as

$$\begin{aligned} &\frac{\partial}{\partial\kappa} J(\rho, \kappa+\omega, \kappa-\omega) \\ &\qquad = (1-\rho)\left(\cosh(\kappa-\omega)^{-\rho} - \cosh(\kappa+\omega)^{-\rho}\right) \sinh((2-\rho)\kappa) \geq 0, \end{aligned}$$

where the positivity follows from the fact that $|\kappa-\omega| \leq |\kappa+\omega|$, thus $\cosh(\kappa-\omega) \leq \cosh(\kappa+\omega)$ and $\cosh(\kappa-\omega)^{-\rho} \geq \cosh(\kappa+\omega)^{-\rho}$, and from the fact that $\sinh(x) \geq 0$ holds for $\forall x \geq 0$. As a result, $J(\rho, \kappa+\omega, \kappa-\omega) \geq 0$ holds as required. This concludes the proof. □

Chapter 5

Polarization for the Expected Distance $\left|W(0|Y) - W(1|Y)\right|$

It is difficult not to notice that the channel parameter $\Delta_W(y)$ defined in (2.15) and its absolute value are at the heart of the many proofs we carried in the previous chapters. By identifying a B-DMC $W\colon \mathbb{F}_2 \to \mathcal{Y}$ with the random variable $\Delta_W(Y)$, we were able to solve complex optimization problems involving a family of information measures by employing only elementary analysis techniques. In the following chapters, we will continue to use this channel parameter as a driving force in a number of proofs in order to bring different perspectives to the theory of polar coding and extensions to it.

Note that, with Y distributed according to $q_W(y)$ given by (2.14), the random variable $|W(0|Y) - W(1|Y)|$ in the chapter's title does nothing other than expressing $\Delta_W(Y)$ in an alternative way using the posterior probabilities of the channel's inputs given its output. Its expected value, i.e., $E[|\Delta_W(Y)|]$ evaluated under the distribution $q_W(y)$, is thus the *variational distance* between the distributions $W(y|0)$ and $W(y|1)$ over $y \in \mathcal{Y}$. Let us denote this expectation by $T(W)$ and explain why we are interested in the evolution of this channel parameter under the polar transform.

We start by relating $T(W)$ to the maximum likelihood decoding error probability $P_{\text{e, ML}}(W)$ of a single bit transmission over the channel W. For shorthand notation, we define for the likelihood ratio $L_W(Y)$ random variable the following probabilities:

$$\mathbb{P}_{q_W}\left[L_W \gtrsim 1\right] := \mathbb{P}_{q_W}\left[L_W > 1\right] + \frac{1}{2}\mathbb{P}_{q_W}\left[L_W = 1\right],$$
$$\mathbb{P}_{q_W}\left[L_W \lesssim 1\right] := \mathbb{P}_{q_W}\left[L_W < 1\right] + \frac{1}{2}\mathbb{P}_{q_W}\left[L_W = 1\right],$$

such that $\mathbb{P}_{q_W}\left[L_W \gtrsim 1\right] + \mathbb{P}_{q_W}\left[L_W \lesssim 1\right] = 1$. The subscript indicates the distribution

with respect to which the sets are weighted. Then, we write

$$\begin{aligned} T(W) := E[|\Delta_W|] &= \frac{1}{2}\sum_{y\in\mathcal{Y}}|W(y|0) - W(y|1)| \\ &= \frac{1}{2}\left(\mathbb{P}_{W(.|0)}\left[L_W \lneq 1\right] - \mathbb{P}_{W(.|0)}\left[L_W \gneq 1\right]\right. \\ &\qquad \left. - \mathbb{P}_{W(.|1)}\left[L_W \lneq 1\right] + \mathbb{P}_{W(.|1)}\left[L_W \gneq 1\right]\right) \\ &= \frac{1}{2}\left(1 - 2\mathbb{P}_{W(.|0)}\left[L_W \gneq 1\right] + 1 - 2\mathbb{P}_{W(.|1)}\left[L_W \lneq 1\right]\right) \\ &= 1 - 2P_{\text{e, ML}}(W). \end{aligned} \tag{5.1}$$

Suppose now that the channel is almost perfect with $I(W) > 1 - \gamma$, where $\gamma > 0$ is small. Noting that $I(W)$ can be written as

$$\begin{aligned} I(W) &= \sum_{y\in\mathcal{Y}} \frac{1}{2}W(y|0)\log(1 + \Delta_W(y)) + \frac{1}{2}W(y|1)\log(1 - \Delta_W(y)) \\ &= 1 - \mathbb{E}\left[h_2\left(\frac{1+\Delta_W}{2}\right)\right], \end{aligned}$$

it follows from the inequality $h_2((1+\Delta)/2) \geq 1 - |\Delta|$ that $I(W) \leq T(W)$. So, we have $P_{\text{e, ML}}(W) < \gamma$ when $I(W) > 1 - \gamma$. As pointed out before in the introduction of this thesis, transmitting data uncoded over the channel does ensure reliable communication in this situation,.

As usual, we define the processes $T_n(W) := T(W_n)$ and $P_{\text{e, ML}}(W_n)$ associated to the channel polarization process. The second process tracks the error probability of the likelihood ratio based decision rule of the synthetic channels. In this chapter, we investigate, starting with a symmetric B-DMC, the evolution and the convergence properties of $T_n(W)$, $P_{\text{e, ML}}(W_n)$, and various other random processes related to the likelihood ratios of the synthetic channels.

What's Coming, Doc?

The first result we state will be Proposition 5.1 which encompasses in its statement the polarization, loss, and extremality properties of the polar transform for the process $T_n(W)$. Subsequently, we will show in Proposition 5.2 that, for symmetric B-DMCs, $T_n(W)$ and the considered probability processes are bounded sub/super martingales converging to the extremes of their bounded intervals. In particular, we will draw the conclusion that the decision rule using the likelihood ratios of the synthetic channels leads to a bounded submartingale decision error process converging to the extremes $\{0, 0.5\}$, independent of the transmitted input sequence (due to symmetry). Then,

we will apply this knowledge to quickly revisit the theory of channel polarization for symmetric B-DMCs. In the final section, we will cast suspicion upon the audience preparing them for the act of Chapter 7: The Mismatched Capacity of Polar Codes, with the mismatched polar successive cancellation decoder as the leading actor.

As a side note, the extremality of the BEC given in Proposition 5.1 will be later used in Theorem 6.15, where we show that the information sets of symmetric B-DMCs can be squeezed between the information sets of non-trivial BECs. Some of the derivations will also help us in Appendix 8.C when dealing with an approximation to the computations of the likelihood ratios of the synthetic channels.

5.1 Properties of the Polar Transform for the Likelihood Ratios

Recall that in Chapter 4 we defined the random sequence $B_1, \ldots, B_n$ which is drawn i.i.d. according to a Bernoulli distribution with probabilities equal to $1/2$. The process $T_n(W)$ can be described by the following recursion

$$T_{n+1} := \begin{cases} T_n^-, & \text{if } B_{n+1} = 1 \\ T_n^+, & \text{if } B_{n+1} = 0 \end{cases},$$

where $T_n^- := T(W_n^-)$ and $T_n^+ := T(W_n^+)$. The main results of this chapter are reported in the following two propositions. Their proofs are given in Subsection 5.1.3.

Proposition 5.1 (Polarization, Loss, and Extremality Properties)**.** *Let W be a symmetric B-DMC. Then,*

$$\begin{aligned} T_n^- &= T_n^2, \\ T_n^+ &\in \left[T_n, 2T_n - T_n^2\right]. \end{aligned}$$

Moreover, amongst all B-DMCs of the same variational distance $T_0 = T(W)$, the BEC is an extremal channel in the evolution of T_n.

Proposition 5.2 (Gain/Loss & Convergence Laws)**.** *For a symmetric B-DMC W,*

(i) The process T_n is a bounded supermartingale in the interval $[0, 1]$ and converges a.s. to $\{0, 1\}$.

(ii) The process $P_{\text{e, ML}}(W_n)$ is a bounded submartingale in the interval $[0, 0.5]$ and converges a.s. to $\{0, 0.5\}$.

(iii) The process $\mathbb{P}_{q_{W_n}}(L_{W_n} = 1)$ is a bounded submartingale in the interval $[0, 1]$ and converges a.s. to $\{0, 1\}$.

To prove these results, we first review the likelihood ratio recursion of the polar transform and study the one-step properties of this recursion.

5.1.1 Likelihood Ratio Recursion of the Polar Transform

For each $n \geq 0$, the likelihood ratios of the $N = 2^n$ channels $W_N^{(i)} : \mathbb{F}_2 \to \mathcal{Y}^N \times \mathbb{F}_2^{i-1}$, for $i = 1, \ldots, N$, are denoted as $L_{W_N^{(i)}}(y_1^N, u_1^{i-1})$. Recall that the polar successive cancellation decoder decodes the received output in N stages. The decoder sets the estimate $\hat{u}_i$ to its known value on the noisy synthetic channels and uses the decision rule described in (1.8) otherwise.

For a symmetric B-DMC, [2, Corollary 1] shows that the decision error probability of the genie-aided polar decoder is independent of the transmitted input sequence. Hence the analysis of the error probability process and the other likelihood ratio probability processes can be carried by assuming that the all zeros sequence is sent through the channel. For simplicity, we let $L_N^{(i)}(y_1^N) := L_{W_N^{(i)}}(y_1^N, 0_1^{i-1})$. Using (1.10) and (1.11), these likelihood ratios can be computed in a recursive fashion as follows:

$$\begin{aligned} L_{2N}^{(2i-1)}(y_1^{2N}) &= \frac{L_N^{(i)}(y_1^N) + L_N^{(i)}(y_{N+1}^{2N})}{1 + L_N^{(i)}(y_1^N) L_N^{(i)}(y_{N+1}^{2N})}, \\ L_{2N}^{(2i)}(y_1^{2N}) &= L_N^{(i)}(y_1^N) L_N^{(i)}(y_{N+1}^{2N}), \end{aligned}$$

for all $i = 1, \ldots, N = 2^n$. Upon these observations, for symmetric B-DMCs, it will be sufficient for the proofs to focus on the following likelihood ratio process:

$$L_{n+1}(Y_1^{2N}) := \begin{cases} L_n^-(Y_1^{2N}), & \text{if } B_{n+1} = 1 \\ L_n^+(Y_1^{2N}), & \text{if } B_{n+1} = 0 \end{cases}, \tag{5.2}$$

where

$$\begin{aligned} L_n^-(Y_1^{2N}) &= \frac{L_n(Y_1^N) + L_n(Y_{N+1}^{2N})}{1 + L_n(Y_1^N) L_n(Y_{N+1}^{2N})}, \\ L_n^+(Y_1^{2N}) &= L_n(Y_1^N) L_n(Y_{N+1}^{2N}). \end{aligned}$$

Note that by construction Y_1^N and Y_{N+1}^{2N} are i.i.d. random variables observed at the output of the channel $W^N(.|0_1^N)$.

5.1.2 One-Step Properties of the Recursion

The next two lemmas introduce useful expressions for the quantities of interest. Their proofs are given in Appendix 5.A.

Lemma 5.3. *Let L_1 and L_2 be independent random variables. Then,*

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \lesseqgtr 1\right] = \mathbb{P}\left[L_1 \lesseqgtr 1\right]\mathbb{P}\left[L_2 \lesseqgtr 1\right] + \mathbb{P}\left[L_1 \gtreqless 1\right]\mathbb{P}\left[L_2 \gtreqless 1\right], \tag{5.3}$$

$$\begin{aligned}\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \gtreqless 1\right] &= \mathbb{P}\left[L_1 \lesseqgtr 1\right]\mathbb{P}\left[L_2 \gtreqless 1\right] + \mathbb{P}\left[L_1 \gtreqless 1\right]\mathbb{P}\left[L_2 \lesseqgtr 1\right] \\ &= \mathbb{P}\left[L_1 \gtreqless 1\right] + \mathbb{P}\left[L_2 \gtreqless 1\right] - 2\mathbb{P}\left[L_1 \gtreqless 1\right]\mathbb{P}\left[L_2 \gtreqless 1\right], \end{aligned} \tag{5.4}$$

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} = 1\right] = \mathbb{P}\left[L_1 = 1\right] + \mathbb{P}\left[L_2 = 1\right] - \mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right]. \tag{5.5}$$

Lemma 5.4. *Let L_1 and L_2 be independent random variables satisfying the following condition:*

$$\mathbb{P}\left[L_i = \ell\right] = \frac{1}{\ell}\mathbb{P}\left[L_i = \frac{1}{\ell}\right], \quad \forall \ell, \forall i = 1, 2.$$

Then,

$$\mathbb{P}\left[L_1L_2 \lesseqgtr 1\right] = \mathbb{P}\left[L_1 \lesseqgtr 1\right]\mathbb{P}\left[L_2 \lesseqgtr 1\right] + \sum_{\ell_1 \gtreqless 1, \ell_2 \gtreqless 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\max\{\ell_1, \ell_2\}, \tag{5.6}$$

where we abused the notation to define (note the $\gtreqless$ sign in the summation index)

$$\begin{aligned}&\sum_{\ell_1 \gtreqless 1, \ell_2 \gtreqless 1} \mathbb{P}\left[L = \ell_1\right]\mathbb{P}\left[L = \ell_2\right]\max\{\ell_1, \ell_2\} \\ &\quad := \sum_{\ell_1 > 1, \ell_2 > 1} \mathbb{P}\left[L = \ell_1\right]\mathbb{P}\left[L = \ell_2\right]\max\{\ell_1, \ell_2\} + \frac{1}{4}\mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right] \\ &\qquad + \frac{1}{2}\mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 < 1\right] + \frac{1}{2}\mathbb{P}\left[L_1 < 1\right]\mathbb{P}\left[L_2 = 1\right],\end{aligned}$$

and

$$\mathbb{P}\left[L_1L_2 \gtreqless 1\right] = \mathbb{P}\left[L_1 \gtreqless 1\right]\mathbb{P}\left[L_2 \gtreqless 1\right] + \sum_{\ell_1 \gtreqless 1, \ell_2 \gtreqless 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\min\{\ell_1, \ell_2\}, \tag{5.7}$$

where

$$\begin{aligned}&\sum_{\ell_1 \gtreqless 1, \ell_2 \gtreqless 1} \mathbb{P}\left[L = \ell_1\right]\mathbb{P}\left[L = \ell_2\right]\min\{\ell_1, \ell_2\} \\ &\quad := \sum_{\ell_1 > 1, \ell_2 > 1} \mathbb{P}\left[L = \ell_1\right]\mathbb{P}\left[L = \ell_2\right]\min\{\ell_1, \ell_2\} + \frac{1}{4}\mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right] \\ &\qquad + \frac{1}{2}\mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 > 1\right] + \frac{1}{2}\mathbb{P}\left[L_1 > 1\right]\mathbb{P}\left[L_2 = 1\right].\end{aligned}$$

Using the above lemmas, we will derive the likelihood ratio versions of the polarization and conservation properties of the polar transform.

Proposition 5.5 (Polarization Property)**.** *Let L_1 and L_2 be as defined in Lemma 5.4. Without loss of generality, suppose $\mathbb{P}[L_1 \gtrsim 1] \leq \mathbb{P}[L_2 \gtrsim 1]$. Given that $\mathbb{P}[L_i \gtrsim 1] \leq \mathbb{P}[L_i \lesssim 1]$, for $i = 1, 2$, the recursive likelihood ratio transformation satisfies*

$$\mathbb{P}[L_1 L_2 \gtrsim 1] \leq \mathbb{P}[L_1 \gtrsim 1] \leq \mathbb{P}[L_2 \gtrsim 1] \leq \mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \gtrsim 1\right],$$

$$\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \lesssim 1\right] \leq \mathbb{P}[L_2 \lesssim 1] \leq \mathbb{P}[L_1 \lesssim 1] \leq \mathbb{P}[L_1 L_2 \lesssim 1].$$

Proof. We first prove the inequalities for the minus transform. By assumption $\mathbb{P}[L_i \gtrsim 1] \in [0, 0.5]$, for $i = 1, 2$. So, using the expression of Lemma 5.3, we get

$$\begin{aligned}\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \gtrsim 1\right] &= \mathbb{P}[L_1 \gtrsim 1] + \mathbb{P}[L_2 \gtrsim 1] - 2\mathbb{P}[L_1 \gtrsim 1]\,\mathbb{P}[L_2 \gtrsim 1] \\ &\geq \max\{\mathbb{P}[L_2 \gtrsim 1], \mathbb{P}[L_1 \gtrsim 1]\}.\end{aligned}$$

(For $f(a, b) = a + b - 2ab$ with $a, b \in [0, 0.5]$ such that $a \leq b$, we have $f(a, b) \geq \max\{a, b\}$ as the function is increasing in a and $f(0, b) = b$). This also implies

$$\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \lesssim 1\right] \leq \min\{\mathbb{P}[L_2 \lesssim 1], \mathbb{P}[L_1 \lesssim 1]\}.$$

Next, we prove the inequalities for the plus transform. Using Lemma 5.4, we have

$$\begin{aligned}&\mathbb{P}[L_1 L_2 \gtrsim 1] \\ =&\mathbb{P}[L_1 \gtrsim 1]\,\mathbb{P}[L_2 \gtrsim 1] + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}[L_1 = \ell_1]\,\mathbb{P}[L_2 = \ell_2] \min\{\ell_1, \ell_2\} \\ \leq&\mathbb{P}[L_1 \lesssim 1]\,\mathbb{P}[L_2 \gtrsim 1] + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}[L_1 = \ell_1]\,\mathbb{P}[L_2 = \ell_2]\,\ell_2 \\ =&\mathbb{P}[L_1 \gtrsim 1]\,\mathbb{P}[L_2 \gtrsim 1] + \mathbb{P}[L_1 \gtrsim 1]\,\mathbb{P}[L_2 \lesssim 1] \\ =&\mathbb{P}[L_1 \gtrsim 1] \leq \mathbb{P}[L_2 \gtrsim 1].\end{aligned}$$

This also implies

$$\mathbb{P}[L_1 L_2 \lesssim 1] \geq \max\{\mathbb{P}[L_2 \lesssim 1], \mathbb{P}[L_1 \lesssim 1]\},$$

and concludes the proof. □

Proposition 5.6 (Gain/Loss Properties)**.** *Let L_1 and L_2 be as defined in Lemma 5.4. Then, the following set of inequalities hold:*

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \gtrsim 1\right] + \mathbb{P}\left[L_1L_2 \gtrsim 1\right] \geq \mathbb{P}\left[L_1 \gtrsim 1\right] + \mathbb{P}\left[L_2 \gtrsim 1\right], \tag{5.8}$$

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \lesssim 1\right] + \mathbb{P}\left[L_1L_2 \lesssim 1\right] \leq \mathbb{P}\left[L_1 \lesssim 1\right] + \mathbb{P}\left[L_2 \lesssim 1\right], \tag{5.9}$$

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} = 1\right] + \mathbb{P}\left[L_1L_2 = 1\right] \geq \mathbb{P}\left[L_1 = 1\right] + \mathbb{P}\left[L_2 = 1\right]. \tag{5.10}$$

Hence, we also have

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} < 1\right] + \mathbb{P}\left[L_1L_2 < 1\right] \leq \mathbb{P}\left[L_1 < 1\right] + \mathbb{P}\left[L_2 < 1\right],$$

$$\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \geq 1\right] + \mathbb{P}\left[L_1L_2 \geq 1\right] \geq \mathbb{P}\left[L_1 \geq 1\right] + \mathbb{P}\left[L_2 \geq 1\right].$$

Proof. We start by proving the inequality in (5.8). Using the expressions derived in Lemmas 5.3 and 5.4, we get

$$\begin{aligned}
&\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} \gtrsim 1\right] + \mathbb{P}\left[L_1L_2 \gtrsim 1\right] \\
=&\mathbb{P}\left[L_1 \gtrsim 1\right] + \mathbb{P}\left[L_2 \gtrsim 1\right] - 2\mathbb{P}\left[L_1 \gtrsim 1\right]\mathbb{P}\left[L_2 \gtrsim 1\right] \\
&+ \mathbb{P}\left[L_1 \gtrsim 1\right]\mathbb{P}\left[L_2 \gtrsim 1\right] + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\min\{\ell_1, \ell_2\} \\
=&\mathbb{P}\left[L_1 \gtrsim 1\right] + \mathbb{P}\left[L_2 \gtrsim 1\right] - \mathbb{P}\left[L_1 \gtrsim 1\right]\mathbb{P}\left[L_2 \gtrsim 1\right] \\
&\qquad + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\min\{\ell_1, \ell_2\} \\
\geq&\mathbb{P}\left[L_1 \gtrsim 1\right] + \mathbb{P}\left[L_2 \gtrsim 1\right] - 2\mathbb{P}\left[L_1 \gtrsim 1\right]\mathbb{P}\left[L_2 \gtrsim 1\right],
\end{aligned}$$

where the inequality follows from

$$\sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\min\{\ell_1, \ell_2\} \geq \mathbb{P}\left[L_1 \gtrsim 1\right]\mathbb{P}\left[L_2 \gtrsim 1\right]. \tag{5.11}$$

This also proves (5.9) in view of the relation $\mathbb{P}\left[L \lesssim 1\right] = 1 - \mathbb{P}\left[L \gtrsim 1\right]$. Finally, to prove (5.10), we write

$$\begin{aligned}
&\mathbb{P}\left[\frac{L_1+L_2}{1+L_1L_2} = 1\right] + \mathbb{P}\left[L_1L_2 = 1\right] \\
&\geq \mathbb{P}\left[L_1 = 1\right] + \mathbb{P}\left[L_2 = 1\right] - \mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right] + \mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right],
\end{aligned}$$

where we used Lemmas 5.3, 5.4, and $\mathbb{P}\left[L_1L_2 = 1\right] \geq \mathbb{P}\left[L_1 = 1\right]\mathbb{P}\left[L_2 = 1\right]$. □

The final proposition of this subsection shows that the condition in the hypothesis of the Polarization Property Lemma 5.5 is preserved under the polar transform.

Proposition 5.7. *Suppose that $\mathbb{P}[L_i \gtrsim 1] \leq \mathbb{P}[L_i \lesssim 1]$, for $i = 1, 2$. Then, the polar transform preserves this inequality relation for the likelihood ratios, i.e,*

$$\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \gtrsim 1\right] \leq \mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \lesssim 1\right] \quad \textit{and} \quad \mathbb{P}[L_1 L_2 \gtrsim 1] \leq \mathbb{P}[L_1 L_2 \lesssim 1].$$

Proof. From Lemma 5.3, we get

$$\begin{aligned}\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \lesssim 1\right] - \mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \gtrsim 1\right] \\ = \left(\mathbb{P}[L_1 \lesssim 1] - \mathbb{P}[L_1 \gtrsim 1]\right)\left(\mathbb{P}[L_2 \lesssim 1] - \mathbb{P}[L_2 \gtrsim 1]\right) \geq 0 \quad (5.12)\end{aligned}$$

where the non-negativity follows by the hypothesis of the proposition. The claim for the plus transform can be obtained straightforwardly by inspecting (5.6) and (5.7) derived in Lemma (5.4). □

5.1.3 Proofs of Propositions 5.1 and 5.2

Before giving the proofs, we introduce a property of symmetric channels that will be used. If W is a symmetric B-DMC, then by using $W(y|0) = W(y|1)/L_W(y)$, one can derive the following property:

$$\mathbb{P}_{W(y|0)}[L_W = \ell] = \frac{1}{\ell}\mathbb{P}_{W(y|0)}\left[L_W = \frac{1}{\ell}\right]. \quad (5.13)$$

Moreover, we note that for a symmetric B-DMC W, $T_n(W)$ reduces to

$$T_n(W) = \mathbb{P}_{W_n(.|0)}[L_{W_n} \lesssim 1] - \mathbb{P}_{W_n(.|0)}[L_{W_n} \gtrsim 1],$$

For simplicity, we drop the subscripts to denote: $T_n = \mathbb{P}[L_n \lesssim 1] - \mathbb{P}[L_n \gtrsim 1]$. Similarly, $\mathbb{P}[L_n \gtrsim 1] := P_{\text{e, ML}}(W_n)$ and $\mathbb{P}[L_n = 1] := \mathbb{P}_{q_{W_n}}[L_{W_n} = 1]$.

Proof of Proposition 5.1. First, we prove the relation $T_n^+ \leq 2T_n - T_n^2$. By using (5.6) and (5.7), we obtain

$$\begin{aligned}T_n^+ &= \mathbb{P}[L_n \lesssim 1]^2 - \mathbb{P}[L_n \gtrsim 1]^2 \\ &\quad + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}[L_n = \ell_1]\,\mathbb{P}[L_n = \ell_2]\,(\max\{\ell_1, \ell_2\} - \min\{\ell_1, \ell_2\}) \\ &= T_n + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}[L_n = \ell_1]\,\mathbb{P}[L_n = \ell_2]\,(\max\{\ell_1, \ell_2\} - \min\{\ell_1, \ell_2\}).\end{aligned}$$

Note that

$$\begin{aligned} T_n - T_n^2 = T_n(1-T_n) &= \left(\mathbb{P}\left[L_n \lesssim 1\right] - \mathbb{P}\left[L_n \gtrsim 1\right]\right) 2\mathbb{P}\left[L_n \gtrsim 1\right] \\ &= 2\mathbb{P}\left[L_n \gtrsim 1\right]\mathbb{P}\left[L_n \lesssim 1\right] - 2\mathbb{P}\left[L_n \gtrsim 1\right]^2, \end{aligned}$$

as $1 - T_n = 2\mathbb{P}\left[L_n \gtrsim 1\right]$. Therefore,

$$\begin{aligned} 2T_n - T_n^2 - T_n^+ = {} & 2\mathbb{P}\left[L_n \gtrsim 1\right]\mathbb{P}\left[L_n \lesssim 1\right] - 2\mathbb{P}\left[L_n \gtrsim 1\right]^2 + T_n - T_n \\ & - \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_n = \ell_1\right]\mathbb{P}\left[L_n = \ell_2\right]\left(\max\{\ell_1,\ell_2\} - \min\{\ell_1,\ell_2\}\right). \end{aligned}$$

Since,

$$2\mathbb{P}\left[L_n \gtrsim 1\right]\mathbb{P}\left[L_n \lesssim 1\right] = \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_n = \ell_1\right]\mathbb{P}\left[L_n = \ell_2\right]\left(\max\{\ell_1,\ell_2\} + \min\{\ell_1,\ell_2\}\right),$$

we get

$$\begin{aligned} 2T_n - T_n^2 - T_n^+ = {} & -2\mathbb{P}\left[L_n \gtrsim 1\right]^2 \\ & + \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_n = \ell_1\right]\mathbb{P}\left[L_n = \ell_2\right]\left(\max\{\ell_1,\ell_2\} + \min\{\ell_1,\ell_2\}\right) \\ & - \sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_n = \ell_1\right]\mathbb{P}\left[L_n = \ell_2\right]\left(\max\{\ell_1,\ell_2\} - \min\{\ell_1,\ell_2\}\right) \\ = {} & 2\sum_{\ell_1 \gtrsim 1, \ell_2 \gtrsim 1} \mathbb{P}\left[L_n = \ell_1\right]\mathbb{P}\left[L_n = \ell_2\right]\min\{\ell_1,\ell_2\} - 2\mathbb{P}\left[L_n \gtrsim 1\right]^2 \geq 0, \end{aligned}$$

where the non-negativity follows by (5.11).

Now, we prove the rest of the proposition. From (5.12) (taking L_1 and L_2 identically distributed as L_n), we immediately get $T_n^- = T_n^2$. We also note that Proposition 5.5 implies $T_n^+ \geq T_n$. Upon noticing that being a BEC is preserved under the polar transform [2] with $T_n^+ = 2T_n - T_n^2$, the final claim follows. □

Proof of Proposition 5.2. The symmetry assumption on the channel W implies via (5.13) the following inequality:

$$\mathbb{P}\left[L_W \lesssim 1\right] > \mathbb{P}\left[L_W \gtrsim 1\right]. \tag{5.14}$$

So, by Proposition 5.7, we get $\mathbb{P}\left[L_n \gtrsim 1\right] \leq \mathbb{P}\left[L_n \lesssim 1\right]$, for all $n = 1, 2, \ldots$, and thus, the probabilities are constrained as follows: $\mathbb{P}\left[L_n \gtrsim 1\right] \in [0, 0.5]$, $\mathbb{P}\left[L_n \lesssim 1\right] \in [0.5, 1]$, and $\mathbb{P}\left[L_n = 1\right] \in [0, 1]$. From this the boundedness statements follow.

Proposition 5.1 and the inequalities proved in Proposition 5.6 show the processes are the claimed sub/super martingales. From general results on bounded martingales,

it follows the processes converge a.s. The only part left is to prove the convergence is to the extremes of the bounded intervals. For the process T_n, we know by Proposition 5.1 that $T_n^- = T_n^2$. One can complete the proof that T_n converges to the extremes using this relation in a similar fashion as in the proof of [2, Proposition 9] of the convergence to the extremes of the Bhattacharyya process of the synthetic channels associated with the polar transformations: As $\mathbb{E}[|T_{n+1} - T_n|] \xrightarrow[n\to\infty]{} 0$ holds, we have

$$\mathbb{E}[|T_{n+1} - T_n|] \geq \frac{1}{2}\mathbb{E}[T_n\,(1 - T_n)] \xrightarrow[n\to\infty]{} 0,$$

whence $T_\infty \in \{0, 1\}$. Similarly, as by Lemma 5.3 (taking L_1 and L_2 identically distributed as L_n), we have $\mathbb{P}\left[L_n^- = 1\right] = 2\mathbb{P}\left[L_n = 1\right] - \mathbb{P}\left[L_n = 1\right]^2$, one can show $\mathbb{P}\left[L_\infty = 1\right] \in \{0, 1\}$. Once T_n and $\mathbb{P}\left[L_n = 1\right]$ converge to these values, the remaining probabilities can only converge to the extremes claimed by the proposition. □

5.2 Channel Polarization and Rate of Convergence Revisited

Now, we revisit the theory of channel polarization for symmetric B-DMCs by looking to the four possible combinations of the pair T_∞ and $P_{\text{e, ML}}(W_\infty) = \mathbb{P}\left[L_\infty \gneq 1\right]$, two of which we hopefully 'never' end up with.

1. $T_\infty = 1, \mathbb{P}\left[L_\infty \gneq 1\right] = 0.5$: As $T_\infty = \mathbb{P}[L_\infty \lneq 1] - \mathbb{P}[L_\infty \gneq 1] = 1$, we find $\mathbb{P}[L_\infty \lneq 1] = 1.5$, which is a contradiction. So, this case is not possible.
2. $T_\infty = 1, \mathbb{P}\left[L_\infty \gneq 1\right] = 0$: We look at a perfect channel.
3. $T_\infty = 0, \mathbb{P}\left[L_\infty \gneq 1\right] = 0.5$: We look at a completely noisy channel.
4. $T_\infty = 0, \mathbb{P}\left[L_\infty \gneq 1\right] = 0$: The two constraints tell us $\mathbb{P}[L_\infty \lneq 1] = 0$, and the second one implies $\mathbb{P}[L_\infty \lneq 1] = 1 - \mathbb{P}[L_\infty \gneq 1] = 1$. We end up again with a contradiction. So, this case is not possible either.

Note that we still need the preservation of the sum capacities to show that the fraction of perfect channels is $I(W)$. Next, we show that the results on the rate of convergence of polar codes [30] can be stated in terms of T_n.

Proposition 5.8. *For any $\beta < 1/2$,*

$$\lim_{n\to\infty} \mathbb{P}[T_n < 2^{-2^{n\beta}}] = 1 - I(W).$$

Proof. The conditions (z.1), (z.2), (z.3) in [30] still hold with the Bhattacharyya process Z_n replaced by T_n, and the condition $\mathbb{P}\left[Z_\infty = 0\right] = I_0$ in (z.3) replaced by $\mathbb{P}\left[T_\infty = 0\right] = 1 - I_0$. □

5.3 Detective, What If We Track the Wrong Process?

In this final section, we make a small digression to explore the difficulties in extending the previous results to mismatched processes. The impatient reader may simply skip this part and come back while reading Chapter 7 where we study the mismatched capacity of polar codes.

So far, we have shown that the processes tracking the evolution of the synthetic channels' likelihood ratio probabilities exhibit martingale properties. Inspired by these results, we define the mismatched version of those processes and investigate their evolution under the polar transformations.

We know that for the channel processes $W_n(\cdot|u)$, for $u \in \{0,1\}$, we can write

$$\mathbb{P}_{q_{W_n}}[\cdot] = \frac{1}{2}\mathbb{P}_{W_n(\cdot|0)}[\cdot] + \frac{1}{2}\mathbb{P}_{W_n(\cdot|1)}[\cdot].$$

As before, we will use the following notations

$$\begin{aligned}\mathbb{P}_{W_n(\cdot|u)}[L_{V_n} \gtrsim 1] &:= \mathbb{P}_{W_n(\cdot|u)}[L_{V_n} > 1] + \frac{1}{2}\mathbb{P}_{W_n(\cdot|u)}[L_{V_n} = 1],\\ \mathbb{P}_{W_n(\cdot|u)}[L_{V_n} \lesssim 1] &:= \mathbb{P}_{W_n(\cdot|u)}[L_{V_n} < 1] + \frac{1}{2}\mathbb{P}_{W_n(\cdot|u)}[L_{V_n} = 1],\end{aligned}$$

such that $\mathbb{P}_{W_n(\cdot|u)}[L_{V_n} \gtrsim 1] + \mathbb{P}_{W_n(\cdot|u)}[L_{V_n} \lesssim 1] = 1$. In addition, we define two other channel parameters:

$$\begin{aligned}T^0(W_n, V_n) &:= \mathbb{P}_{W_n(\cdot|0)}[L_{V_n} \lesssim 1] - \mathbb{P}_{W_n(\cdot|0)}[L_{V_n} \gtrsim 1]\\ &= \mathbb{P}_{W_n(\cdot|0)}[L_{V_n} < 1] - \mathbb{P}_{W_n(\cdot|0)}[L_{V_n} > 1],\end{aligned} \tag{5.15}$$

$$\begin{aligned}T^1(W_n, V_n) &:= \mathbb{P}_{W_n(\cdot|1)}[L_{V_n} \gtrsim 1] - \mathbb{P}_{W_n(\cdot|1)}[L_{V_n} \lesssim 1]\\ &= \mathbb{P}_{W_n(\cdot|1)}[L_{V_n} > 1] - \mathbb{P}_{W_n(\cdot|1)}[L_{V_n} < 1].\end{aligned} \tag{5.16}$$

Note that both $T^0, T^1 \in [-1, 1]$. The following lemma studies the evolution of these parameters under the minus polar transform. The proof is given in Appendix 5.B.

Lemma 5.9.

$$\begin{aligned}T^0(W^-, V^-) &= \frac{1}{2}T^0(W,V)^2 + \frac{1}{2}T^1(W,V)^2,\\ T^1(W^-, V^-) &= T^0(W,V)T^1(W,V).\end{aligned}$$

Now, replacing the parameters with the mismatched versions, i.e.,

$$T_n(W,V) := T(W_n, V_n) = \frac{T^0(W_n, V_n) + T^1(W_n, V_n)}{2}, \tag{5.17}$$

$T_n^- := T(W_n^-, V_n^-)$, and $T_n^+ := T(W_n^+, V_n^+)$, we note that the process satisfies $T_n^- = T_n^2$ by Lemma 5.9. So, if the process converges almost surely, the only possibility is to $\{0, 1\}$. Moreover, in this case the mismatched capacities of the synthetic channels will either be 0 or 1. The main problem here is to prove that T_n converges almost surely as we do not have a good characterization for T_n^+.

Finally, the next proposition investigates the convergence properties of the process $\mathbb{P}_{W_n}[L_{V_n} = 1]$. The proof is given in Appendix 5.B.

Proposition 5.10. *$\mathbb{P}_{q_{W_n}}[L_{V_n} = 1]$ is a bounded submartingale process taking values in $[0, 1]$ and the process converges a.s. to $\{0, 1\}$.*

Proposition 5.10 gives a partial hint on the convergence points of T_n: A synthetic channel has $T(W_N^{(i)}, V_N^{(i)}) = 0$ whenever $\mathbb{P}_{W_N^{(i)}}\left[L_{V_N^{(i)}} = 1\right] = 1$ holds.

Appendix

In this Appendix, we prove Lemma 5.3 and Lemma 5.4 in the first part and Lemma 5.9 and Proposition 5.10 in the second part.

5.A Proofs of Lemmas 5.3 and 5.4

Proof of Lemma 5.3. The claim in (5.5) can be easily verified. To prove the remaining ones, we first observe that

$$\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} > 1\right] = \mathbb{P}[L_1 < 1]\,\mathbb{P}[L_2 > 1] + \mathbb{P}[L_1 > 1]\,\mathbb{P}[L_2 < 1]. \tag{5.18}$$

Using (5.5) and (5.18), we obtain

$$\begin{aligned}\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \gneq 1\right] &= \mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} > 1\right] + \frac{1}{2}\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} = 1\right] \\ &= \mathbb{P}[L_1 \lneq 1]\,\mathbb{P}[L_2 \gneq 1] + \mathbb{P}[L_1 \gneq 1]\,\mathbb{P}[L_2 \lneq 1],\end{aligned}$$

as

$$\begin{aligned}&\mathbb{P}[L_1 \lneq 1]\,\mathbb{P}[L_2 \gneq 1] \\ &\qquad = \mathbb{P}[L_1 < 1]\,\mathbb{P}[L_2 > 1] + \frac{1}{2}\mathbb{P}[L_1 = 1] - \frac{1}{4}\mathbb{P}[L_1 = 1]\,\mathbb{P}[L_2 = 1].\end{aligned}$$

As a result, we also get

$$\mathbb{P}\left[\frac{L_1 + L_2}{1 + L_1 L_2} \lneq 1\right] = \mathbb{P}[L_1 \lneq 1]\,\mathbb{P}[L_2 \lneq 1] + \mathbb{P}[L_1 \gneq 1]\,\mathbb{P}[L_2 \gneq 1]. \qquad \square$$

Proof of Lemma 5.4. We start by proving (5.6) using the lemma's assumption:

$$
\begin{aligned}
&\mathbb{P}\left[L_1 L_2 \lesssim 1\right] \\
=&\sum_{\ell_1<1}\sum_{\ell_2<1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&+\sum_{\ell_1<1}\sum_{1\le\ell_2<1/\ell_1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&+\sum_{\ell_1\ge 1}\sum_{\ell_2\le 1/\ell_1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&-\frac{1}{2}\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2=1\right]. \\
=&\mathbb{P}\left[L_1<1\right]\mathbb{P}\left[L_2<1\right]-\frac{1}{2}\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2=1\right] \\
&+\sum_{\ell_1>1}\sum_{1\le\ell_2<\ell_1}\ell_1\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&+\sum_{\ell_1\ge 1}\sum_{\ell_2\ge\ell_1}\ell_2\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
=&\mathbb{P}\left[L_1<1\right]\mathbb{P}\left[L_2<1\right]-\frac{1}{2}\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2=1\right] \\
&+\sum_{\ell_1>1}\sum_{1<\ell_2<\ell_1}\ell_1\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&+\mathbb{P}\left[L_2=1\right]\sum_{\ell_1>1}\ell_1\mathbb{P}\left[L_1=\ell_1\right] \\
&+\sum_{\ell_1>1}\sum_{\ell_2\ge\ell_1}\ell_2\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right] \\
&+\mathbb{P}\left[L_1=1\right]\sum_{\ell_2>1}\ell_2\mathbb{P}\left[L_1=\ell_2\right]+\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2=1\right] \\
=&\mathbb{P}\left[L_1<1\right]\mathbb{P}\left[L_2<1\right]+\frac{1}{2}\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2=1\right] \\
&+\sum_{\ell_1>1}\sum_{\ell_2>1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right]\max\{\ell_1,\ell_2\} \\
&+\mathbb{P}\left[L_1=1\right]\mathbb{P}\left[L_2<1\right]+\mathbb{P}\left[L_1<1\right]\mathbb{P}\left[L_2=1\right] \\
=&\mathbb{P}\left[L_1\lesssim 1\right]\mathbb{P}\left[L_2\lesssim 1\right]+\sum_{\ell_1\gtrsim 1}\sum_{\ell_2\gtrsim 1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right]\max\{\ell_1,\ell_2\}.
\end{aligned}
$$

To prove (5.7), we first note that

$$
\begin{aligned}
&\sum_{\ell_1\gtrsim 1}\sum_{\ell_2\gtrsim 1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right]\left(\max\{\ell_1,\ell_2\}+\min\{\ell_1,\ell_2\}\right) \\
=&\sum_{\ell_1\gtrsim 1}\sum_{\ell_2\gtrsim 1}\mathbb{P}\left[L_1=\ell_1\right]\mathbb{P}\left[L_2=\ell_2\right]\left(\ell_1+\ell_2\right) \\
=&\mathbb{P}\left[L_1\lesssim 1\right]\mathbb{P}\left[L_2\gtrsim 1\right]+\mathbb{P}\left[L_1\gtrsim 1\right]\mathbb{P}\left[L_2\lesssim 1\right].
\end{aligned}
$$

Since $1 = \mathbb{P}\left[L_1 L_2 \lneq 1\right] + \mathbb{P}\left[L_1 L_2 \gneq 1\right] = \left(\mathbb{P}\left[L \lneq 1\right] + \mathbb{P}\left[L \gneq 1\right]\right)^2$, we get

$$\mathbb{P}\left[L_1 L_2 \gneq 1\right] = \mathbb{P}\left[L_1 \gneq 1\right]\mathbb{P}\left[L_2 \gneq 1\right] + \sum_{\ell_1 \gneq 1}\sum_{\ell_2 \gneq 1} \mathbb{P}\left[L_1 = \ell_1\right]\mathbb{P}\left[L_2 = \ell_2\right]\min\{\ell_1, \ell_2\}. \qquad \square$$

5.B Proofs of Lemma 5.9 and Proposition 5.10

Proof of Lemma 5.9. We let $W_0 := W(.|0)$ and $W_1 := W(.|1)$.

$$\begin{aligned} T^0(W^-, V^-) &= \sum_{\substack{y_1^2: \\ L_{V^-}<1}} W^-(y_1^2|0) - \sum_{\substack{y_1^2: \\ L_{V^-}>1}} W^-(y_1^2|0) \\ &= \frac{1}{2}\mathbb{P}_{W_0}\left[L_V < 1\right]^2 + \frac{1}{2}\mathbb{P}_{W_0}\left[L_V > 1\right]^2 + \frac{1}{2}\mathbb{P}_{W_1}\left[L_V < 1\right]^2 \\ &\quad + \frac{1}{2}\mathbb{P}_{W_1}\left[L_V > 1\right]^2 - \mathbb{P}_{W_0}\left[L_V < 1\right]\mathbb{P}_{W_0}\left[L_V > 1\right] \\ &\quad - \mathbb{P}_{W_1}\left[L_V < 1\right]\mathbb{P}_{W_1}\left[L_V > 1\right] \\ &= \frac{1}{2}T^0(W, V)^2 + \frac{1}{2}T^1(W, V)^2. \end{aligned}$$

$$\begin{aligned} T^1(W^-, V^-) &= \sum_{\substack{y_1^2: \\ L_{V^-}>1}} W^-(y_1^2|1) - \sum_{\substack{y_1^2: \\ L_{V^-}<1}} W^-(y_1^2|1) \\ &= \mathbb{P}_{W_0}\left[L_V < 1\right]\mathbb{P}_{W_1}\left[L_V > 1\right] + \mathbb{P}_{W_0}\left[L_V > 1\right]\mathbb{P}_{W_1}\left[L_V < 1\right] \\ &\quad - \mathbb{P}_{W_0}\left[L_V < 1\right]\mathbb{P}_{W_1}\left[L_V < 1\right] - \mathbb{P}_{W_0}\left[L_V > 1\right]\mathbb{P}_{W_1}\left[L_V > 1\right] \\ &= T^1(W, V)T^0(W, V). \end{aligned}$$ $\square$

Proof of Proposition 5.10. For the minus transformation, we get

$$\begin{aligned} \mathbb{P}_{q_W}\left[L_{V^-} = 1\right] &= \sum_{\substack{y_1^2: \\ L_{V^-}=1}} \frac{1}{2}W^-(y_1^2|0) + \frac{1}{2}W^-(y_1^2|1) \\ &= \sum_{\substack{y_1^2: \\ L_{V^-}=1}} \frac{1}{4}\left(W(y_1|0)W(y_2|0) + W(y_1|1)W(y_2|1)\right) \\ &\quad + \frac{1}{4}\left(W(y_1|1)W(y_2|0) + W(y_1|0)W(y_2|1)\right) \\ &= \mathbb{P}_{W_0}\left[L_V = 1\right] + \mathbb{P}_{W_1}\left[L_V = 1\right] - \frac{1}{4}\left(\mathbb{P}_{W_0}\left[L_V = 1\right] + \mathbb{P}_{W_1}\left[L_V = 1\right]\right)^2 \\ &= 2\mathbb{P}_{q_W}\left[L_V = 1\right] - \mathbb{P}_{q_W}\left[L_V = 1\right]^2. \end{aligned}$$

For the plus transformation, we get

$$\begin{aligned}
\mathbb{P}_{q_W}\left[L_{V^+}=1\right] &= \sum_{\substack{y_1^2u_1:\\ L_{V^+}=1}} \frac{1}{2}W^+(y_1^2u_1|0)+\frac{1}{2}W^+(y_1^2u_1|1)\\
&\geq \sum_{\substack{y_1:\\ L_V=1}}\sum_{\substack{y_2:\\ L_V=1}} \frac{1}{4}\left(W(y_1|0)W(y_2|0)+W(y_1|1)W(y_2|0)\right)\\
&\qquad + \sum_{\substack{y_1:\\ L_V=1}}\sum_{\substack{y_2:\\ L_V=1}} \frac{1}{4}\left(W(y_1|1)W(y_2|1)+W(y_1|0)W(y_2|1)\right)\\
&\geq \frac{1}{4}\left(\mathbb{P}_{W_0}\left[L_V=1\right]+\mathbb{P}_{W_1}\left[L_V=1\right]\right)^2 = \mathbb{P}_{q_W}\left[L_V=1\right]^2.
\end{aligned}$$

Hence after a single step, the following inequality holds:

$$\mathbb{P}_{q_W}\left[L_{V^-}=1\right]+\mathbb{P}_{q_W}\left[L_{V^+}=1\right] \geq 2\mathbb{P}_{q_W}\left[L_V=1\right].$$

By the recursive structure, the inequality holds at every step, proving the claim that the process is a submartingale. The boundedness statement is trivial. One can complete the proof that the convergence is to the extremes in a similar fashion as in the proof of [2, Proposition 9] of the convergence to the extremes of the Bhattacharyya process of the synthetic channels associated with the polar transformations:

$$\begin{aligned}
&\mathbb{E}\Big[\big|\mathbb{P}_{q_{W_{n+1}}}\left[L_{V_{n+1}}=1\right]-\mathbb{P}_{q_{W_n}}\left[L_{V_n}=1\right]\big|\Big] \xrightarrow[n\to\infty]{} 0\\
&\qquad\Longrightarrow\ \mathbb{E}\Big[\big|\mathbb{P}_{q_{W_{n+1}}}\left[L_{V_{n+1}}=1\right]-\mathbb{P}_{q_{W_n}}\left[L_{V_n}=1\right]\big|\Big]\\
&\qquad\qquad \geq \frac{1}{2}\mathbb{E}\big[\mathbb{P}_{q_{W_n}}\left[L_{V_n}=1\right]\left(1-\mathbb{P}_{q_{W_n}}\left[L_{V_n}=1\right]\right)\big] \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

showing $\mathbb{P}_{q_{W_\infty}}\left[L_{V_\infty}=1\right]\in\{0,1\}$. □

Chapter 6

Order Preserving Properties of the Polar Transform

Three nice properties of the polar transform were emphasized in the last two chapters; by showing these properties on different information measures, we revisited and generalized the main lines of the original theory. The primary objective was to better understand how the polar transform operates on a *given* B-DMC. This implicit assumption behind the developed theory must, however, be questioned at the design stage of polar codes. From next chapter onward, we will deviate from the single point to point communication channel model to study the robustness of polar coding under more complex communication scenarios. In this transition chapter, we embark the investigation starting from the information sets of polar codes. Our objective is to characterize how the polar transform operates on the set of B-DMCs in a rigorous framework.

To set up, think about the elegant principle behind the construction of the information set of a polar code for a given B-DMC W. Independent copies of the channel W are combined and split by applying the polar transform in a recursive fashion. After a long sequence of such operations, the synthesized channels cluster eventually in two states, perfect or noisy. As the main idea behind the construction of the information set is to ensure that the overall error probability of the decoding procedure is small, the information set of a polar code of block-length $N = 2^n$, with $n = 1, 2, \ldots$, for the channel W, denoted as $\mathcal{A}_N(W)$, is specified by picking from the set $\{+, -\}^n$ the indices of the channels which are good for uncoded transmission, i.e.,

$$\mathcal{A}_N(W) = \big\{ s^n \in \{+, -\}^n : \; W^{s^n} \text{ is '}\textit{good}\text{'} \big\}. \tag{6.1}$$

The economy of concepts used to describe this task is stunning and accomplishing the task itself sounds quite simple. Nevertheless, a hidden difficulty arises: comput-

ing efficiently the transition probabilities of the synthetic channels of huge output alphabet sizes. Initially, [2] solved this problem by proposing to approximate the computations by estimating the good channels via their Bhattacharyya distance with the help of the Monte Carlo method. However, this approach had two limitations: complexity and reliability of the Monte Carlo estimates. Mori and Tanaka [31] proposed a better method based on density evolution, but implementing the method with sufficient precision still required further investigation. Thanks to Tal and Vardy, this problem of practical interest was thought out in [3] with an algorithm to carry the computations approximately (but within guaranteed bounds) and efficiently. The fact that polar codes can be explicitly constructed, and so efficiently, is certainly a big step toward the practice of polar coding.

Another 'not so hidden' characteristic of (6.1) is the reliance of the definition on a specific channel. This apparent observation led to a question of both theoretical and practical interest after the invention of polar codes: How large is $\mathcal{A}_N(W) \cap \mathcal{A}_N(V)$, for two given channels W and V. Two partial orders have been pointed out in [2] which order the information sets of polar codes: Any BEC provides good indices for all other B-DMCs having smaller Bhattacharyya parameters, and any channel which is degraded with respect to another B-DMC provides good indices for the upgraded channel[1]. In this chapter, we will show that these partial orderings can be studied in the context of stochastic orders known as convex orderings. Interestingly, it will turn out that the solution to the efficient computation problem found in [3] is closely tied to the notion of convex ordering.

What's Coming, Doc?

Many channel parameters can be used to quantify "good" in (6.1); for instance, the symmetric capacity or the Bhattacharyya parameter are among popular choices. In fact, any channel parameter appearing in a meaningful upper bound to $P_{\text{e, ML}}(W)$ defined in (1.9) is eligible as this bound would apply individually to the synthetic channels and would thus serve to upper bound the error probability of polar codes via the union bound. In this chapter, we consider a family of such quantifiers generated by the following class of functions

$$\mathcal{F}_{\text{cx, s}} := \big\{ f_s \colon [-1,1] \to [0,1] : \ f_s \text{ is } \textit{symmetric}^2 \text{ and } \textit{convex} \\ \text{such that } f_s(0) = 0 \text{ and } f_s(1) = 1 \big\}. \quad (6.2)$$

The functions in $\mathcal{F}_{\text{cx, s}}$ will take as argument the channel parameter $\Delta_W(y)$ we defined in (2.15) as the normalized difference between $W(y|0)$ and $W(y|1)$.

[1] W is upgraded with respect to V if and only if V is degraded with respect to W

[2] A function $f(\delta)$ is called *symmetric* if $f(\delta) = f(-\delta)$, for all $\delta \in \mathbb{R}$

Recall from the previous chapter that $T(W) = \mathbb{E}\left[|\Delta_W|\right]$ computes the variational distance between the two distributions $W(y|0)$ and $W(y|1)$ and satisfies $T(W) = 1 - 2P_{\text{e, ML}}(W)$. As $T(W) > 1 - \gamma$ implies $P_{\text{e, ML}}(W) < \gamma$, to have the latter small, the channel at hand must have a large variational distance (close to 1). Equivalently, it would be sufficient that

$$T_{f_s}(W) := \mathbb{E}\left[f_s\left(\Delta_W\right)\right]$$

be large for any $f_s \in \mathcal{F}_{\text{cx, s}}$, since

$$T(W) \geq T_{f_s}(W)$$

always holds. Upon noticing $T_{f_s}(W) \in [0, 1]$, we conclude via $T(W) = 1 - 2P_{\text{e, ML}}(W)$ that the parameters $T_{f_s}(W)$ generate a family of upper bounds to $P_{\text{e, ML}}(W)$. Based on these results, we refine the vague definition in (6.1) as follows.

Definition 6.1. Let $f_s \in \mathcal{F}_{\text{cx, s}}$ and $\gamma \in (0, 1)$ be a threshold. W is called 'γ-good' if $T_{f_s}(W) \geq 1 - \gamma$ holds. Accordingly, the *information set* definition is adapted as

$$\mathcal{A}_N^{f_s,\gamma}(W) := \left\{s^n \in \{+,-\}^n : \; T_{f_s}(W^{s^n}) \geq 1 - \gamma\right\}.$$

For instance, the particular choice of $f_s(\delta) = 1 - h_2(\frac{1+\delta}{2})$, where $h_2(.)$ denotes the binary entropy function, or $f(\delta) = 1 - \sqrt{1 - \delta^2}$ lead to information set definitions based on the values of the symmetric capacities and the Bhattacharyya parameters of the synthetic channels, respectively.

In this chapter, we will show that, in essence, taking $\Delta_W(Y)$ as argument, the class of symmetric convex functions generates a partial ordering for B-DMCs which orders the information sets of polar codes:

$$\mathcal{A}_N^{f_s,\gamma}(V) \subset \mathcal{A}_N^{f_s,\gamma}(W), \forall N, \forall \gamma \quad \text{if} \quad T_{f_s}(V) \leq T_{f_s}(W), \text{for all } f_s \in \mathcal{F}_{\text{cx, s}}.$$

This result will follow as a corollary to Theorem 6.5 which will show that the generalized polar transform $\langle W_1, W_2\rangle^{\pm}$ described in (4.13) and (4.14) preserves symmetric convex orderings.

Once the theorem will be proved, we will compare this ordering with the stochastic degradation ordering already known to order the information sets of polar codes. We will show that while for symmetric channels this ordering is equivalent to stochastic degradation, a strictly weaker partial order is obtained when at least one of the channels is asymmetric. In particular, we will illustrate this by an example which studies both orderings between a Z-channel and a binary symmetric channel whose inputs are used with equal frequency. In the process, we will also present tools which can be useful for verifying the symmetric convex ordering: the cut criterion due to

[32] and channel symmetrization.

In the final two sections, we will show that after the polar transform is applied, the channels are ordered with the symmetric convex ordering and we will discuss how the symmetric convex ordering is useful for the efficient construction of polar codes. Finally, keep in mind as a spoiler alert that the results of Theorem 6.5 will find applications in Chapter 8 and Chapter 9 when dealing with the problem of universal polar coding with channel knowledge at the decoder.

6.1 A Novel Partial Ordering for B-DMCs: The Symmetric Convex Ordering

First and foremost, we designate the novel ordering.

Definition 6.2. We say that two B-DMCs $W : \mathbb{F}_2 \to \mathcal{Y}_1$ and $V : \mathbb{F}_2 \to \mathcal{Y}_2$ satisfy the *symmetric convex ordering* if

$$\mathbb{E}\left[f_s(\Delta_V)\right] \leq \mathbb{E}\left[f_s(\Delta_W)\right],$$

for all functions $f_s \in \mathcal{F}_{\text{cx, s}}$. The ordering is denoted by $V \prec_{\text{cx, s}} W$.

Next, we bridge this definition with a well known stochastic order. Let Δ_1 and Δ_2 be two random variables with distributions F_{Δ_1} and F_{Δ_2}, respectively.

Definition 6.3. [33] Δ_1 is smaller with respect to the *increasing convex ordering* (*decreasing concave ordering*) than Δ_2, written $\Delta_1 \prec_{\text{icx}} \Delta_2$ ($\Delta_1 \prec_{\text{dcv}} \Delta_2$), if

$$\mathbb{E}\left[f(\Delta_1)\right] \leq \mathbb{E}\left[f(\Delta_2)\right], \tag{6.3}$$

for all increasing convex (decreasing concave) functions f, for which the expectations exist.

As any result involving the $\prec_{\text{icx}}$ ordering can be mapped to the $\prec_{\text{dcv}}$ ordering, we will stick to the first one. Alternatively, this ordering can be described by using only the class of symmetric functions.

Proposition 6.4. *$|\Delta_1| \prec_{\text{icx}} |\Delta_2|$ if and only if*

$$\mathbb{E}\left[f(\Delta_1)\right] \leq \mathbb{E}\left[f(\Delta_2)\right],$$

for all convex and symmetric functions f, for which the expectations exist.

Proof. The proof follows by the fact that $f(|\delta|) = f(\delta)$ holds for any symmetric function $f(\delta)$, $\delta \in \mathbb{R}$. □

Thus, the new partial ordering introduced in Definition 6.2 is an increasing convex ordering for the absolute value of the channels' Δ parameters.

Now, we are ready to state the main result.

Theorem 6.5. *Let $W_1 \colon \mathbb{F}_2 \to \mathcal{Y}_1$, $W_2 \colon \mathbb{F}_2 \to \mathcal{Y}_2$, $V_1 \colon \mathbb{F}_2 \to \mathcal{Y}_3$, and $V_2 \colon \mathbb{F}_2 \to \mathcal{Y}_4$ be B-DMCs such that*

$$|\Delta_{V_1}| \prec_{\text{icx}} |\Delta_{W_1}| \quad \textit{and} \quad |\Delta_{V_2}| \prec_{\text{icx}} |\Delta_{W_2}|$$

hold. Then, the polar transform preserves this ordering, i.e., $\left|\Delta_{V_{1,2}^{\pm}}\right| \prec_{\text{icx}} \left|\Delta_{W_{1,2}^{\pm}}\right|$.

Proof of Theorem 6.5. We will use the characterization given in Proposition 6.4 in the proof. After applying the polar transform to the channels, one can derive the following recursion

$$\Delta_{W_{1,2}^{-}}(Y_1Y_2) = \Delta_{W_1}(Y_1)\Delta_{W_2}(Y_2), \tag{6.4}$$

$$\Delta_{W_{1,2}^{+}}(Y_1Y_2U_1) = \frac{\Delta_{W_1}(Y_1) + (-1)^{U_1}\Delta_{W_2}(Y_2)}{1 + (-1)^{U_1}\Delta_{W_1}(Y_1)\Delta_{W_2}(Y_2)}, \tag{6.5}$$

where

$$Y_1Y_2 \sim q_{W_1}(y_1)q_{W_2}(y_2),$$

and

$$Y_1Y_2U_1 \sim q_{W_1}(y_1)q_{W_2}(y_2)\frac{1 + (-1)^{u_1}\Delta_{W_1}(y_1)\Delta_{W_2}(y_2)}{2}.$$

See the proofs of Lemma 4.1 and Lemma 4.2 for a proof.

Let $f(\delta)$ be a function which is convex and symmetric in $\delta \in [-1, 1]$. Note that the function must be increasing in $\delta \in [0, 1]$. For the minus polar transform, we write

$$\sum_{y_1 \in \mathcal{Y}_1} \sum_{y_2 \in \mathcal{Y}_2} q_{W_{1,2}^{-}}(y_1y_2) f\left(\Delta_{W_{1,2}^{-}}(y_1y_2)\right) = \sum_{y_1 \in \mathcal{Y}_1} q_{W_1}(y_1) \sum_{y_2 \in \mathcal{Y}_2} q_{W_2}(y_2) f^{-}\left(\Delta_{W_1}(y_1), \Delta_{W_2}(y_2)\right)$$

where $f^{-}(\delta_1, \delta_2) = f(\delta_1\delta_2)$, for $\delta_1, \delta_2 \in [-1, 1]$. As we assumed $f(\delta)$ to be convex and symmetric in its argument, so is f^{-} in both of its arguments. Similarly for the plus polar transform, we write

$$\sum_{y_1 \in \mathcal{Y}_1} \sum_{y_2 \in \mathcal{Y}_2} \sum_{u_1 \in \mathbb{F}_2} q_{W_{1,2}^{+}}(y_1y_2u_1) f\left(\Delta_{W_{1,2}^{+}}(y_1y_2u_1)\right) = \sum_{y_1 \in \mathcal{Y}_1} q_{W_1}(y_1) \sum_{y_2 \in \mathcal{Y}_2} q_{W_2}(y_2) f^{+}\left(\Delta_{W_1}(y_1), \Delta_{W_2}(y_2)\right),$$

where

$$f^+(\delta_1, \delta_2) = \frac{1+\delta_1\delta_2}{2} f\left(\frac{\delta_1+\delta_2}{1+\delta_1\delta_2}\right) + \frac{1-\delta_1\delta_2}{2} f\left(\frac{\delta_1-\delta_2}{1-\delta_1\delta_2}\right), \tag{6.6}$$

for $\delta_1, \delta_2 \in [-1, 1]$. Lemma 6.18 in Appendix 6.A shows that f^+ is also a convex and symmetric function in both of its arguments.

So, using the assumptions $|\Delta_{V_1}| \prec_{\text{icx}} |\Delta_{W_1}|$ and $|\Delta_{V_2}| \prec_{\text{icx}} |\Delta_{W_2}|$, we deduce that

$$\begin{aligned}
&\sum_{y_3 \in \mathcal{Y}_3} q_{V_1}(y_3) \sum_{y_4 \in \mathcal{Y}_4} q_{V_2}(y_4) f^{\pm}\left(\Delta_{V_1}(y_3), \Delta_{V_2}(y_4)\right) \\
&\qquad \le \sum_{y_3 \in \mathcal{Y}_3} q_{V_1}(y_3) \sum_{y_2 \in \mathcal{Y}_2} q_{W_2}(y_2) f^{\pm}\left(\Delta_{V_1}(y_3), \Delta_{W_2}(y_2)\right) \\
&\qquad = \sum_{y_2 \in \mathcal{Y}_2} q_{W_2}(y_2) \sum_{y_3 \in \mathcal{Y}_3} q_{V_1}(y_3) f^{\pm}\left(\Delta_{V_1}(y_3), \Delta_{W_2}(y_2)\right) \\
&\qquad \le \sum_{y_2 \in \mathcal{Y}_2} q_{W_2}(y_2) \sum_{y_1 \in \mathcal{Y}_1} q_{W_1}(y_1) f^{\pm}\left(\Delta_{W_1}(y_1), \Delta_{W_2}(y_2)\right).
\end{aligned}$$

This proves our claim that $\left|\Delta_{V_{1,2}^{\pm}}\right| \prec_{\text{icx}} \left|\Delta_{W_{1,2}^{\pm}}\right|$. □

Using the refined definition of the information set in (6.1), we get the following corollary to the previous theorem.

Corollary 6.6. *Let $W\colon \mathbb{F}_2 \to \mathcal{Y}_1$ and $V\colon \mathbb{F}_2 \to \mathcal{Y}_2$ be such that $V \prec_{\text{cx, s}} W$ as defined in Definition 6.2. Then,*

$$\mathcal{A}_N^{f_s,\gamma}(V) \subset \mathcal{A}_N^{f_s,\gamma}(W), \tag{6.7}$$

holds for all $f_s \in \mathcal{F}_{\text{cx, s}}$, for any $\gamma \in (0, 1)$, and for all $N = 2^n$ with $n = 1, 2, \ldots$.

Proof. The assumption $V \prec_{\text{cx, s}} W$ on the channels implies via Proposition 6.4 that $|\Delta_V| \prec_{\text{icx}} |\Delta_W|$ holds. Then, (6.7) follows by Theorem 6.5. □

As we pointed out earlier, it is stated in [2] that the information sets of polar codes are ordered for stochastically degraded channels. See [34, Lemma 4.7] for a proof of the fact that stochastic degradation is preserved under the original polar transform and Appendix 2.B for the fact that two stochastically degraded DMCs are ordered in their E_0 parameters. It would therefore be of interest to compare the symmetric convex ordering we introduced with stochastic degradation.

6.2 Exploration

6.2.1 Convex Ordering

The material up to and including Theorem 6.11 is drawn from [33, Section 1.3]. The following definition introduces a special case of the increasing convex ordering.

Definition 6.7. [33, Theorem B] Suppose Δ_1 and Δ_2 have equal mean values, i.e., $\mathbb{E}[\Delta_1] = \mathbb{E}[\Delta_2]$. We say Δ_1 is smaller with respect to the *convex ordering* than Δ_2, written $\Delta_1 \prec_{\text{cx}} \Delta_2$, if and only if

$$\mathbb{E}\left[f(\Delta_1)\right] \leq \mathbb{E}\left[f(\Delta_2)\right],$$

for all convex functions f for which the expectations exist.

Definition 6.8. [33] A *Markov kernel* is a function $K_{\text{M}}(\delta, E)$, $\delta \in \mathbb{R}$, $E \in \mathscr{B}$, such that $K_{\text{M}}(\delta, .)$ is a probability measure on $\mathbb{R}$ for each fixed δ and $K_{\text{M}}(., E)$ is a measurable function for each fixed E. K_{M} is mean value preserving if the mean value of the probability measure $K_{\text{M}}(\delta, .)$ is equal to δ.

An alternative description of the convex ordering due to Blackwell [35] is given in [33, Theorem C]. Below is the statement of this theorem.

Theorem 6.9. *[35]* $\Delta_1 \prec_{\text{cx}} \Delta_2$ *if and only if there exists a mean value preserving Markov kernel* K_{M} *such that* $F_{\Delta_2} = K_{\text{M}} F_{\Delta_1}$*, i.e.,*

$$F_{\Delta_2}(\delta_2) = \mathbb{E}\big[K_{\text{M}}\big(\Delta_1, (-\infty, \delta_2]\big)\big].$$

A random variable Δ is called *symmetric* if the distribution of Δ satisfies $F_{\Delta}(\delta) = 1 - F_{\Delta}(-\delta)$, for all $\delta \in \mathbb{R}$. In the next proposition, we exploit this symmetry property.

Proposition 6.10. *Let* Δ_1 *and* Δ_2 *be symmetric random variables, then* $\Delta_1 \prec_{\text{cx}} \Delta_2$ *if and only if* $|\Delta_1| \prec_{\text{icx}} |\Delta_2|$.

Proof. The 'only if part' follows by definition. So, we only need to prove the 'if part'. Let $f(\delta)$ be a convex function in $\delta \in \mathbb{R}$. As Δ_1 is symmetric, we can write

$$\mathbb{E}\left[f(\Delta_1)\right] = \mathbb{E}\left[\frac{f(\Delta_1) + f(-\Delta_1)}{2}\right] = \mathbb{E}\left[f_s(\Delta_1)\right],$$

where $f_s(\delta) = \left(f(\delta) + f(-\delta)\right)/2$ is a convex symmetric function. In particular, $f_s(.)$ is increasing on $\mathbb{R}_+$. Hence using $|\Delta_1| \prec_{\text{icx}} |\Delta_2|$, we get

$$\mathbb{E}\left[f(\Delta_1)\right] = \mathbb{E}\left[f_s(|\Delta_1|)\right] \leq \mathbb{E}\left[f_s(|\Delta_2|)\right] = \mathbb{E}\left[f(\Delta_2)\right]. \qquad \square$$

Now, we show that for symmetric channels the convex ordering is equivalent to the stochastic degradation ordering. Let $V\colon \mathbb{F}_2 \to \mathcal{Y}_2$ be stochastically degraded with respect to $W\colon \mathbb{F}_2 \to \mathcal{Y}_1$. Then, by definition 2.6, we know that there exists a channel $P\colon \mathcal{Y}_1 \to \mathcal{Y}_2$ such that (2.27) holds. In this case, one can derive the following:

$$\Delta_V(z) = \frac{V(z|0) - V(z|1)}{V(z|0) + V(z|1)} = \sum_{y\in\mathcal{Y}_1} \bar{P}(y|z)\Delta_W(y),$$

for $z \in \mathcal{Y}_2$, where

$$\bar{P}(y|z) = \frac{q_W(y)P(z|y)}{\sum\limits_{y\in\mathcal{Y}_1} q_W(y)P(z|y)}$$

corresponds to the inputs posterior probabilities given the output of the channel P. So, for any convex function $f(.)$, we obtain

$$\begin{aligned}
\mathbb{E}\left[f(\Delta_V)\right] &= \sum_{z\in\mathcal{Y}_2} q_V(z) f(\Delta_V(z)) \\
&= \sum_{z\in\mathcal{Y}_2} \left(\sum_{y\in\mathcal{Y}_1} q_W(y)P(z|y)\right) f\left(\sum_{y\in\mathcal{Y}_1} \bar{P}(y|z)\Delta_W(y)\right) \\
&\leq \sum_{z\in\mathcal{Y}_2}\sum_{y\in\mathcal{Y}_1} q_W(y)P(z|y) f(\Delta_W(y)) \\
&= \sum_{y\in\mathcal{Y}_1} q_W(y) f(\Delta_W(y)) = \mathbb{E}\left[f(\Delta_W)\right], \qquad (6.8)
\end{aligned}$$

where the inequality follows by Jensen's inequality. In particular, the ordering holds with equality for the function $f(\delta) = \delta$. Hence, degradation preserves the mean value, i.e., $E[\Delta_W] = E[\Delta_V]$. By Definition 6.7, we conclude the order relation $\Delta_V \prec_{\text{cx}} \Delta_W$ holds for stochastically degraded channels.

To show the reverse implication, suppose the channels satisfy $\Delta_V \prec_{\text{cx}} \Delta_W$. By Theorem 6.9, there exists a Markov kernel K_{M} such that

$$\sum_{y\in\mathcal{Y}_1} K_{\text{M}}(z,y) = 1, \qquad (6.9)$$

$$\Delta_V(z) = \sum_{y\in\mathcal{Y}_1} K_{\text{M}}(z,y)\Delta_W(y), \qquad (6.10)$$

$$\mathbb{P}\left[\Delta_W(y) = \delta_y\right] = \sum_{z\in\mathcal{Y}_2} K_{\text{M}}(z,y)\mathbb{P}\left[\Delta_V(z) = \delta_z\right], \qquad (6.11)$$

for all $y \in \mathcal{Y}_1$ and $z \in \mathcal{Y}_2$. Note that (6.11) is equivalent to

$$q_W(y) = \sum_{z\in\mathcal{Y}_2} K_{\text{M}}(z,y) q_V(z), \qquad (6.12)$$

and from (6.10), we get

$$V(z|0) - V(z|1) = \sum_{y \in \mathcal{Y}_1} \widetilde{K}_{\mathrm{M}}(z, y) \left[W(y|0) - W(y|1) \right], \tag{6.13}$$

where

$$\widetilde{K}_{\mathrm{M}}(z, y) = \frac{q_V(z)}{q_W(y)} K_{\mathrm{M}}(z, y). \tag{6.14}$$

Now, observe that via (6.12), we have

$$\sum_{z \in \mathcal{Y}_2} \widetilde{K}_{\mathrm{M}}(z, y) = 1.$$

Moreover, taking the denominator $q_W(y)$ in (6.14) to the other side, summing over $y \in \mathcal{Y}_1$, and using (6.9), we get

$$V(z|0) + V(z|1) = \sum_{y \in \mathcal{Y}_1} \widetilde{K}_{\mathrm{M}}(z, y) \left[W(y|0) + W(y|1) \right]. \tag{6.15}$$

Combining (6.13) and (6.15) gives

$$V(z|x) = \sum_{y \in \mathcal{Y}_1} \widetilde{K}_{\mathrm{M}}(z, y) W(y|x),$$

for $x \in \mathbb{F}_2$. This proves that convex ordering implies stochastic degradation as $\widetilde{K}_{\mathrm{M}}(z, y)$ is of the form $P(z|y)$ given in Definition 2.6. The equivalence is proved.

6.2.2 Tools for Verifying the Symmetric Convex Ordering

As the symmetric convex ordering between two channels can be described via the increasing convex ordering of their $|\Delta|$ parameters, we can borrow any tool from the literature used to verify the latter. In the next theorem, a 'simple' criterion, known as the *Karlin-Novikoff cut criterion* [32], is given for the increasing convex ordering.

Theorem 6.11. *[33, Theorem E] Suppose that for Δ_1, Δ_2 with finite first moments $m_{\Delta_1} = \mathbb{E}[\Delta_1]$ and $m_{\Delta_2} = \mathbb{E}[\Delta_2]$, we have $m_{\Delta_1} \leq m_{\Delta_2}$ and*

$$F_{\Delta_1}(\delta) \leq F_{\Delta_2}(\delta), \quad \text{for } \delta \leq c, \tag{6.16}$$

$$F_{\Delta_1}(\delta) \geq F_{\Delta_2}(\delta), \quad \text{for } \delta > c, \tag{6.17}$$

for some $c \in \mathbb{R}$, then $\Delta_1 \prec_{\mathrm{icx}} \Delta_2$.

Although we will not make use of in this thesis, a more general version of the cut criterion called Karlin-Novikoff-Stoyan-Taylor crossing conditions for stop-loss order can be found in [36]. The theorem provides a necessary and sufficient condition

for the *stop-loss order* which is the name given to the increasing convex ordering in the actuarial science literature.

In the comparison process, the following idea will also be useful in the process of checking our ordering.

Definition 6.12. [34, Definition 1.3] For any B-DMC $W\colon \mathcal{X} \to \mathcal{Y}$, the *symmetrized* B-DMC $W_s\colon \mathcal{X} \to \mathcal{Y} \times \mathcal{X}$ is defined as

$$W_s(y, z|x) = \frac{1}{2} W(y|x \oplus z).$$

6.2.3 Novelty of the Ordering by an Example

We saw that any channel V which satisfies the relation $\Delta_V \prec_{\text{cx}} \Delta_W$ with any other channel W is in fact stochastically degraded with respect to W. It is also clear by definition that the convex ordering between the channel random variables implies the symmetric convex ordering introduced in Definition 6.2. So, we need to study the reverse implication to decide whether the symmetric convex ordering condition of Theorem 6.5 gives a strictly weaker condition than convex ordering (stochastic degradation). At this point, by recalling the equivalence stated in Proposition 6.10, we notice that this is not the case for symmetric channels as the two orders $\Delta_V \prec_{\text{cx}} \Delta_W$ and $|\Delta_V| \prec_{\text{icx}} |\Delta_W|$ are equivalent for symmetric channels. The purpose of this subsection is to show that no equivalence exists between the symmetric convex ordering and stochastic degradation if one of the two channels is asymmetric. If we can find a pair of B-DMCs that does not satisfy stochastic degradation, but satisfies the symmetric convex ordering, we will be done. Such a pair is illustrated in the next example.

Example 6.13. Let W be a *Z-channel*, as shown in Figure 6.1, with crossover probability $r \in [0, 1]$ and V be a BSC with crossover probability $p \in [0, 0.5]$. In this example, we will answer the following three questions:

($q1$) Suppose V is a stochastically degraded version of W. What is the best possible BSC (with the smallest p) which satisfies this condition?

($q2$) Suppose instead that the channels satisfy the symmetric convex ordering $|\Delta_V| \prec_{\text{icx}} |\Delta_W|$. What is the best possible BSC which satisfies this condition?

($q3$) Suppose we first symmetrize W according to Definition 6.12 to construct W_s. Suppose now V is a stochastically degraded version of W_s. What is the best possible BSC which satisfies this condition?

Then, we will compare the three BSCs to decide which ordering results in a better

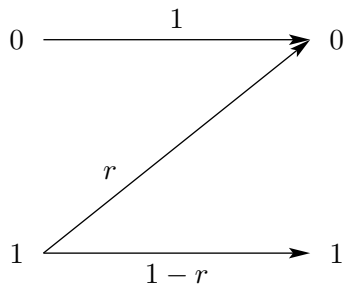

Figure 6.1: Z-channel.

channel with a smaller p, and thus, leads to a larger information subset for the Z-channel. Here are the answers.

$(a1)$ *Stochastic degradation:* Let us derive the range of possible values of p in terms of r under this assumption. For this purpose, we define the asymmetric binary channel P degrading W to V by

$$V(z|x) = \sum_{y \in \mathbb{F}_2} W(y|x)P(z|y). \tag{6.18}$$

First we note that $P(0|0) = 1 - p$ and $P(0|1) = p$ are the only possibilities. Let $P(0|1) = \alpha$. Then, using (6.18), we get

$$V(0|1) = p = (1 - r)\alpha + r(1 - p),$$

which implies

$$p = \frac{r + (1 - r)\alpha}{1 + r}. \tag{6.19}$$

Noting that the right hand side of (6.19) is increasing in $\alpha \in [0, 1]$, we conclude that

$$\frac{r}{1 + r} \leq p \leq \frac{1}{1 + r}$$

whenever we impose stochastic degradation on the channels. Picking the BSC having the smallest crossover probability $p = r/(1 + r)$ answers the first question.

$(a2)$ $|\Delta_V| \prec_{\text{icx}} |\Delta_W|$*:* Now, we will derive the range of possible values of p in terms of r under this assumption by using the cut-criterion given in Theorem 6.11. We start by computing the values of $E[|\Delta_V|]$ and $E[|\Delta_W|]$ in terms of the channel parameters. For the BSC, we have $E[|\Delta_V|] = 1 - 2p$. For the Z-channel, we have

$$|\Delta_W(y)| = \begin{cases} \dfrac{1 - r}{1 + r}, & \text{if } y = 0 \\ 1, & \text{if } y = 1 \end{cases}, \tag{6.20}$$

$q_W(0) = (1 + r)/2$, and $q_W(1) = (1 - r)/2$. So, we compute $E[|\Delta_W|] = 1 - r$. As

any B-DMC together with any BSC with crossover probability p will always satisfy the conditions (6.16) and (6.17) of Theorem 6.11 with $c = |1 - 2p|$ and F being the cumulative distribution of the BSC, we can see by the theorem's statement that the condition $E[|\Delta_V|]] \leq E[|\Delta_W|]]$ is necessary in our example for $|\Delta_V| \prec_{\text{icx}} |\Delta_W|$ to hold. This in turn implies that $p \geq r/2$. Hence, the best possible BSC in this case has crossover probability $p = r/2$. This answers the second question.

$(a3)$ *Channel symmetrization:* We first note a more general result: a given B-DMC W' and its symmetrized version W'_s always satisfy $|\Delta_{W'_s}(y, z)| = |\Delta_{W'}(y)|$ with $|\Delta_{W'_s}(y, z)|$ distributed as $q_{W'}(y)/2$, for $z \in \mathbb{F}_2$. Therefore, for any function $f(\delta)$ defined for $\delta \in [0, 1]$, we have

$$\mathbb{E}\left[f(|\Delta_{W'}|)\right] = \mathbb{E}\left[f(|\Delta_{W'_s}|)\right].$$

We conclude that for any two B-DMCs W' and V': $|\Delta_{V'}| \prec_{\text{icx}} |\Delta_{W'}|$ if and only if $|\Delta_{V'_s}| \prec_{\text{icx}} |\Delta_{W'_s}|$. Moreover, as the channels in this last condition are symmetric, we know the condition holds if and only if $\Delta_{V'_s} \prec_{\text{cx}} \Delta_{W'_s}$, i.e., the symmetrized versions of the channels are ordered by stochastic degradation. So, we have the same answer as in the previous case: the best possible BSC in this case has also crossover probability $p = r/2$.

Let us compare the results. Noting that $r/2 \leq r/(1 + r)$ holds for any $r \in [0, 1]$, and with equality if and only if $r = 0, 1$, we conclude that, for $r \in (0, 1)$, the BSC with smallest crossover probability is found by the symmetric convex ordering and this BSC is not stochastically degraded with respect to the Z-channel. For instance, when $r = 1/2$, the crossover probabilities of the best BSC we found in the second case is $1/4$ compared to $1/3$ in the first one. Finally, we also showed that one can verify the symmetric convex ordering by first symmetrizing the asymmetric channels and then checking for stochastic degradation. The example proves that for general B-DMCs the symmetric convex ordering is strictly weaker than stochastic degradation.

6.3 How to prepare a BEC sandwhich?

In this section, we discuss two other orderings related to BECs. First we note the following property of the BEC.

Proposition 6.14. *Amongst the set of symmetric B-DMCs with a given fixed value of the channels' variational distance T, the BEC U of erasure probability $1 - T(U)$ maximizes the symmetric capacity and minimizes the Bhattacharyya parameter.*

Proof. The proof follows by noting $T(U) = 1 - Z(U)$ for the BEC and using

the following upper bounds to the uncoded error probability over any B-DMC W: $(1-T(W))/2 \leq (1-I(W))/2$ and $(1-T(W))/2 \leq Z(W)/2$. □

For simplicity, we define $Z = |\Delta_W|$. Suppose a BEC BEC with erasure probability $\epsilon \in [0,1]$ and a B-DMC W satisfy $\mathbb{E}[Z] \leq E[Z_{BEC}]$. As Z_{BEC} is $\{0,1\}$ valued with $P(Z_{BEC}=0)=\epsilon$, Z_{BEC} and any arbitrary random variable Z satisfy the conditions (6.16) and (6.17) of Theorem 6.11 with G being the cumulative distribution of Z_{BEC}. As a result, the assumption $\mathbb{E}[Z] \leq E[Z_{BEC}]$ implies $Z \prec_{\text{icx}} Z_{BEC}$. By Theorem 6.5, we know that this ordering is preserved under the polar transform

Another case of increasing convex ordering slightly different than Theorem 6.5 happens when BEC and W are such that the Bhattacharyya parameters of the channels satisfy $Z(W) \leq Z(BEC)$ (beware that the random variable Z is different than $Z(W)$!). Let us define the random variable $B = \sqrt{1-Z^2}$. Then, $Z(W) = \mathbb{E}[B_W]$. Hence, the channels satisfy $\mathbb{E}[B] \leq \mathbb{E}[B_{BEC}]$. One more time, letting G denote the cumulative distribution of B_{BEC} in Theorem 6.11, we see that $\mathbb{E}[B] \leq \mathbb{E}[B_{BEC}]$ implies $B \prec_{\text{icx}} B_{BEC}$. Finally, it is well known from [2, Proposition 6] that this ordering is also preserved under the polar transform.

Using these two BEC orderings, the following theorem shows that the information set of a given symmetric B-DMC can be squeezed between the information sets of two BECs.

Theorem 6.15. *For any given symmetric B-DMC W with parameter values $T(W)$ and $Z(W)$, define the BEC U such that $T(U) = T(W)$ and the BEC V such that $Z(V) = Z(W)$. Then, we have $Z(U_N^{(i)}) \leq Z(W_N^{(i)}) \leq Z(V_N^{(i)})$ for any $i = 1, \ldots$ and $N = 2^n$ with $n \geq 0$. Furthermore, this implies the following ordering of the information sets:*

$$\mathcal{A}_N^{f_B,\gamma}(V) \subset \mathcal{A}_N^{f_B,\gamma}(W) \subset \mathcal{A}_N^{f_B,\gamma}(U), \quad \forall \gamma \in [0,1].$$

for the function $f_B(\delta) = 1 - \sqrt{1-\delta^2}$.

Proof. It is already known that the BEC V provides universally good indices: $\mathcal{A}_N^{f_B,\gamma}(V) \subset \mathcal{A}_N^{f_B,\gamma}(W)$ [2]. The proof that the BEC U provides 'universally bad' indices follows by the extremality properties stated in Propositions 5.1 and 6.14. □

6.4 Polarization Property

The following lemma proves that the polarization property of the polar transform holds for all the channel parameters $T_{f_s}(W)$, where $f_s \in \mathcal{F}_{\text{cx, s}}$.

Lemma 6.16 (Polarization Property)**.** *For any two B-DMCs W_1 and W_2,we have*

$$T_{f_s}(W_{1,2}^-) \leq T_{f_s}(W_1) \leq T_{f_s}(W_{1,2}^+),$$
$$T_{f_s}(W_{1,2}^-) \leq T_{f_s}(W_2) \leq T_{f_s}(W_{1,2}^+),$$

for any $f_s \in \mathcal{F}_{\text{cx, s}}$.

Proof. The idea behind the proof of this lemma is exactly the same idea we used in the proof of Lemma 4.5. First, note that the channels $W_{1,2}^{\pm}$ and $W_{2,1}^{\pm}$ have the same T_{f_s} values. Thus, it would be sufficient to show the first set of inequalities.

As for any realizations δ_1 and δ_2 of the random variables $\Delta_{W_1}(y_1)$ and $\Delta_{W_2}(y_2)$, respectively, $|\delta_1\delta_2| \leq |\delta_1|$ holds, we have

$$f_s(\delta_1\delta_2) \leq f_s(\delta_1),$$

for any $f_s \in \mathcal{F}_{\text{cx, s}}$. Taking expectations of both sides, we get $T_{f_s}(W_{1,2}^-) \leq T_{f_s}(W_1)$.

On the other side, we have

$$\frac{1+\delta_1\delta_2}{2} f\left(\frac{\delta_1+\delta_2}{1+\delta_1\delta_2}\right) + \frac{1-\delta_1\delta_2}{2} f_s\left(\frac{\delta_1-\delta_2}{1-\delta_1\delta_2}\right) \leq f_s\left(\frac{\delta_1+\delta_2}{2} + \frac{\delta_1-\delta_2}{2}\right) = f_s(\delta_1),$$

by Jensen's inequality. Taking expectations, $T_{f_s}(W_1) \leq T_{f_s}(W_{1,2}^+)$ follows. □

Using Definition 6.2, the following corollary follows to the lemma.

Corollary 6.17. *The channels $W_{1,2}^-$, W_1, $W_{1,2}^+$ satisfy the symmetric convex ordering:*

$$\left|\Delta_{W_{1,2}^-}\right| \prec_{\text{icx}} |\Delta_{W_1}| \prec_{\text{icx}} \left|\Delta_{W_{1,2}^+}\right|.$$

The same result holds for the channel W_2.

6.5 Efficient Construction of Polar Codes

In the beginning of this chapter, we acknowledged the difficulty in computing efficiently the exact transition probabilities of the synthetic channels when these have very large output alphabets. In this section, we make a quick look into how, despite this underlying difficulty, the information sets of polar codes can still be efficiently constructed. We interpret, from the perspective of the current chapter, the idea of the approximation algorithm used in [3]:

- The information set construction algorithm starts by computing the exact transition probabilities of the synthetic channels until the size of the output alphabets violate a maximum permissible bound.

- Once this happens, the synthetic channels with large output alphabets are replaced by channels (i) which are 'close' to the original channels, (ii) which have permissible output alphabet sizes, and (iii) whose children synthesized by the sequence of polar transformations still remain 'close' to their exact versions. Thus, one of the key ingredients is to use an approximation inducing an ordering which is preserved under the polar transform. In [3], stochastic degradation is used for that purpose, and it is shown that the algorithm performs well— a further analysis of the algorithm carried out in [37] bounds the maximum approximation loss of the algorithm and shows that the algorithm works with almost linear complexity in the block-length.

  As the symmetric convex ordering is a (weaker) partial ordering also preserved under the polar transform, it can be used as an alternative approximation method for the synthetic channels. Although we have not implemented such an algorithm to evaluate its performance, we claim that similar guarantees can be obtained given the fact that both convex ordering (stochastic degradation) and symmetric convex ordering are induced via the fusion (merging) of the outputs.

- Let us also discuss the role of Theorem 6.15: the exact and approximate computations can be abandoned once the gap between the information sets of the two specific 'squeezer' BECs defined in the theorem's statement is sufficiently small. In that case, the algorithm can proceed by using the BEC recursion for some channel parameters such as the Bhattacharyya distance, and eventually terminate.

In addition, the results of this chapter shows that the idea of the algorithm proposed in [3] can be applied to approximate the channels synthesized by the more general polar transform $\langle W_1, W_2 \rangle^{\pm}$. However, the readers should wait until Chapter 9 to see that such a construction combining non-identical channels does still make sense, and whence the algorithm remains useful.

Appendix

6.A Lemma 6.18

Lemma 6.18. *Let $f(\delta)$ be a convex and symmetric function in $\delta \in [-1, 1]$. Then, the function defined in* (6.6) *is also a convex and symmetric function.*

Proof. For simplicity, we first define

$$f_1(\delta_1, \delta_2) = (1 + \delta_1\delta_2) f\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right),$$
$$f_2(\delta_1, \delta_2) = (1 - \delta_1\delta_2) f\left(\frac{\delta_1 - \delta_2}{1 - \delta_1\delta_2}\right),$$

for $\delta_1, \delta_2 \in [-1, 1]$. Hence, (6.6) equals to $f^+(\delta_1, \delta_2) = \frac{1}{2} f_1(\delta_1, \delta_2) + \frac{1}{2} f_2(\delta_1, \delta_2)$. As $f^+(\delta_1, \delta_2) = f^+(\delta_2, \delta_1)$, it is sufficient to prove the lemma for one of the variables. One can easily prove that the function is symmetric in $\delta_1 \in [-1, 1]$, i.e, $f^+(\delta_1, \delta_2) = f^+(-\delta_1, \delta_2)$ by using the symmetry of the function $f(\delta)$ in $\delta \in [-1, 1]$.

We will prove the rest of the lemma for smooth functions f. As such functions are dense, this is without loss of generality. Let f' and f'' denote the first and the second derivatives of $f(\delta)$ with respect to the variable δ, respectively. Then, we get

$$\frac{\partial}{\partial \delta_1} f_1(\delta_1, \delta_2) = \delta_2 f\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right) + \frac{1 - {\delta_2}^2}{1 + \delta_1\delta_2} f'\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right),$$

$$\begin{aligned}\frac{\partial^2}{\partial {\delta_1}^2} f_1(\delta_1, \delta_2) =& \delta_2 f'\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right) \frac{1 - {\delta_2}^2}{(1 + \delta_1\delta_2)^2} - \delta_2 \frac{1 - {\delta_2}^2}{(1 + \delta_1\delta_2)^2} f'\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right) \\ &+ \frac{1 - {\delta_2}^2}{1 + \delta_1\delta_2} f''\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right) \frac{1 - {\delta_2}^2}{(1 + \delta_1\delta_2)^2} \\ =& \frac{(1 - {\delta_2}^2)^2}{(1 + \delta_1\delta_2)^3} f''\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right).\end{aligned}$$

Similarly, we get

$$\frac{\partial}{\partial \delta_1} f_2(\delta_1, \delta_2) = -\delta_2 f\left(\frac{\delta_1 - \delta_2}{1 - \delta_1\delta_2}\right) + \frac{1 - {\delta_2}^2}{1 - \delta_1\delta_2} f'\left(\frac{\delta_1 - \delta_2}{1 - \delta_1\delta_2}\right),$$

$$\frac{\partial^2}{\partial {\delta_1}^2} f_2(\delta_1, \delta_2) = \frac{(1 - {\delta_2}^2)^2}{(1 - \delta_1\delta_2)^3} f''\left(\frac{\delta_1 - \delta_2}{1 - \delta_1\delta_2}\right).$$

Summing these we obtain

$$\begin{aligned}\frac{\partial^2}{\partial {\delta_1}^2} f^+(\delta_1, \delta_2) = \frac{1}{2} \frac{(1 - {\delta_2}^2)^2}{(1 + \delta_1\delta_2)^3} f''\left(\frac{\delta_1 + \delta_2}{1 + \delta_1\delta_2}\right) \\ + \frac{1}{2} \frac{(1 - {\delta_2}^2)^2}{(1 - \delta_1\delta_2)^3} f''\left(\frac{\delta_1 - \delta_2}{1 - \delta_1\delta_2}\right) \geq 0,\end{aligned}$$

where the sign of $f^+(\delta_1, \delta_2)$ can be deduced from the convexity of the function $f(\delta)$ in $\delta \in [-1, 1]$. This proves that $f^+(\delta_1, \delta_2)$ is convex in $\delta_1 \in [-1, 1]$ and completes the proof. □

Chapter 7

The Mismatched Capacity of Polar Codes

Consider a single use of a B-DMC $W\colon \mathbb{F}_2 \to \mathcal{Y}$ to transmit a single bit. Assume that maximum likelihood decoding with respect to a possibly mismatched B-DMC $V\colon \mathbb{F}_2 \to \mathcal{Y}$ is used as a decoding metric. Given that the symbol 0 is transmitted, the error probability resulting from such a transmission is given by

$$P_{\text{e, ML}}(W, V|0) := \sum_{\substack{y\colon \\ L_V(y)>1}} W(y|0) + \frac{1}{2} \sum_{\substack{y\colon \\ L_V(y)=1}} W(y|0),$$

where $L_V(y) = V(y|1)/V(y|0)$. Similarly, the error probability given that the symbol 1 is transmitted is given by

$$P_{\text{e, ML}}(W, V|1) := \sum_{\substack{y\colon \\ L_V(y)<1}} W(y|1) + \frac{1}{2} \sum_{\substack{y\colon \\ L_V(y)=1}} W(y|1).$$

Assuming both inputs are equally likely, the average error probability satisfies

$$P_{\text{e, ML}}(W, V) := \frac{1}{2} P_{\text{e, ML}}(W, V|0) + \frac{1}{2} P_{\text{e, ML}}(W, V|1). \tag{7.1}$$

Suppose now once again (as we did in the beginning of Chapter 5) that the channel is almost perfect with $I(W) > 1 - \gamma$, where $\gamma > 0$ is small. As there is a mismatch in the decoding procedure, this time $P_{\text{e, ML}}(W, V)$ is not necessarily small even when $I(V) > 1 - \gamma$. For instance, take two BSCs of crossover probabilities $\epsilon = \gamma$ and $1 - \epsilon$. Although both channels have high symmetric capacities, an error will occur with high probability in case of mismatched decoding. For this reason, when there is a mismatch in the decoding metric, we cannot always ensure reliable communication by simply transmitting data uncoded over the channel as in the matched case. Nor

can we claim right away that: "low error probability and low complexity polar decoding is also possible even when the good synthetic channels are not going to be decoded with their own channel metrics during successive cancellation decoding". Thus, it is not obvious at all what could be achieved by polarizing the true channel W and constructing a polar code for the channel when the polar decoder is not to compute its decision functions using the true channel law.

In this chapter, we will study the performance of mismatched polar decoders. The *mismatched polar successive cancellation decoder* can be described as the matched decoder using a chain of estimators

$$\hat{u}_i = \begin{cases} u_i, & \text{if } \; i \in \mathcal{A}_N^c \\ f_M^{(i)}(y_1^N, \hat{u}_1^{i-1}), & \text{if } \; i \in \mathcal{A}_N \end{cases},$$

for $i = 1, \ldots, N$, where the decision functions $f_M^{(i)}(y_1^N, \hat{u}_1^{i-1})$ still apply the maximum likelihood rule but with respect to mismatched channels $V_N^{(i)}$ synthesized by polarizing a B-DMC V different than the true communication channel W. Let $C_P(W, V)$ denote the transmission capacity of the channel W when the outputs are decoded with a mismatched polar decoder designed with respect to another channel V. We will call $C_P(W, V)$ as the *polar mismatched capacity*. The primary objective of this chapter is to study the effects of a decoding mismatch on the transmission capacity of polar coding. As hinted before in the final section of Chapter 5, we will see that things do not generalize trivially when we deviate from the classical scenario.

7.1 Reliable Communication with a Given Decision Rule

In various communication scenarios, we encounter sub-optimal decoders due to partial/missing channel information or practical implementation constraints (complexity, feasibility requirements). To give an example of an obstacle on the way of optimal decoding, we can consider the case where a high signal to noise ratio channel is used with a large constellation with points indexed by k-bit symbols $s(0...0), ..., s(1...1)$, and the receiver is interested in recovering the 1st of these k bits. Then, the true likelihood ratio requires the computation of the sums

$$\sum_{b_2,\ldots,b_k} W(y|s(0, b_2, ..., b_k)) \quad \text{and} \quad \sum_{b_2,\ldots,b_k} W(y|s(1, b_2, ..., b_k)),$$

each containing an exponential number of terms (in k). The receiver hardware may not permit such computations, and consequently the decoder designer may be forced to use a simpler metric $V(y|1)/V(y|0)$ which approximates the true one. In

such cases even when the receiver is informed of the true channel W, the decoding operation proceeds on the basis of a mismatched channel V. Regardless of the nature of the obstacle, sub-optimal decoders might perform worse than the optimal decoders which minimize the average decoding error probability and possibly result in capacity loss. Modeling such sub-optimal scenarios via *'reliable communication with a given decision rule'* and establishing coding theorems for them allows one to assess the extent of any loss.

To allow their study within a unified framework, decoders can be categorized based on generic definitions of the decision functions. *Alpha (α) decoders* are such a class where the decision rules are based on the joint type of the codewords and the received sequence. Given a codeword set $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$, where $\mathbf{x}_m \in \mathcal{X}^N$ for all $m = 1, \ldots, M$, an α decoder assigns a received sequence $\mathbf{y}$ the message $i = 1, \ldots, M$ if and only if $\alpha(\mathbf{x}_i, \mathbf{y}) < \alpha(\mathbf{x}_j, \mathbf{y})$, for $\forall j \neq i$. If there is no such i the decoder declares an erasure. The derivation of the error exponent is based on a useful tool called the method of types [13]. For the single point-to-point channel $W : \mathcal{X} \to \mathcal{Y}$, the *random coding exponent of α decoders* used with constant composition codes of type $\widehat{P}$ and rate R, is given by Csiszár and Narayan [12] as

$$E_{r,\alpha}(R, \widehat{P}, W) := \inf_{\substack{\widehat{P}_1 \in \mathcal{P}^N(\widehat{P}, W) \\ \widehat{P}_2 \in \mathcal{P}^N(\widehat{P}, \widehat{P}_1) \\ \alpha(\widehat{P}\widehat{P}_2) \leq \alpha(\widehat{P}\widehat{P}_1)}} \left(\mathrm{Div}(\widehat{P}_1 \| W | \widehat{P}) + |I(\widehat{P}; \widehat{P}_2) - R|^\dagger \right), \tag{7.2}$$

where $\widehat{P}_1(y|x)$, $\widehat{P}_2(y|x)$ are noise compositions, $\mathrm{Div}(\widehat{P}_1 \| W | \widehat{P})$ is the conditional divergence, and

$$\mathcal{P}^N(\widehat{P}, W) := \{\widehat{P}_1(y|x) : \ \widehat{P}\widehat{P}_1(x, y) \in \mathcal{P}^N(\mathcal{X}, \mathcal{Y}), \widehat{P}\widehat{P}_1(y) = \widehat{P}W(y)\}.$$

The input type $\widehat{P}$ and the channel transition probability W induce a joint distribution $\widehat{P}W(x, y) = \widehat{P}(x)W(y|x)$, the corresponding output marginals are given by $\widehat{P}W(y)$, and $\mathcal{P}^N(\mathcal{X}, \mathcal{Y})$ denotes the set of joint N-types. See Definition 8.20 for the definition of an N-type. We refer to [11] and [13] for more details on these notions. A universal decoding rule for DMCs, initially proposed by Goppa [38], is given by the MMI decoder belonging to the class of α decoders. The MMI decoder reduces $E_{r,\alpha}(R, \hat{P}, W)$ to [11], [13]

$$\inf_{\widehat{P}_1 \in \mathcal{P}^N(\widehat{P}, W)} \left(\mathrm{Div}(\widehat{P}_1 \| W | \widehat{P}) + |I(\widehat{P}; \widehat{P}_1) - R|^\dagger \right). \tag{7.3}$$

As the exponent is positive if and only if $R < I(\widehat{P}; W)$, universally attainable transmission rates are obtained for any DMC with finite input and output alphabets. The problem of reliable communication under channel uncertainty is already solved for DMCs, but only in theory. Typicality decoders are too complex to be implementable.

Further downstream, Csiszár and Narayan [12] studied the performance of a more restricted class of decoders using additive decision rules. A d-decoder is an α-decoder whose decision function is computed using the additive extension of a single letter metric $d(x, y)$. Even though the MMI decoder is no longer treated within this class, d-decoders still provide a broad enough framework to allow the study of rules such as the optimal or mismatched maximum likelihood decoding rules, or of difficult information-theoretical problems such as the Shannon capacity of a graph and the zero-error capacity of a DMC [39]. The transmission capacity of the channel W when decoded with an additive metric d is denoted by $C_d(W)$. When the metric d corresponds to the maximum likelihood decoder with respect to a channel V, the decoder is called a *mismatched maximum likelihood decoder*. We denote the corresponding *mismatched capacity* by $C(W, V)$. No closed form single letter expression is known for $C(W, V)$ or $C_d(W)$. Single-letter lower bounds have been derived, but no converse for any of the lower bounds exists, except for some special cases. Binary input binary output channels are such a case where $C_d(W) = C(W)$ or 0 depending on whether or not the mismatch metric is in 'harmony' with the channel behavior [12]. Another exception is the class of binary input discrete memoryless channels. Balakirsky [40], [41] derived a converse and gave a computable expression for $C_d(W)$ when W is a B-DMC.

Successive cancellation decoders sitting at the center of this chapter can be considered as another large decoder family based on successive cancellation decoding procedures. Offering a quite different decoding paradigm than additive decoders, a successive cancellation decoder will decode the received output y_1^N in N stages using a chain of estimators from $i = 1, \ldots, N$ each possibly depending on the previous ones. The estimators $\hat{u}_i$ can base their decisions on arbitrary single letter metrics of the form $d_i(u_i, y_1^N \hat{u}_1^{i-1})$. The polar successive cancellation decoder, however, owes its fame not only for yielding polar coding theorems proving the 'symmetric capacity achievingness' of polar codes for a large class of channels, but also for inheriting the low complexity structure of the recursive code construction process. On the road to low complexity universal decoding, we will see in this and the next chapter that similar conclusions apply for the mismatched polar successive cancellation decoder in many cases of interest: While extending the theory of channel polarization and polar codes to mismatched processes, the mismatched polar decoder preserves the low complexity structure of the original polar decoder.

What's Coming, Doc?

The main purpose of this chapter is to answer the following questions:

- Is there an expression for $C_P(W, V)$? Is it computable? Are there single-letter lower bounds for $C_P(W, V)$?

To answer these questions, we introduce two mismatched channel parameters:

$$I(W,V) := \sum_{x\in\mathbb{F}_2}\sum_{y\in\mathcal{Y}} \frac{1}{2} W(y|x) \log \frac{V(y|x)}{\frac{1}{2}V(y|0)+\frac{1}{2}V(y|1)}, \tag{7.4}$$

and

$$D(W,V) := \sum_{y\in\mathcal{Y}} q_W(y)\sqrt{|\Delta_V(y)|}, \tag{7.5}$$

where $q_W(y)$ and $\Delta_V(y)$ are defined in (2.14) and (2.15), respectively. We will report the following results related to these two parameters:

- Proposition 7.7 will show that the quantity $I(W,V)$ turns out to be, as the symmetric capacity of the channel, a conserved quantity under the polar transform. As a consequence, the process $I_n(W,V) := I(W_n,V_n)$ associated to both polarization processes of the channels W and V will be, as its matched counterpart, a martingale process. However, identifying the convergence points of this process will be a challenging task. Although we will conjecture that, whenever $I(W,V) > -\infty$, $I_\infty(W,V) \in \{0,1\}$ a.s., we will not only rely on this conjecture to characterize $C_P(W,V)$.
- Instead, we will couple the analysis of $I_n(W,V)$ with the analysis of the process $D_n(W,V) := D(W_n,V_n)$. Proposition 7.11 will show that $D_n(W,V)$ is a bounded supermartingale converging a.s. to a $\{0,1\}$ valued random variable. The coupling will also reveal that the convergence points of $D_\infty(W,V)$ have the following operational meaning: $D_\infty = 0$ a.s. refers to completely noisy synthetic channels $W_N^{(i)}$ over which the genie aided mismatched maximum likelihood decoding with respect to $V_N^{(i)}$ will fail with high probability, and $D_\infty = 1$ a.s. refers to almost perfect synthetic channels where uncoded transmission will result in a vanishing error probability $P_{\text{e, ML}}(W_N^{(i)}, W_N^{(i)})$.
- Furthermore, we will show that the polarization of mismatched processes happens sufficiently fast: Theorem 7.4 will argue that the same rate of convergence result derived in [30] for matched processes holds.

These results will constitute the basis of the achievability and coding theorems with mismatched polar decoding: Theorem 7.18, Theorem 7.19, and Theorem 7.20 will state the following results:

(i) $C_P(W,V)$ equals to the fraction of indices for which $D_\infty = 1$ a.s.,

(ii) $I(W,V)$ is a single letter lower bound to $C_P(W,V)$,

(ii) This lower bound naturally generates a sequence of tighter lower bounds that we conjecture to be asymptotically converging to $C_P(W,V)$,

(iii) The block decoding error probability of the mismatched polar decoder is in the order of the square root of the block-length.

As a consequence, whenever $I(W,V) > 0$, or any of the lower bounds are positive, strictly positive communication rates can be achieved with mismatched polar decoding. Following this, we will discuss how the definition of the information set of a polar code can be adapted to the mismatched polar decoding scenario.

In the final section, we will compare $C_P(W,V)$ with the classical mismatched capacity $C(W,V)$. For the sake of the comparison, we will study the evolution under the polar transform of '*Balakirsky's converse*' for B-DMCs: We will show that $C(W,V)$ can be either created or lost, and in general it is not preserved after applying the polar transform to W and V. Hence, we will not be able to conclude any 'martingale conservation and convergence laws' for the process $C_n(W,V) := C(W_n, V_n)$, as Arıkan did for $I_n(W)$ by using the one-step conservation property of $I(W)$[1]. Instead, we will conclude that no general order between the polar mismatched capacity $C_P(W,V)$ and the classical mismatched capacity $C(W,V)$ can be formulated: there are examples for which $C_P(W,V) > C(W,V)$ and also examples where we expect the reverse inequality to hold. Motivated by this, we will propose polarization as a novel architecture to boost $C(W,V)$ since communication rates higher than $C(W,V)$ can be achieved in some cases by integrating the polarization architecture of Arıkan into the classical mismatched communication scenarios.

7.2 Mismatched Conservation & Convergence

Let us consider again a single use of the B-DMC $W\colon \mathbb{F}_2 \to \mathcal{Y}$ to transmit a 0. Given that the symbol 0 is transmitted, one can derive the following upper bound to the decoding error probability resulting from such a transmission:

$$\begin{aligned} P_{\text{e, ML}}(W,V|0) &= \sum_{\substack{y\colon \\ L_V(y)\geq 1}} W(y|0) + \frac{1}{2} \sum_{\substack{y\colon \\ L_V(y)=1}} W(y|0) \\ &\leq \sum_{\substack{y\colon \\ L_V(y)\geq 1}} W(y|0) \left(\log\left(1+L_V(y)\right) - \log 2 + \log 2\right) \\ &\leq \sum_{y\in\mathcal{Y}} W(y|0) \left(1 - \log\left(\frac{2}{1+L_V(y)}\right)\right) = 1 - I(W,V|0), \end{aligned}$$

where

$$I(W,V|0) = \sum_{y\in\mathcal{Y}} W(y|0) \log\left(\frac{2V(y|0)}{V(y|0)+V(y|1)}\right).$$

[1]This does not mean that the process $C_n(W,V)$ does not converge. In fact, the convergence of $D_n(W,V)$ can be used to infer the convergence of $C_n(W,V)$.

Similarly, the error probability given that the symbol 1 is transmitted is bounded by

$$P_{\text{e, ML}}(W, V|1) \leq 1 - I(W, V|1),$$

where

$$I(W, V|1) = \sum_{y \in \mathcal{Y}} W(y|1) \log \left(\frac{2V(y|1)}{V(y|0) + V(y|1)} \right).$$

Assuming both inputs are equally likely, the average error probability satisfies

$$P_{\text{e, ML}}(W, V) \leq 1 - I(W, V), \tag{7.6}$$

with $I(W, V) = \left(I(W, V|0) + I(W, V|1)\right)/2$. Without any surprise, the matched quantity $I(W) := I(W, W)$, for $V = W$, gives the symmetric capacity of the channel. We shall soon study the evolution of the mismatched process $I_n(W, V)$ associated to the channel polarization processes W_n and V_n. Prior to that, let us make a couple of observations related to $I(W, V)$.

Proposition 7.1. $I(W) \geq I(W, V)$.

Proof. We show that $I(W) - I(W, V) \geq 0$. Let $W_0 := W(\cdot|0)$ and $W_1 := W(\cdot|1)$. The difference can be written as:

$$\frac{1}{2} \operatorname{Div}\left(W_0 \| V_0\right) + \frac{1}{2} \operatorname{Div}\left(W_1 \| V_1\right) - \operatorname{Div}\left(\frac{W_0 + W_1}{2} \middle\| \frac{V_0 + V_1}{2} \right) \geq 0,$$

where the inequality follows from the convexity of the Kullback–Leibler divergence $\operatorname{Div}(P_1 \| P_2)$ in the pair (P_1, P_2). □

Thus, the bound $I(W, V) \leq 1$ necessarily holds. Notice, however, that $I(W, V)$ can take on negative values and is in fact unbounded from below.

It is also worth mentioning that the definition of the mismatched parameter $I(W, V)$ matches the *generalized mutual information* definition under the uniform input distribution in Fischer [42]. For completeness, we included in Appendix 7.A the derivation of $I(W, V)$ based on Gallager's error exponent derivation technique [5][2]. It is shown in [42] that $C(W, V) \geq I(W, V)$ holds, see also [12, Remarks i)]. Thus, $|I(W, V)|^{\dagger}$ is an achievable rate with mismatched maximum likelihood decoding. As a final observation, we show that $I(W, V)$ can be expressed in terms of our favorite channel parameter $\Delta_V(y)$ defined in (2.15).

[2]Kaplan and Shamai [43] derived a more general version of Fischer's generalized mutual information by using Gallager's technique. See also Merhav et. al. [44] for the definition of this more general expression.

Lemma 7.2.

$$I(W,V) = \sum_{y\in\mathcal{Y}} \frac{1}{2} W(y|0) \log\left(1 + \Delta_V(y)\right) + \frac{1}{2} W(y|1) \log\left(1 - \Delta_V(y)\right). \quad (7.7)$$

Proof. The proof follows by noting that

$$1 + \Delta_V(y) = \frac{2V(y|0)}{V(y|0) + V(y|1)} \quad \text{and} \quad 1 - \Delta_V(y) = \frac{2V(y|1)}{V(y|0) + V(y|1)},$$

and using these in (7.4). □

Along with $I(W,V)$, we will also analyze the following mismatched channel parameter $D(W,V)$. Note that when $W = V$, a few simple manipulations show that

$$D(W) = \sum_{y\in\mathcal{Y}} q_W(y) \sqrt{|W(0|y)^2 - W(1|y)^2|} = \sum_{y\in\mathcal{Y}} q_W(y) \sqrt{|W(0|y) - W(1|y)|},$$

where $W(.|y)$ denotes the posterior probabilities of the channel inputs given its output. Thus, $D(W)$ is somewhat similar to $T(W) = \sum_{y\in\mathcal{Y}} q_W(y)|W(0|y) - W(1|y)|$, the 'variational distance' we introduced in Chapter 5. In Section 5.3, we did attempt to extend the convergence result of Proposition 5.2 about $T_n(W)$ to the mismatched version $T_n(W,V)$ defined in (5.17). However, we closed the chapter with empty hands. We will soon find out that the 'convergence and conservation laws' we hoped to write through the polarization analysis of the process $T_n(W,V)$ will follow from the analysis of $I_n(W,V)$ and $D_n(W,V)$. Furthermore, the analysis will help us prove the following two theorems.

Theorem 7.3. *Let W and V be two B-DMCs. Then,*

$$\mathbb{P}[D_\infty(W,V) \to 1] \geq I(W,V). \quad (7.8)$$

Theorem 7.4 (Rate of Convergence)**.** *For any $\beta < 1/2$,*

$$\lim_{n\to\infty} \mathbb{P}[D_n(W,V) < 2^{-2^{n\beta}}] \leq 1 - I(W,V). \quad (7.9)$$

The main achievability and coding theorems with mismatched polar decoding, which we will state in Section 7.3, will rely heavily on the above technical results. Their proofs are given in Subsection 7.2.4. Till then, we have a long way to go.

7.2.1 A 'Conservation Law'

We begin by applying the useful Δ trick to derive suitable expressions for $I(W^-, V^-)$ and $I(W^+, V^+)$ in the next two lemmas.

Lemma 7.5. *Let W and V be two B-DMCs. Then,*

$$\begin{aligned} I(W^-, V^-) = &\frac{1}{4}\sum_{y_1 y_2} W(y_1|0)W(y_2|0)\log\left(1+\Delta_V(y_1)\Delta_V(y_2)\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|0)W(y_2|1)\log\left(1-\Delta_V(y_1)\Delta_V(y_2)\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|1)W(y_2|0)\log\left(1-\Delta_V(y_1)\Delta_V(y_2)\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|1)W(y_2|1)\log\left(1+\Delta_V(y_1)\Delta_V(y_2)\right). \end{aligned} \tag{7.10}$$

Proof. Using the definition of the minus transformation in (4.1), the proof follows upon observing

$$\begin{aligned} 1+\Delta_V(y_1)\Delta_V(y_2) &= \frac{2V(y_1|0)V(y_2|0)+2V(y_1|1)V(y_2|1)}{\sum_u V(y_1|u)V(y_2|u\oplus 1)+V(y_1|u\oplus 1)V(y_2|u)}, \\ 1-\Delta_V(y_1)\Delta_V(y_2) &= \frac{2V(y_1|0)V(y_2|1)+2V(y_1|1)V(y_2|0)}{\sum_u V(y_1|u)V(y_2|u\oplus 1)+V(y_1|u\oplus 1)V(y_2|u)}. \end{aligned}$$ □

Lemma 7.6. *Let W and V be two B-DMCs. Then,*

$$\begin{aligned} I(W^+, V^+) = &\frac{1}{4}\sum_{y_1 y_2} W(y_1|0)W(y_2|0)\log\left(1+\frac{\Delta_V(y_1)+\Delta_V(y_2)}{1+\Delta_V(y_1)\Delta_V(y_2)}\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|0)W(y_2|1)\log\left(1+\frac{\Delta_V(y_1)-\Delta_V(y_2)}{1-\Delta_V(y_1)\Delta_V(y_2)}\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|1)W(y_2|0)\log\left(1-\frac{\Delta_V(y_1)-\Delta_V(y_2)}{1-\Delta_V(y_1)\Delta_V(y_2)}\right) \\ &+ \frac{1}{4}\sum_{y_1 y_2} W(y_1|1)W(y_2|1)\log\left(1-\frac{\Delta_V(y_1)+\Delta_V(y_2)}{1+\Delta_V(y_1)\Delta_V(y_2)}\right). \end{aligned} \tag{7.11}$$

Proof. Using the definition of the plus transformation in (4.2), the proof follows

upon observing

$$1+\frac{\Delta_V(y_1)+\Delta_V(y_2)}{1+\Delta_V(y_1)\Delta_V(y_2)}=\frac{2V(y_1|0)V(y_2|0)}{\sum_u V(y_1|u)V(y_2|u)},$$
$$1-\frac{\Delta_V(y_1)+\Delta_V(y_2)}{1+\Delta_V(y_1)\Delta_V(y_2)}=\frac{2V(y_1|1)V(y_2|1)}{\sum_u V(y_1|u)V(y_2|u)},$$
$$1+\frac{\Delta_V(y_1)-\Delta_V(y_2)}{1-\Delta_V(y_1)\Delta_V(y_2)}=\frac{2V(y_1|0)V(y_2|1)}{\sum_u V(y_1|u)V(y_2|u\oplus 1)},$$
$$1-\frac{\Delta_V(y_1)-\Delta_V(y_2)}{1-\Delta_V(y_1)\Delta_V(y_2)}=\frac{2V(y_1|1)V(y_2|0)}{\sum_u V(y_1|u)V(y_2|u\oplus 1)}.$$ □

Now, using the expressions (7.7), (7.10), and (7.11) derived in Lemmas 7.2, 7.5, and 7.6, respectively, we can readily see that the polar transform conserves the quantity $I(W,V)$.

Proposition 7.7 (Conservation Property)**.** *For any two B-DMCs W and V,*

$$I(W^-,V^-)+I(W^+,V^+)=2I(W,V).$$

This result lays the foundation of the subsequent martingale argument.

Proposition 7.8 (Conservation Law)**.** *The process $I_n(W,V)$ is a bounded martingale such that $I_n(W,V)\leq 1$, for all $n\geq 0$. Furthermore, the process converges a.s. to a limiting random variable $I_\infty(W,V)$ such that $\mathbb{E}[I_\infty(W,V)]\geq I(W,V)$ holds.*

Proof. The martingale property follows by Proposition 7.7 and the boundedness by Proposition 7.1. The remaining claims hold by general results on bounded martingales upon noticing that $1-I(W,V)$ is a non-negative supermartingale, see [27, 11.7]. □

Although the above proposition tells us $I_n(W,V)$ will converge a.s., we do not know to which values the convergence will be. In Subsection 7.2.3, we will give our conjecture regarding this point.

7.2.2 A 'Convergence Law'

Let us put aside $I(W,V)$ and consider back the operation of the polar decoder in terms of the channel parameter Δ, our savior, instead of the likelihood ratio.

The genie-aided successive cancellation decoder can as well start by computing N numbers $\Delta(y_1)\dots\Delta(y_N)$ taking values in the interval $[-1,1]$, and then transform these numbers using a Fast Fourier Transform like circuitry, see [2, Figure 10], into a final set of N numbers corresponding to the Δ parameters of the synthetic channels. Due to the recursive nature of the operations, the decoder operates each time instant on two numbers and performs either the minus operation on the two numbers, say Δ_1 and Δ_2, by computing

$$\Delta^{-} = \Delta_1\Delta_2,$$

or one of the plus operations by computing

$$\Delta^{+_1} = \frac{\Delta_1+\Delta_2}{1+\Delta_1\Delta_2}, \quad \Delta^{+_2} = \frac{\Delta_1-\Delta_2}{1-\Delta_1\Delta_2}.$$

While the minus operation is a straightforward multiplication of two numbers, both plus operations look more complex. The following lemma will help us to obtain a simple upper bound on these seemingly complex operations.

Lemma 7.9. *For a, b in the interval $[0,1]$,*

$$\sqrt{\frac{|a^2-b^2|}{1-a^2b^2}} \le \sqrt{\frac{a^2+b^2}{1+a^2b^2}} \le a+b-ab.$$

Proof. For the first inequality, we can assume without loss of generality that $x = a^2 \ge b^2 = y$, and we only need to check

$$\frac{x-y}{1-xy} \le \frac{x+y}{1+xy},$$

for $x, y \in [0,1]$, or equivalently, $(x-y)(1+xy) \le (x+y)(1-xy)$. But this last simplifies to $x^2y \le y$, which clearly holds.

For the second inequality, squaring both sides and multiplying by $(1+a^2b^2)$ we see that the inequality is equivalent to

$$(a+b-ab)^2(1+a^2b^2) - a^2 - b^2 \ge 0.$$

The left hand side factorizes as $a(1-a)b(1-b)\big(2-ab(1+a+b-ab)\big)$. Thus the lemma will be proved once we show that

$$t(1+s-t) \le 2,$$

where $s = a+b$ and $t = ab$. Note that $0 \le s \le 2$ and $0 \le t \le 1$. Thus, $t(1+s-t) \le t(3-t) \le 2$. □

Now, if we take $a = \sqrt{|\Delta_1|}$ and $b = \sqrt{|\Delta_2|}$ in Lemma 7.9, we get

$$\sqrt{\left|\frac{\Delta_1 + \Delta_2}{1 + \Delta_1\Delta_2}\right|} \leq \sqrt{|\Delta_1|} + \sqrt{|\Delta_2|} - \sqrt{|\Delta_1|}\sqrt{|\Delta_2|}, \tag{7.12}$$

and

$$\sqrt{\left|\frac{\Delta_1 - \Delta_2}{1 - \Delta_1\Delta_2}\right|} \leq \sqrt{|\Delta_1|} + \sqrt{|\Delta_2|} - \sqrt{|\Delta_1|}\sqrt{|\Delta_2|}. \tag{7.13}$$

This brings us to the next lemma.

Lemma 7.10. *Suppose Δ_1, Δ_2 are independent $[-1, 1]$ valued random variables with $\mathbb{E}\left[\sqrt{|\Delta_i|}\right] = \mu_i$. Then*

$$\mathbb{E}\left[\sqrt{\left|\frac{\Delta_1 + \Delta_2}{1 + \Delta_1\Delta_2}\right|}\right] \leq \mu_1 + \mu_2 - \mu_1\mu_2,$$

and

$$\mathbb{E}\left[\sqrt{\left|\frac{\Delta_1 - \Delta_2}{1 - \Delta_1\Delta_2}\right|}\right] \leq \mu_1 + \mu_2 - \mu_1\mu_2.$$

Proof. The lemma follows by taking the expectations of both sides of (7.12) and (7.13) and noting the independence of Δ_1 and Δ_2. □

Those familiar with the polar coding literature could by now foresee what we are about to state. In the following proposition, we state the convergence properties of the process $D_n(W, V)$.

Proposition 7.11 (Loss & Convergence Law)**.** *Let W and V be B-DMCs. Then, the process $D_n(W, V)$ is a bounded supermartingale which converges a.s. to a $\{0, 1\}$ valued limiting random variable $D_\infty(W, V)$.*

Proof. We will prove that

$$D(W^+, V^+) + D(W^-, V^-) \leq 2D(W, V)$$

holds after a single step. The general result showing the process is a supermartingale will follow by the recursive structure. From (6.4) and (6.5), we know that

$$\Delta_{V^-}(Y_1Y_2) = \Delta_V(Y_1)\Delta_V(Y_2), \tag{7.14}$$

$$\Delta_{V^+}(Y_1Y_2U_1) = \frac{\Delta_V(Y_1) + (-1)^{U_1}\Delta_V(Y_2)}{1 + (-1)^{U_1}\Delta_V(Y_1)\Delta_V(Y_2)}, \tag{7.15}$$

where $Y_1 \sim q_W(y_1)$ and $Y_2 \sim q_W(y_2)$. So, we get

$$D(W^-, V^-) = \mathbb{E}\left[\sqrt{|\Delta_{V^-}|}\right] = \mathbb{E}\left[\sqrt{|\Delta_V|}\right]^2 = D(W, V)^2,$$

and

$$\begin{aligned} D(W^+, V^+) &= \mathbb{E}\left[\sqrt{|\Delta_{V^+}|}\right] \\ &\leq 2\mathbb{E}\left[\sqrt{|\Delta_V|}\right] - \mathbb{E}\left[\sqrt{|\Delta_V|}\right]^2 \\ &= 2D(W, V) - D(W, V)^2, \end{aligned}$$

where the inequality holds by Lemma 7.10. As $\sqrt{|\Delta_{V_n}|} \in [0, 1]$, this proves that the process is a bounded supermartingale, and by general results on bounded martingales, the process converges a.s. to a limiting random variable $D_\infty(W, V)$ under the distribution q_{W_n}. One can prove the convergence is to $\{0, 1\}$ using the squaring property of the minus transformation in a similar fashion as in the proof of [2, Proposition 9] of the convergence points of the Bhattacharyya process of the synthetic channels associated to the polar transformations. □

We now deduce the following result on the distribution of Δ_{V_∞}.

Corollary 7.12. *The distribution of Δ_{V_∞} measured under any (matched or mismatched) output distribution is supported at $\{-1, 0, 1\}$.*

Proof. The corollary follows by Proposition 7.11 as the convergence points of $D_n(W, V)$ are the extreme points of the interval from which the process $\sqrt{|\Delta_{V_n}|}$ takes values. □

7.2.3 Detective, Smells Like a Mystery

So far, we did analyze $I_n(W, V)$ and $D_n(W, V)$ independently. In the following lemma, we first show how these two channel parameters are coupled in general.

Lemma 7.13. $I(W, V) \leq \dfrac{1}{\ln 2} D(W, V)$.

Proof. The result follows from the inequalities

$$\log(1 + \Delta) \leq \sqrt{|\Delta|}/\ln 2 \quad \text{and} \quad \log(1 - \Delta) \leq \sqrt{|\Delta|}/\ln 2, \quad \text{for } \Delta \in [-1, 1].$$

□

Next, we state a conjecture about the convergence points of $I_n(W, V)$ and about how the recursive application of the polar transform affects the coupling between $I_n(W, V)$ and $D_n(W, V)$ in the limit.

Conjecture 7.14. Let W and V be two B-DMCs such that $I(W,V) > -\infty$. Then, $I_\infty(W,V)$ is a.s. $\{0,1\}$ valued with

$$\mathbb{P}[I_\infty(W,V) \to 1] = \mathbb{P}[D_\infty(W,V) \to 1]. \tag{7.16}$$

To simplify the analysis, we introduce the next definition.

Definition 7.15. When for some permutation π on the output alphabet $\mathcal{Y}$ satisfying $\pi = \pi^{-1}$, we simultaneously have $W(y|0) = W(\pi(y)|1)$ and $V(y|0) = V(\pi(y)|1)$ for all output letters y, we say that the channels are *symmetrized by the same permutation*.

Heuristic Proof. Let us discuss some ideas for proving the conjecture. In trying to prove that the process $I_n(W,V)$ cannot converge to any point in the interval $(0,1)$, a natural strategy is to follow the proof method for the matched case which shows that $I_n(W)$ cannot converge to any point in the interval $(0,1)$. There, two properties are used: (i) $I_n(W)$ is a martingale, and (ii) for any $\gamma > 0$ there is a $\xi > 0$ such that $|I(W^-) - I(W)| < \xi$ implies that $I(W) \notin (\gamma, 1-\gamma)$. The martingale property already holds in the mismatched case, so to prove the conjecture it would be sufficient to show that the second property holds for the mismatched case too, i.e., for any $\gamma > 0$ there is a $\xi > 0$ such that

$$|I(W,V) - I(W^-,V^-)| < \xi \quad \text{implies} \quad I(W,V) \notin (\gamma, 1-\gamma). \tag{7.17}$$

For simplicity, let us assume W and V are both symmetrized by the same permutation. Then using this symmetry in Lemma 7.2, we get

$$I(W,V) = \sum_{y \in \mathcal{Y}} W(y|0) \log\left(1 + \Delta_V(y)\right).$$

In this case, the statement in (7.17) is equivalent to showing that when Δ_1 and Δ_2 are i.i.d. random variables taking values in $[-1,1]$,

$$|E[\log(1+\Delta_1)] - E[\log(1+\Delta_1\Delta_2)]| < \xi \quad \text{implies} \quad E[\log(1+\Delta_1)] \notin (\gamma, 1-\gamma). \tag{7.18}$$

In other words, the pair $(I, I^-) := E\Big[\big(\log(1+\Delta_1), \log(1+\Delta_1\Delta_2)\big)\Big]$ avoids the vicinity of the line segment connecting $(0,0)$ to $(1,1)$ except at the end-points. Furthermore, since we want to show this result when the process $D_n(W,V)$ converges to 1, it is sufficient to prove this for random variables Δ_1 that satisfy $E[|\Delta_1|] > 1 - \xi'$.

Unfortunately, this approach allows us to construct, for any $\gamma > 0$, a Δ random variable giving a counterexample to the statement (7.18). The example is as follows:

- The distribution of Δ is supported on four points $-(1-\delta_1)$, $-(1-\delta_2)$, and $(1-\delta_1)$, $(1-\delta_2)$. The probability masses p, 0, 0 and $1-p$ on these four points are chosen to make $I = 1/2$,

$$p = p(\delta_1, \delta_2) = \frac{\log(2-\delta_2) - (1/2)}{\log(2-\delta_2) - \log \delta_1}.$$

- One then sees that

$$I^- = I^-(\delta_1, \delta_2) = p^2 \log(1 + (1-\delta_1)^2) + (1-p)^2 \log(1 + (1-\delta_2)^2) + 2p(1-p)\log(\delta_1 + \delta_2 - \delta_1\delta_2).$$

- Requiring $I^- = I = 1/2$ constrains (δ_1, δ_2) to lie on a one dimensional curve; furthermore this curve passes through the origin— this can be seen either via a numerical plot, or approximating p and I^- for small values of δ_1, δ_2 and observing that $\delta_1 \approx \delta_2^2/2$ will yield an I^- value of $1/2$. Consequently, it is possible to choose δ_1 and δ_2 to be arbitrarily small and thus make the $D = E[|\Delta|]$ as close to 1 as desired while keeping $I = I^- = 1/2$.

At this point, the reader may be curious as to why we discuss this strategy to proving the conjecture only to show that the strategy fails. The reason is two-fold: (i) the strategy is the most direct way to the proof and so we have eliminated a false lead; (ii) even though this 'one-step' reasoning fails, it is possible that a 'two-step' reasoning, namely imposing the condition that not only $I(W^-, V^-)$ is close to $I(W, V)$ but also $I(W^{--}, V^{--})$ and $I(W^{+-}, V^{+-})$ are also close to $I(W, V)$, and finding the set of possible values of $I(W, V)$ under this condition (together with the $D's$ all being close to 1) may succeed. □

7.2.4 Proofs of Theorem 7.3 and Theorem 7.4

We have come a long way, but at last we are ready to prove the theorems. Thanks to the previous derivations and results, the proofs will be done at one fell swoop!

Proof of Theorem 7.3. By Proposition 7.8, we know that $\mathbb{E}[I_\infty(W, V)] \geq I(W, V)$ holds. By Lemma 7.13, we have $D_\infty(W, V) \geq I_\infty(W, V) \ln 2$. Since by Proposition 7.11, D_∞ is a.s. $\{0, 1\}$ valued, this implies $D_\infty(W, V) \geq I_\infty(W, V)$, and consequently, $\mathbb{P}[D_\infty(W, V) \to 1] = \mathbb{E}[D_\infty(W, V)] \geq \mathbb{E}[I_\infty(W, V)] \geq I(W, V)$. □

Proof of Theorem 7.4. The result follows by [30, Theorem 1] as the conditions (z.1), (z.2), and (z.3) stated in [30] hold taking $I_0 = \mathbb{P}[D_\infty(W, V) = 0] \leq 1 - I(W, V)$ and $Z_n = D_n(W, V)$. □

7.3 Achievability/Coding Theorems with Mismatched Polar Decoding

Let $P_{\mathrm{e}}(W, V, \mathcal{A}_N)$ denote the best achievable block decoding error probability over the ensemble of all possible choices of the set $\mathcal{A}_N^{\mathrm{c}}$ under mismatched polar successive cancellation decoding with respect to the channel V when the true channel is W. The following two propositions derive upper bounds to this error probability.

Proposition 7.16.

$$P_{\mathrm{e}}(W, V, \mathcal{A}_N) \leq \sum_{i \in \mathcal{A}_N} P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}),$$

where $P_{\mathrm{e,\,ML}}(W_N^{(i)}, W_N^{(i)})$ is the error probability of the 'genie-aided' mismatched decoder for the i-th synthetic channel. See (7.1) *for the definition of $P_{\mathrm{e,\,ML}}(W, V)$.*

Proof. Following similar derivations to the analysis carried in [2] for the matched counterpart, we can upper bound this error probability by:

$$\begin{aligned}
P_{\mathrm{e}}(W, V, \mathcal{A}_N) &= \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ \hat{U}_1^{i-1} = u_1^{i-1}, \hat{U}_i \neq u_i \right\} \right] \\
&= \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ \hat{U}_1^{i-1} = u_1^{i-1}, f_M^{(i)}(y_1^N, \hat{U}_1^{i-1}) \neq u_i \right\} \right] \\
&\leq \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ f_M^{(i)}(y_1^N, u_1^{i-1}) \neq u_i \right\} \right] \\
&\leq \sum_{i \in \mathcal{A}_N} P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}).
\end{aligned}$$

□

Proposition 7.17.

$$P_{\mathrm{e}}(W, V, \mathcal{A}_N) \leq \sum_{i \in \mathcal{A}_N} \left(1 - I(W_N^{(i)}, V_N^{(i)}) \right). \tag{7.19}$$

Proof. The upper bound in Proposition 7.16 can be further extended by invoking for each of the synthetic channels the bound derived in (7.6), i.e.,

$$P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}) \leq 1 - I(W_N^{(i)}, V_N^{(i)}).$$

From this, (7.19) follows. □

Now, we list consecutively three theorems. Their proofs are given at the end.

In the first theorem, we give an expression for the polar mismatched capacity and provide a family of improving lower bounds to this capacity.

Theorem 7.18. *Let W and V be two B-DMCs. Then, the polar mismatched capacity is given by*

$$C_P(W,V) = \mathbb{P}[D_\infty(W,V) \to 1].$$

Furthermore, the following family of lower bounds

$$\mathbb{P}[D_\infty(W,V) = 1] \geq \frac{1}{2^n} \sum_{s^n \in \{+,-\}^n} \left| I(W^{s^n}, V^{s^n}) \right|^\dagger \tag{7.20}$$

holds for all $n = 0, 1, \ldots$.

The next theorem complements the previous achievability theorem with a coding theorem for B-DMCs with mismatched polar decoding.

Theorem 7.19. *For any $R < C_P(W,V)$ with $C_P(W,V) > 0$, there exists a sequence of information sets $\mathcal{A}_N \subset \{1, \ldots, N\}$, for $N = 2^n$ with $n = 0, 1, \ldots$, such that $|\mathcal{A}_N| \geq \lfloor NR \rfloor$ and*

$$Pe(W,V,\mathcal{A}_N) = O(2^{-\sqrt{N}}). \tag{7.21}$$

In a final theorem, we claim that the family of lower bounds in (7.20) is indeed asymptotically tight.

Theorem 7.20. *Let W and V be two B-DMCs. Under Conjecture 7.14, the polar mismatched capacity equals*

$$C_P(W,V) = \mathbb{P}[I_\infty(W,V) \to 1] = \lim_{n \to \infty} \frac{1}{2^n} \sum_{s^n \in \{+,-\}^n} \left| I(W^{s^n}, V^{s^n}) \right|^\dagger.$$

This last theorem should rather be interpreted as an elegant argument for the polar mismatched capacity. Otherwise, we believe it is not crucial from a practical point of view the moment we know that $C_P(W,V) = \mathbb{P}[D_\infty(W,V) = 1]$. In fact, identifying the indices of the 'good' synthetic channels would require the same type of computations whether the I parameters or D parameters of the channels are involved in the procedure. Now, it is time to prove the theorems.

Proof of Theorem 7.18. Let $\mathbb{P}_{W_N^{(i)}(.|u_i)}$ denote the probability of a set weighted under the distribution $W_N^{(i)}(.|u_i)$, for $u_i \in \{0,1\}$. The mismatched decoding error

probability over the i-th synthetic channel is given by

$$P_{\text{e, ML}}(W_N^{(i)}, V_N^{(i)}) = \mathbb{P}_{W_N^{(i)}(.|0)}\left[\Delta_{V_N^{(i)}} < 0\right] + \frac{1}{2}\mathbb{P}_{W_N^{(i)}(.|0)}\left[\Delta_{V_N^{(i)}} = 0\right] + \mathbb{P}_{W_N^{(i)}(.|1)}\left[\Delta_{V_N^{(i)}} > 0\right] + \frac{1}{2}\mathbb{P}_{W_N^{(i)}(.|1)}\left[\Delta_{V_N^{(i)}} = 0\right].$$

We will first study the implication of the result about the convergence of the process D_n on the convergence of the process $P_{\text{e, ML}}(W_n, V_n)$. Recall that by Proposition 7.11, $D_\infty \in \{0, 1\}$ a.s. We will consider the cases $D_n \to 0$ and $D_n \to 1$ separately and show that the two cases still correspond to a bad channel and good channel, respectively, in the mismatched scenario:

1. First note that $\mathbb{E}[|\Delta_n|] \to 0$ holds when $D_n \to 0$ and the convergence is to 0 both when the expectation is taken under $W_n(.|0)$ and $W_n(.|1)$. It is clear that $P_{\text{e, ML}}(W_n, V_n) \to 1$ in this case, and thus, the corresponding synthetic channels are bad for communication. In addition, we also notice that the value of $I_n(W, V)$ can only converge to a non-positive value over these channels as by Jensen's inequality we have

$$\mathbb{E}_{W_n(.|0)}[\log(1 + \Delta_{V_n})] \leq 0 \quad \text{and} \quad \mathbb{E}_{W_n(.|1)}[\log(1 - \Delta_{V_n})] \leq 0.$$

2. To show that $D_n \to 1$ corresponds to good channels we proceed as follows: for each of the $N = 2^n$ channels at the n-th stage of polarization, compute $\mathbb{P}_{W_N^{(i)}(.|0)}\left[\Delta_{V_N}^{(i)} \in [-1, -1 + \xi)\right]$, and for $\beta \in (0, 1)$, let $M_n(\beta)$ be the fraction of channels for which this value is larger than β:

$$M_n(\beta) := \frac{1}{2^n}\#\Big\{i \in \{1, \ldots, N = 2^n\} : \mathbb{P}_{W_N^{(i)}(.|0)}\left[\Delta_{V_N^{(i)}} \in [-1, -1 + \xi)\right] > \beta\Big\}.$$

Note that as D_n converges a.s. to a $\{0, 1\}$ valued random variable,

$$\mathbb{P}_{W_n(.|0)}\left[\Delta_{V_n} \in (-1 + \xi, -1 + \eta)\right] \to 0;$$

thus, for large n, the value of $M_n(\beta)$ is independent of the choice of ξ. Furthermore, by the martingale property of I_n:

$$I_0 := I(W, V) = \frac{1}{2}\frac{1}{2^n}\sum_{i=1}^{2^n}\mathbb{E}_{W_N^{(i)}(.|0)}\left[\log\left(1 + \Delta_{V_N^{(i)}}\right)\right] + \underbrace{\frac{1}{2}\frac{1}{2^n}\sum_{i=1}^{2^n}\mathbb{E}_{W_N^{(i)}(.|1)}\left[\log\left(1 - \Delta_{V_N^{(i)}}\right)\right]}_{\leq \frac{1}{2}\log 2}.$$

Thus,

$$I_0 \leq \frac{1}{2}M_n(\beta)\left(\beta \log \xi + \log 2\right) + \frac{1}{2}(1 - M_n(\beta)) \log 2 + \frac{1}{2}\log 2 \\ \leq \frac{1}{2}M_n(\beta)\beta \log \xi + \frac{3}{2}\log 2.$$

By the remark that $M_n(\beta)$ is not changed by the choice of ξ, we conclude that for any $\xi > 0$, $M_n(\beta)$ must vanish as n gets large, for otherwise, the right hand side will fail to be larger than I_0 for small enough ξ. Consequently when $D_n \to 1$, $\mathbb{P}_{W_n(.|0)}\left[\Delta_{V_n} \in (1-\xi, 1]\right] \to 1$. In a similar way, the argument can be repeated to show that $\mathbb{P}_{W_n(.|1)}\left[\Delta_{V_n} \in (-1, -1+\xi]\right] \to 1$ holds as well in this case. Thus, the value of $P_{\text{e, ML}}(W_n, V_n)$ must be vanishing over the synthetic channels for which $D_n \to 1$ and the channels are good.

Now, consider the sequence of information sets defined as

$$\mathcal{A}_N^{\gamma}(W, V) := \left\{i \in \{1, \ldots, N\} : \; D(W_N^{(i)}, V_N^{(i)}) \geq 1 - \gamma\right\}, \tag{7.22}$$

where $N = 2^n$ with $n = 1, 2, \ldots$ is the block-length and $\gamma \in (0, 1)$ is a desired threshold. By Proposition 7.16, the mismatched decoding error probability over the channel W of a polar code with information set $\mathcal{A}_N^{\gamma}(W, V)$, for a given N and γ, and using a mismatched successive cancellation decoder operating with the parameters of the channel V will be upper bounded by

$$P_{\text{e}}(W, V, \mathcal{A}_N^{\gamma}(W, V)) \leq \frac{1}{2^N} \sum_{i \in \mathcal{A}_N^{\gamma}(W,V)} P_{\text{e, ML}}(W_N^{(i)}, V_N^{(i)}),$$

Taking $N \to \infty$ and $\gamma \to 0$, we get $P_{\text{e}}(W, V, \mathcal{A}_N^{\gamma}(W, V)) \to 0$ via the previous discussion. We conclude that $C_P(W, V) = \mathbb{P}[D_{\infty}(W, V) \to 1]$ holds as claimed.

Next, we prove that the family of lower bounds in (7.20) holds. To begin with, we know that

$$\mathbb{P}[D_{\infty}(W, V) = 1] \geq I(W, V), \tag{7.23}$$

by Theorem 7.3. Next, we discuss a trivial improvement of this lower bound. The bound in (7.23) can be improved initially as

$$\mathbb{P}[I_{\infty}(W, V) = 1] \geq \left|I(W, V)\right|^{\dagger},$$

proving (7.20) for $n = 0$. Going a further step, we can improve the bound to

$$\mathbb{P}[I_{\infty}(W, V) = 1] \geq \frac{1}{2}\left|I(W^-, V^-)\right|^{\dagger} + \frac{1}{2}\left|I(W^+, V^+)\right|^{\dagger},$$

and more generally to

$$\mathbb{P}[I_\infty(W,V)=1] \geq \frac{1}{2^n} \sum_{s^n \in \{+,-\}^n} \left|I(W^{s^n},V^{s^n})\right|^\dagger,$$

for any $n = 0, 1, \ldots$. This concludes the proof. □

Proof of Theorem 7.19. Consider once more the sequence of information sets we defined in (7.22) in the previous proof. The claim on the construction of $\mathcal{A}_N^\gamma(W,V)$ follows by Theorem 7.18 and Theorem 7.4 (now, we can replace the left hand side of (7.9) with $C_P(W,V)$); selecting the threshold $\gamma \simeq 2^{-\sqrt{N}}$ ensures $|A_N^\gamma(W,V)| \geq NR$, for $R < C_P(W,V)$. Taking $N \to \infty$, we get $\gamma \to 0$, and we know from the proof of Theorem 7.18 that $P_{\mathrm{e}}(W,V,\mathcal{A}_N^\gamma(W,V)) \to 0$. To prove (7.21), we simply note that the speed of polarization of $D_n(W,V)$ we found to be $O(2^{-\sqrt{N}})$ also determines the speed of polarization of the resulting mismatched polar decoding error probability process. However, to achieve a positive rate, $C_P(W,V) > 0$ must hold. The family of lower bounds in (7.20) show that, in fact, $I(W,V) > 0$ is a sufficient condition for $C_P(W,V) > 0$. □

Proof of Theorem 7.20. The result follows by Theorem 7.18 and Conjecture 7.14. Observe that once $I(W^{s^n},V^{s^n}) \approx 1$ for a particular synthetic channel, then the mismatched capacity of the channel, which should be larger, satisfies $C(W^{s^n},V^{s^n}) \approx 1$. On the other hand, as in general $I(W^{s^n},V^{s^n})$ is only a lower bound to $C(W^{s^n},V^{s^n})$, we cannot conclude at first glance that whenever $I(W^{s^n},V^{s^n}) \approx 0$ holds for a particular synthetic channel, we also have $C(W^{s^n},V^{s^n}) \approx 0$. Nevertheless, we conclude so by the conjecture. □

7.3.1 Code Construction

Here, we discuss how the information sets described in (7.22) can be constructed under various communication scenarios. Suppose first that the exact channel knowledge is not available and feedback is allowed at the decoder. Assume the code designer not knowing the actual communication channel, say W, decides to implement the polar decoder with respect to a channel V. To handle this scenario, we propose to use the original polar code construction idea of Arıkan [2] in an 'online' fashion. The method is based on the estimation of the parameters $D(W_N^{(i)},V_N^{(i)})$ by a Monte Carlo approach. To that end, the encoder needs to perform multiple transmissions of an input and the decoder has to compute an estimate of the parameters by averaging. Once the information set is constructed, the decoder shall reveal this information to the encoder by feedback. If the encoder knows the channel W, the estimation of the parameters $D(W_N^{(i)},V_N^{(i)})$, can be carried 'offline' at the encoder and feedback from

the decoder will not be necessary. If besides the encoder, the decoder also knows the true channel, the information set can be computed 'offline' at the decoder side as well.

7.3.2 Channel Symmetry

So far, we only discussed the specification of the information set. However, polar coding also relies on the specification of the values of the frozen inputs. Recall that in the matched case, when the channel is symmetric, the frozen values can simply be fixed to 0 as any other choice of the frozen values is as good as this one. We will now show that the notion of channel symmetry defined in Definition 7.15 also simplifies considerably the task when there is a decoding mismatch.

The following proposition can be proved using the symmetry properties derived in [2, Corollary 1].

Proposition 7.21. *Let $W : \mathbb{F}_2 \to \mathcal{Y}$ and $V : \mathbb{F}_2 \to \mathcal{Y}$ be B-DMCs symmetrized by the same permutation. Then, for any $u_1^N \in \{0,1\}^N$,*

$$P_{\text{e, ML}}(W_N^{(i)}, V_N^{(i)}) = \sum_{y_1^N \in \mathcal{Y}^N} W_N(y_1^N|u_1^N)\mathbf{1}\big\{L_{V_N^{(i)}}(y_1^N, u_1^{i-1}) > 1\big\} + \frac{1}{2}\sum_{y_1^N \in \mathcal{Y}^N} W_N(y_1^N|u_1^N)\mathbf{1}\big\{L_{V_N^{(i)}}(y_1^N, u_1^{i-1}) = 1\big\}.$$

From the last proposition, we deduce that if the true channel and the mismatched channel used in the decision procedure are symmetrized by the same permutation, all choices of the frozen values are equally favorable, and thus, the values can simply be all fixed to 0.

7.3.3 Complexity

Polar codes with mismatched polar decoding use exactly the same encoding and decoding architectures as of polar codes with matched polar decoding. Therefore, as explained by Arıkan [2], these components of the communication system can be implemented in $O(N \log N)$ complexity.

As for the complexity of the code construction, no general low-complexity algorithm exists for the Monte Carlo approach (online or offline). The same complexity issues which were raised in [3], [31] after Arıkan published his work [2] do apply. Nevertheless, we believe computationally more efficient alternatives can be proposed for the offline Monte Carlo approach by extending the efficient code con-

struction method proposed in [3] to the mismatched case. This topic requires further investigation.

7.4 Polar vs. Classical Mismatched Capacity

Some readers might be suspicious upon reading the title of this section of whether this will be a fair comparison or not. Let us share our opinion on this matter. By the converse to the channel coding theorem, we know that no matter what coding scheme (encoder and decoder) we are using, we cannot communicate at a rate higher than the capacity of the channel. In that respect, $C(W)$ is *the fundamental* channel parameter. On the other hand, if we initially force the decoder to follow a specific decision rule during the decoding procedure, a new channel parameter is likely to acquire an operational meaning. As a response to setting the decoding rule over the channel W to mismatched maximum likelihood decoding with respect to another channel V, the literature would 'replace' the capacity of the channel W with the mismatched capacity $C(W, V)$, and for B-DMCs, point out a computable expression to it due to Balakirsky [40]. In the same spirit, introducing a mismatch in the decoding procedure of the polar decoder, we have just pointed out in the previous section an expression for the polar mismatched capacity $C_P(W, V)$. In addition, we observe that the decision functions used by the mismatched polar decoder resemble mismatched maximum likelihood decision functions. This particular resemblance due to the nature of the mismatch introduced in both of the decoding rules justifies our fair reflex for comparing $C_P(W, V)$ with $C(W, V)$. We note, however, that the mismatched maximum likelihood decoder and the mismatched polar decoder are different since the decision functions of the mismatched polar decoder treat the future frozen bits as random variables[3]. In this section, we will see that despite the similitude, the behavior of both sub-optimal decoders are quite different in nature; the mismatched polar decoder might outperform the mismatched maximum likelihood decoder in some cases.

7.4.1 Review of Balakirsky's Results

In this section, we revisit the results of [40] on the mismatched capacity of B-DMCs. Let $W\colon \mathbb{F}_2 \to \mathcal{Y}$ be a B-DMC. We fix an input distribution $P(x)$ on $\mathcal{X}$. Some standard definitions follow:

$$H(PW) := -\sum_{y\in\mathcal{Y}} PW(y)\log PW(y),$$

[3]Note that the same comment is made in [2] regarding the similarity between the polar successive cancellation decoding rule and the maximum likelihood decoding rule which is optimal.

with $PW(y) = \sum_{x \in \mathcal{X}} P(x)W(y|x)$.

$$H(P|W) := -\sum_{y \in \mathcal{Y}} P(x)W(y|x) \log W(y|x).$$

So, $I(P;W) = H(PW) - H(W|P)$.

The following result due to Balakirsky gives a closed form expression for $C_d(W)$ when W is a B-DMC.

Theorem 7.22. *[40] For any B-DMC $W : \mathbb{F}_2 \to \mathcal{Y}$ and any $d(x,y)$,*

$$C_d(W) = \max_P I_d(P;W),$$

where

$$I_d(P;W) := \min_{\substack{W' : \\ PW'(y)=PW(y) \\ d(P,W') \le d(P,W)}} I(P;W'),$$

with $d(P,W) := \sum_{x \in \mathbb{F}_2} \sum_{y \in \mathcal{Y}} P(x)W(y|x)d(x,y)$.

Using this closed from expression, Balakirsky studies the computation of $C_d(W)$ for symmetric B-DMCs when the d-decoder preserves the symmetry structure of the communication channel. In the following two examples, we revisit his examples.

Example 7.23. *[40, Examples, Statement 1] For a binary input binary output channel W and any given d-decoding rule $I_d(P,W) = I(P,W)\mathbf{1}\{A\}$ holds with*

$$A = \{\text{sign}\,(1 - W(0|0) + W(1|1)) \\ = \text{sign}\,(d(0,0) + d(1,1) - d(0,1) - d(1,0))\}.$$

So, $C_d(W) = C(W)$ or 0.

Example 7.24. *[40, Examples, Statement 2] Let $W : \mathbb{F}_2 \to \mathcal{Y}$ be a B-DMC with $\mathcal{Y} = \{0, 1, \ldots, L-1\}$. Suppose the transition probability matrix of the channel W and the corresponding metrics for the additive d-decoder are given by*

$$W = \begin{bmatrix} w_0 & w_1 & \ldots & w_{L-1} \\ w_{L-1} & w_{L-2} & \ldots & w_0 \end{bmatrix},$$

and

$$d = \begin{bmatrix} d_0 & d_1 & \ldots & d_{L-1} \\ d_{L-1} & d_{L-2} & \ldots & d_0 \end{bmatrix}. \tag{7.24}$$

Then, the mismatched capacity is achieved for P_{unif} on $\{0,1\}$ and is given by

$$C_d(W) = H(P_{\text{unif}}W) - H(W'|P_{\text{unif}}),$$

where W' is given by

$$W' = \begin{bmatrix} w'_0 & w'_1 & \dots & w'_{L-1} \\ w'_{L-1} & w'_{L-2} & \dots & w'_0 \end{bmatrix},$$

with

$$w'_y = (w_y + w_{L-1-y}) \frac{e^{-\alpha.d_y}}{e^{-\alpha.d_y} + e^{-\alpha.d_{L-1-y}}}, \tag{7.25}$$

for $y \in \mathcal{Y}$, and the parameter $\alpha \geq 0$ is chosen to satisfy the condition:

$$\sum_{y\in\mathcal{Y}} w'_y d_y = \sum_{y\in\mathcal{Y}} w_y d_y. \tag{7.26}$$

7.4.2 No Conservation Property for Balakirsky's Converse

For the rest of this section, we restrict the additive decoders to mismatched maximum likelihood decoders. The goal of this subsection is to show that:

($g1$) There are pairs of B-DMCs $W\colon \mathbb{F}_2 \to \mathcal{Y}$ and $V\colon \mathbb{F}_2 \to \mathcal{Y}$ for which

$$C(W^+, V^+) + C(W^-, V^-) > 2C(W, V).$$

($g2$) Furthermore, there exist cases for which

$$C_P(W, V) > C(W, V).$$

($g3$) However, there are also cases for which

$$C(W^+, V^+) + C(W^-, V^-) < 2C(W, V).$$

To that end, we shall study the evolution of the mismatched capacity of B-DMCs under the one-step polar transform when the communication channel and the mismatched channel used in the decision procedure satisfy a certain symmetry property.

Let V be a B-DMC symmetrized by the same permutation as the channel W defined in Example 2. Recall that we introduced this notion of symmetry in Definition

7.15. Suppose the transition probability matrix of V is given by

$$V = \begin{bmatrix} v_0 & v_1 & \dots & v_{L-1} \\ v_{L-1} & v_{L-2} & \dots & v_0 \end{bmatrix}.$$

Then, the corresponding additive decoder can be defined as in (7.24) by letting $d_y = -\log v_y$, for $y = 0, \dots, L-1$. In this case, the mismatched capacity equals

$$C(W, V) = I(P_{\text{unif}}, W')$$

where the transition probabilities of W' can be computed by replacing the relations in (7.25) and (7.26) with the following relations:

$$w'_j = (w_j + w_{L-1-j}) \frac{v_j^\alpha}{v_j^\alpha + v_{L-1-j}^\alpha},$$
$$\sum_{y \in \mathcal{Y}} w'_y \log v_j = \sum_{y \in \mathcal{Y}} w_y \log v_j.$$

We begin with an example which will help us achieve our first two goals by illustrating specific pairs of BSCs for which $C(W, V) = 0$, but $C_P(W, V) = I(W)$.

Example 7.25. Let W be a BSC of crossover probability $p \in (0, 0.5)$ and V be the BSC of crossover probability $1 - p$. In this example, we will answer the following three questions:

($q1$) Suppose we transmit over the channel W, but do mismatched maximum likelihood decoding with respect to the channel V as shown in Figure 7.1. What is the mismatched capacity $C(W, V)$?

($q2$) Suppose we first apply the polar transform to synthesize the channels W^+, W^- and V^+, V^-, and then we communicate using the architecture given in Figure 7.2. What is the mismatched capacity $C(W^-, V^-)$ in this case?

($q3$) Suppose we communicate over the channel W using polar coding, and we do mismatched polar decoding with respect to the channel V. The communication architecture is shown in Figure 7.3. What is the mismatched capacity of polar coding $C_P(W, V)$?

Here are the corresponding answers:

($a1$) In this case, the crossover probabilities of the BSCs are not in harmony. By the result given in Example 7.23, we conclude that $C(W, V) = 0$.

($a2$) It is known that after applying the minus polar transform to a BSC of crossover probability $\alpha \in [0, 1]$, the synthesized channel is also a BSC, and with

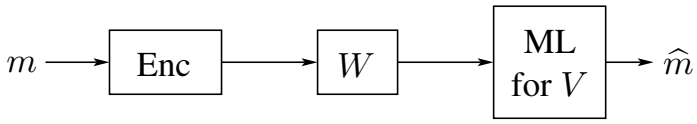

Figure 7.1: Classical mismatched decoding.

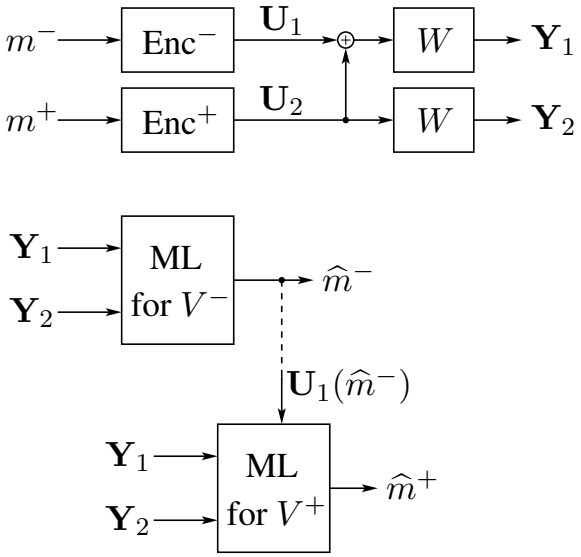

Figure 7.2: One-step polarization architecture for mismatched decoding.

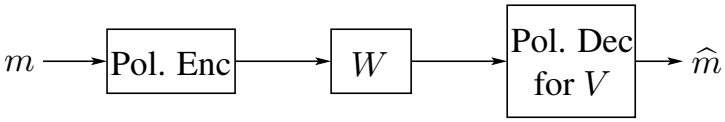

Figure 7.3: Polar mismatched decoding.

crossover probability $2\alpha(1-\alpha)$. So, both W^- and V^- are the same BSC with crossover probability $p^- = 2p(1-p)$. Therefore, the mismatched capacity of the minus channel equals its matched capacity, i.e., $C(W^-, V^-) = I(W^-) = 1 - h_2(p^-)$. As a result,

$$C(W^+, V^+) + C(W^-, V^-) \geq C(W^-, V^-) > 2C(W, V) = 0,$$

and we conclude that, in this example, the polar transform strictly improves $C(W, V)$. We have thus achieved our first goal (g1).

$(a3)$ We found that $V^- = W^-$. It is easy to see that while $V^+ \neq W^+$, one has $V^{+-} = W^{+-}$, and indeed, $V^{++-} = W^{++-}$, ... Consequently, for any sequence $s^n \in \{-,+\}^n$ of polar transformations, $V^{s^n} = W^{s^n}$, except when $s^n = +\cdots+$. Thus, $C(W^{s^n}, V^{s^n}) = I(W^{s^n})$ for all $s^n \neq +\cdots+$ and we see that $C_P(W, V) = I(W)$. Thanks to this result, we also reached our second goal (g2).

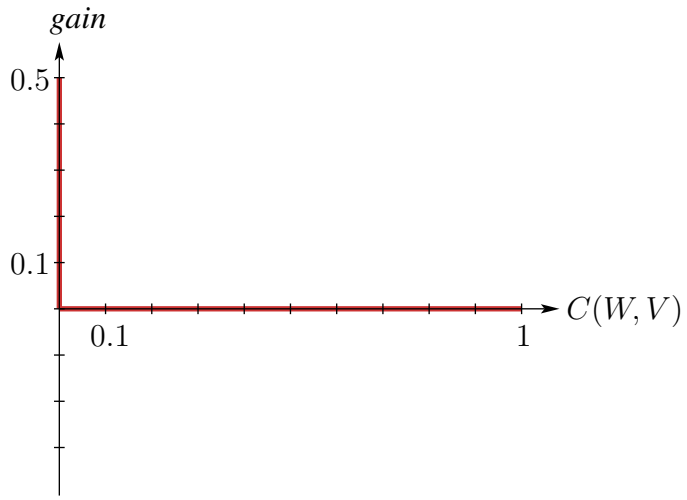

Figure 7.4: $C(W,V)$ vs. *gain* after a one-step polar transform for $|\mathcal{Y}| = \{2,3\}$.

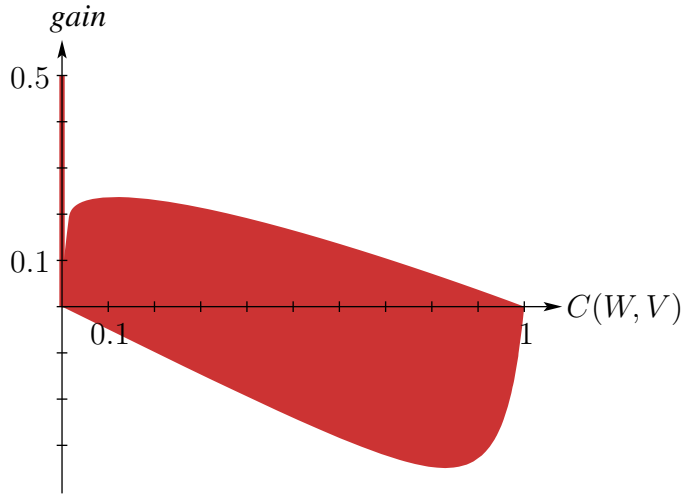

Figure 7.5: $C(W,V)$ vs. *gain* after a one-step polar transform for $|\mathcal{Y}| = 4$.

After such a motivating example, one is curious about whether the improvement we illustrated for the specific pair of BSCs extend to other pairs of mismatched B-DMCs as well: Does polarization create $C(W,V)$? To satisfy one's curiosity, we now present the results of the numerical experiments we carried for random pairs of channels with various output alphabet sizes. Let

$$\textit{gain} := C(W^+,V^+) + C(W^-,V^-) - 2C(W,V)$$

denote the amount of capacity gained after one-step. The numerical experiments show that:

(i) When the output alphabet is binary or ternary, one-step improvement with *gain* > 0 happens only when $C(W,V) = 0$ and can be as large as $1/2$. When

$C(W,V) > 0$, we observe that *gain* $= 0$, so one has neither improvement nor loss. See Figure 7.4.

(ii) When the output alphabet contains four or more symbols, improvement may happen not only when $C(W,V) = 0$ but also when $C(W,V) > 0$; however, there are cases when $C(W^+,V^+) + C(W^-,V^-) < 2C(W,V)$, so one may encounter a loss after a one-step transformation, i.e., *gain* < 0. See Figure 7.5.

(ii) Finally, the numerical experiments also suggest that

$$\sum_{s\in\{+,-\}^n} \frac{1}{2^n} \left| I(W^{s^n}, V^{s^n}) \right|^{\dagger} \leq C(W,V)$$

holds whenever $C(W,V) > 0$, (we did experiment only for $n = 1, 2, 3$). Thus, $C_P(W,V) \leq C(W,V)$ is a likely conjecture for the case where $C(W,V) > 0$.

The above numerical study showed that $C(W,V)$ can be as well lost after applying the polar transform, validating the claim in our third goal (g3). Therefore as opposed to $I_n(W)$, $C_n(W,V)$ does not qualify for being the 'martingale' of the mismatched analogy. In fact, we know from the earlier analysis that $I_n(W,V)$ qualifies for that.

7.4.3 Boosting the Mismatched Capacity via Polarization

The study by Balakirsky [40] gave the initial impulse for the study of this section. We adapted his example [40, Examples, Statement 2] which computes $C_d(W)$ of a symmetric B-DMC W when the additive decoder shares the symmetry structure of W to mismatched decoders, and we carried numerical experiments based on this computation to compare $C(W,V)$ with the sum $C(W^+,V^+) + C(W^-,V^-)$. The experiments revealed that, as opposed to $I(W,V)$, $C(W,V)$ is not necessarily a conserved quantity under the polar transform. Nevertheless, as the choice of a coding scheme is part of a design problem and using the maximum likelihood decoding rule with the metric of a channel V for decoding a long sequence or using the same rule for decoding two long sequences with the metrics of the channel V^- first and then the channel V^+ in a successive cancellation decoding configuration does not differ so much in complexity (asymptotically), the numerical study shows that in some cases the mismatch at the decoder can be better exploited by using the polarization architecture of Figure 7.2. Therefore, communication rates higher than $C(W,V)$ can be achieved in some cases by integrating the polarization architecture of Arıkan into the classical mismatched communication scenario. Furthermore, by studying specific pairs of BSCs W and V such that $C(W,V) = 0$, but $C_P(W,V) > 0$, we showed that there exist channels for which the sequence of polar transformations strictly improve the mismatched capacity of B-DMCs.

Let us explore the new perspective these results brings. One main motivation behind the study of mismatched decoders is the importance to know what can be done in a communication system where the channel is W but the decoder is designed with the belief that it is a different one, say V. The answer to the highest possible transmission rate that can be achieved over the channel under such a disbelief at the decoder side is in fact equal to the capacity of the channel and can be achieved by the MMI decoding rule [11]. Thus, the universal MMI decoder certainly wins out over the other options in terms of achievable rates, but requires in most cases an unpractical amount of resources to be implementable [12]. Bringing the decoding complexity into play, the lower complexity alternatives deserve attention. From a conceptual point of view, the coding rates dictated by Balakirsky's expression $C(W, V)$, the polar mismatched capacity $C_P(W, V)$, and also Balakirsky's mismatched capacity achieving coding schemes used successively over the plus and minus synthetic channels (aka with the polarization architecture) can all be seen as lower bounds to the highest achievable rate. To this effect, since we have seen that all these non-optimal mismatched decoders might result in quite different behaviors, we conclude that in different cases a better lower bound can be obtained, whence a higher transmission rate be achieved, by integrating the one-step architecture to the coding scheme or by polar coding. Moreover, designers shall take into account that even when $C_P(W, V) < C(W, V)$, the $O(N \log N)$ complexity of the encoding and decoding structures of polar coding has a practical edge over the classical codes suitable for mismatched maximum likelihood decoding.

Appendix

7.A *Mismatched $\mathscr{E}_o$ à la Gallager*

In this Appendix, we show that Gallager's style error exponent derivation to assess the performance of random codes with mismatched maximum likelihood decoding also leads to the quantity $I(W, V)$. For the sake of brevity, we refer to [5, Chapter 5] for notations.

Let $W : \mathcal{X} \to \mathcal{Y}$ and $V : \mathcal{X} \to \mathcal{Y}$. We fix the input distribution to $Q(x)$. Suppose the input $\mathbf{x}_m \in \mathcal{X}^N$, for $m \in \{1, \ldots, M = \lceil 2^{NR} \rceil\}$, is transmitted over the channel W and the received output $\mathbf{y} \in \mathcal{Y}^N$ is decoded using a mismatched maximum likelihood decoder with respect to the channel V. We define

$$P_{\mathrm{e},m} := \sum_{\mathbf{x}_m} \sum_{\mathbf{y}} Q^N(\mathbf{x}_m) W^N(\mathbf{y}|\mathbf{x}_m) \mathbb{P}[\text{error}|m, \mathbf{x}_m, \mathbf{y}].$$

Then, the union bound gives

$$\mathbb{P}[\text{error}|m, \mathbf{x}_m, \mathbf{y}] \leq \mathbb{P}\left[\bigcup_{m' \neq m} A_{m'}\right] \leq \left(\sum_{m' \neq m} \mathbb{P}[A_{m'}]\right)^{\rho},$$

for all $0 < \rho \leq 1$, where $A_{m'}$ occurs when $\log \dfrac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)} \geq 0$. Using Chernoff bound with $s > 0$, we get

$$\begin{aligned}
\mathbb{P}[A_{m'}] =& \mathbb{P}\left[\log \frac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)} \geq 0\right] = \mathbb{P}\left[\exp_2\left\{s \log \frac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)}\right\} \geq 1\right] \\
\leq& \mathbb{E}\left[\exp_2\left\{s \log \frac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)}\right\}\right] = \prod_{i=1}^{N} \mathbb{E}\left[\left(\frac{V(y_i|x_i')}{V(y_i|x_i)}\right)^s\right] \\
&= \sum_{\mathbf{x}_{m'}} Q^N(\mathbf{x}_{m'}) \left(\frac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)}\right)^s.
\end{aligned}$$

Therefore, we get:

$$\begin{aligned}
P_{e,m} &= \sum_{\mathbf{x}_m} \sum_{\mathbf{y}} Q^N(\mathbf{x}_m) W^N(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \sum_{\mathbf{x}_{m'}} Q^N(\mathbf{x}_{m'}) \left(\frac{V^N(\mathbf{y}|\mathbf{x}_{m'})}{V^N(\mathbf{y}|\mathbf{x}_m)}\right)^s\right)^{\rho} \\
&\leq (M-1)^{\rho} \sum_{\mathbf{x}_m} \left(\sum_{\mathbf{y}} Q^N(\mathbf{x}_m) \frac{W^N(\mathbf{y}|\mathbf{x}_m)}{V^N(\mathbf{y}|\mathbf{x}_m)^{s\rho}}\right) \times \\
&\qquad\qquad \left(\sum_{\mathbf{x}_{m'}} Q^N(\mathbf{x}_{m'}) V^N(\mathbf{y}|\mathbf{x}_{m'})^s\right)^{\rho} \\
&\leq 2^{\rho N R} \left[\sum_{y} \left(\sum_{x} Q(x) \frac{W(y|x)}{V(y|x)^{s\rho}}\right) \left(\sum_{x} Q(x) V(y|x)^s\right)^{\rho}\right]^N.
\end{aligned}$$

Hence, we could define the *mismatch exponent* as:

$$P_{e,m} \leq \exp_2\left\{-N\left(E_0\left[\rho, s, Q, W, V\right) - \rho R\right]\right\},$$

where

$$E_0(\rho, s, Q, W, V) := -\log\left(\sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} Q(x) \frac{W(y|x)}{V(y|x)^{s\rho}}\right) \left(\sum_{x \in \mathcal{X}} Q(x) V(y|x)^s\right)^{\rho}\right).$$

In the case of binary equally likely inputs $\mathcal{X} = \mathbb{F}_2$ and the choice $s = \frac{1}{1+\rho}$, the

expression simplifies to

$$E_0(\rho, W, V) := -\log\left(\sum_{y\in\mathcal{Y}}\frac{1}{2}\left(\frac{W(y|0)}{V(y|0)^{\frac{\rho}{1+\rho}}}+\frac{W(y|1)}{V(y|1)^{\frac{\rho}{1+\rho}}}\right)\times \left(\frac{1}{2}V(y|0)^{\frac{1}{1+\rho}}+\frac{1}{2}V(y|1)^{\frac{1}{1+\rho}}\right)^{\rho}\right).$$

When $W = V$, $E_0(\rho, W, V)$ reduces to $E_0(\rho, W)$.

When $\rho = 0$, we get

$$\left.\frac{\partial E_0(\rho, W, V)}{\partial\rho}\right|_{\rho=0} = -\frac{1}{2}\sum_y W(y|0)\log\left(\frac{1+V(y|1)/V(y|0)}{2}\right) - \frac{1}{2}\sum_y W(y|1)\log\left(\frac{1+V(y|0)/V(y|1)}{2}\right),$$

which is equal in fact to the channel parameter $I(W, V)$:

$$I(W, V) = \frac{1}{2}\sum_{y\in\mathcal{Y}} W(y|0)\log\left(\frac{V(y|0)}{q_V(y)}\right)+\frac{1}{2}\sum_{y\in\mathcal{Y}} W(y|1)\log\left(\frac{V(y|1)}{q_V(y)}\right),$$

where $q_V(y)$ and $\Delta_V(y)$ are defined in (2.14) and (2.15), respectively.

When $\rho = 1$, we get

$$E_0(1, W, V) = -\log\left(\frac{1+\frac{1}{2}\sum_y W(y|0)\sqrt{\frac{V(y|1)}{V(y|0)}}+\frac{1}{2}\sum_y W(y|1)\sqrt{\frac{V(y|0)}{V(y|1)}}}{2}\right).$$

So, we can define a *mismatched Bhattacharyya distance* as follows:

$$Z(W, V) := \frac{1}{2}\sum_{y\in\mathcal{Y}} W(y|0)\sqrt{L_V(y)}+\frac{1}{2}\sum_{y\in\mathcal{Y}} W(y|1)\frac{1}{\sqrt{L_V(y)}}, \tag{7.27}$$

where $L_V(y) = V(y|1)/V(y|0)$. Both parameters reduce to the ordinary ones in the matched case, i.e., $I(W, W) = I(W)$ and $Z(W, W) = Z(W)$.

At last, we remark that keeping the parameter s in the derivations, Kaplan and Shamai [43] reach a more general version of Fischer's generalized mutual information [42].

Chapter 8

Designing Robust Polar Codes over B-DMCs

The design of a polar code is a channel specific task [2]. A central stage in this procedure is the construction of the information set selecting the synthetic channels' indices that are good for uncoded transmission. Once the information set is constructed, the encoder and the decoder operate with the common knowledge of the information set, knowing over which indices to transmit and decode data and over which to use priorly fixed bits. Besides the information set, the operation of the successive cancellation decoder is channel dependent. The dependence of the system components of polar coding on the communication channel makes computing the information set a challenging problem in scenarios where the channel is unknown or only a partial knowledge exists. An unknown quotation taken from the page 77 of the Legendary Quotebook from 2050 [no citation is available yet] even says: "In any coding problem, it is unrealistic to assume that the true channel W and the design channel V are the same." Thanks to the coding theorem we proved in the previous chapter for the mismatched communication scenario, we know that even if the true channel is unknown, polar codes can still be designed with respect to a mismatched channel and positive communication rates can be achieved by polar coding. In this chapter, we rise to a new challenge and study the robustness of polar codes over compound B-DMCs.

8.1 Communication over a Class of Channels

In practical communication scenarios, only a partial knowledge on the communication channel might be available. For instance, we might only know the possible range of values that the mutual information between the input and output of the channel takes. The study of reliable communication under channel uncertainty be-

comes relevant to tackle these situations in which a complete channel knowledge is missing. The qualitative notion of channel uncertainty is made more concrete with the definition of more complex channel models. The *compound channel* model is one such instance where the unknown channel is restricted to belong to a given class of channels. Both the encoder and the decoder are ignorant of which of the channels in the class occurs, but the channel remains fixed during the communication of a given codeword. Hence, the selected code should ensure good performance for all the channels in the class.

Blackwell et al. [10] found the *capacity of a class of DMCs* $\mathcal{W}$ to be

$$C_{\text{comp}}(\mathcal{W}) = \max_{P(x)} \inf_{W \in \mathcal{W}} I(P; W). \tag{8.1}$$

This is the supremum of all achievable transmission rates when the code is required to perform well in all the channels in the class. The direct part and the converse part are proved in [10, Theorem 1]. To be more specific, a maximum error probability criteria is used to asses performance: a sequence of $(2^{NR}, \xi_N, N)$ codes of block-length N for the class with $M = \lfloor 2^{NR} \rfloor$ codewords and corresponding sequence of disjoint decoding regions $B_m \subset \mathcal{Y}^N$, for $m = 1, \ldots, M$, is required to satisfy for all the codewords $\mathbf{x}_m \in \mathcal{X}^N$ and for every channel $W \in \mathcal{W}$ an upper bound $\sum_{\mathbf{y} \in B_m^c} W(\mathbf{y}|\mathbf{x}_m) \leq \xi_N$, where $\xi_N \to 0$. The key point is that the exponential error bound ξ_N given in the theorem does not depend on the actual channel that occurs during the transmission of a codeword, but depends only on $C_{\text{comp}}(\mathcal{W})$, R, and the alphabet sizes $|\mathcal{X}|$ and $|\mathcal{Y}|$.

A more modern analysis of the compound communication problem which uses the method of types is due to Csiszár and Körner [11]. In this analysis, the constructed code is substituted with a constant composition code from a random type ensemble based on the error exponent given in (7.3) and decoded with the MMI decoder. Though both techniques result in the same expression, the MMI decoder is more robust in the sense that it does not even need the knowledge of the compound set which is required by Blackwell's construction. Yet, neither of the methods is practically implementable.

In general, $C_{\text{comp}}(\mathcal{W})$ is smaller than the infimum of the capacities of any channel in $\mathcal{W}$. Nevertheless, if we restrict the analysis to a symmetric class of channels $\mathcal{W}_s$, then

$$C_{\text{comp}}(\mathcal{W}_s) = \inf_{W \in \mathcal{W}_s} I(P_{\text{unif}}; W), \tag{8.2}$$

where P_{unif} denotes the uniform input distribution. To show this, note that

$$\max_{P(x)} \inf_{W \in \mathcal{W}_s} I(P; W) \leq \inf_{W \in \mathcal{W}_s} \max_{P(x)} I(P; W)$$

holds as a general fact. As the uniform input distribution is the capacity achieving input distribution of all symmetric channels, for any $W \in \mathcal{W}_s$, we have $I(P_{\text{unif}}; W) = \max_{P(x)} I(P; W)$. Thus, the inequality in the opposite direction is also true:

$$C_{\text{comp}}(\mathcal{W}_s) = \max_{P(x)} \inf_{W \in \mathcal{W}_s} I(P; W) \geq \inf_{W \in \mathcal{W}_s} I(P_{\text{unif}}; W).$$

However, even if the compound capacity expression is of the form (8.2), it is not necessarily true that a capacity-achieving code (encoder/decoder) designed for the 'worst' channel in the class will also achieve the same rate if used over all of the other channels in the class. The decoding regions given for the worst channel may not be adequate for the other channels in the class and may not satisfy the required error probability criteria. For instance, the typicality decoder for the worst channel will generally fail for other channels.

In order to accommodate such a worst case design, the compound set of channels must be well-structured. For example, it is sufficient that the class of channels $\mathcal{W}$ be a *convex compact set*. In this case, the compound capacity is given by

$$C_{\text{comp}}(\mathcal{W}) = \max_{P} \min_{W \in \mathcal{W}} I(P; W) = \min_{W \in \mathcal{W}} I(P^*; W) = I(P^*; W^*),$$

where $P^* = \arg\max_{P(x)} I(P; W)$, and this capacity can be achieved by d decoding with respect to $d(x, y) = -\log W^*(y|x)$ [12, Remarks ii)]. Thus, mismatched maximum likelihood decoding with respect to the 'worst' channel in the class is optimal for convex compound sets . This is a consequence of the following property.

Proposition 8.1. *[11, Proof of Corollary 6.10] Given a set of channels $\mathcal{W}$, its convex closure $\overline{\mathcal{W}}$, and a fixed $P(x)$, let $V = \arg\min_{W \in \overline{\mathcal{W}}} I(P; W)$. Then,*

$$I(P; \alpha W + (1 - \alpha)V) \geq I(P; V), \quad \forall \alpha \in [0, 1], \forall\, W \in \mathcal{W}.$$

This implies the following inequality:

$$I(P; W, V) := \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x) W(y|x) \log \frac{V(y|x)}{\sum_{x' \in \mathcal{X}} P(x') V(y|x')} \geq I(P; V), \tag{8.3}$$

where $I(P; W, V)$ denotes the generalized mutual information:

In the previous chapter, we mentioned that $I(W, V)$ is also a lower bound to the mismatched capacity $C(W, V)$ of the channel W with mismatched maximum likelihood decoding with respect to V. In fact, Fischer [42] showed that more generally $C(W, V) \geq I(P; W, V)$ holds, for any distribution $P(x)$ on the inputs of the channel. So for any channel W in a convex compound set $\mathcal{W}$, we have

$$C(W, V) \geq I(P^*; W, V) \geq I(P^*; V) = C_{\text{comp}}(\mathcal{W}),$$

by Proposition 8.1. This proves the achievability claim [12, Remarks ii)].

An example of a convex set of channels is the class of binary symmetric channels with crossover probabilities within an interval. Note that convexity is a sufficient but not necessary condition for (8.3) to hold. One-sided sets defined by Abbe and Zheng [45] generalize this idea. The following definition is adapted from [45, Definition 3].

Definition 8.2. [45] A set $\mathcal{W}$ is called *one-sided* with respect to the input distribution $P(x)$ if the following minimizer is unique:

$$V = \arg \min_{W \in \mathrm{cl}(\mathcal{W})} I(P; W),$$

where $\mathrm{cl}(\mathcal{W})$ is the closure of $\mathcal{W}$, and if the condition in (8.3) holds. (Though this condition is stated as a divergence inequality in [45, Equation 8], one can check that it is equivalent to (8.3).). By [45, Lemma 4] convex sets are one-sided, but there exist one-sided sets that are not convex.

When the analysis is restricted only to the uniform input distribution as in our case, the *symmetric compound capacity* of a class of channels can be defined as

$$I_{\text{comp}}(\mathcal{W}) = \min_{W \in \mathcal{W}} I(W), \tag{8.4}$$

A quick inspection reveals that the inequality (8.3) given in Proposition 8.1 is equivalent to the following condition:

$$I(W, V) := I(P_{\text{unif}}; W, V) \geq I(V), \quad \forall\, W \in \mathcal{W}. \tag{8.5}$$

This is good news as we are already quite familiar with the parameter $I(W, V)$.

What's Coming, Doc?

It is shown in the Appendix A (at the end of the thesis) that there exist linear codes which achieve the symmetric compound capacity of a class of channels. In the light of this information, the question we pose is: Does the class of polar codes, as a subset of linear codes, achieve the symmetric compound capacity of a class of channels? In [34], the compound capacity of polar codes is studied and it is shown that in general polar codes do not achieve the symmetric compound capacity of a class of channels even if the channels in the class are symmetric. The underlying assumption throughout the analysis in [34] is the availability of the actual communication channel to the decoder. This transforms the problem into finding an information set providing suitable indices for communication over any channel in the class. As the lack of knowledge cannot possibly reverse the end result, the assumption simplifies the theoretical analysis for drawing the general conclusion.

On the other hand, the following question has not been addressed so far: what rates can be achieved in the compound setting? Both from conceptual and practical points of view, the effect of a design mismatch on the achievable rates over a class of channels is an important problem to be studied. We will provide a partial characterization of this problem by identifying channel conditions leading to universal polar code designs.

The first three sections of this chapter will report the following conclusions:

- We will start by considering the problem of universal polar coding when the decoder knows the communication channel. In this case, the order preserving property of the polar transform which we derived in Chapter 6 will be applicable right away; we will argue that polar codes are universal over sets of channels ordered by the symmetric convex ordering introduced in Definition 6.2.

- The subsequent section will study the complement problem: universal polar coding with channel knowledge at the encoder. Theorem 7.18 stated in the previous chapter will be the starting point of the study. We will show in Theorem 8.3 that polar codes do achieve the symmetric capacity of convex sets of channels and more generally of one-sided sets of channels introduced by [45].

- The final section of the trilogy will consider the problem of universal polar coding for both the encoding and the decoding procedures. The section will illustrate that while a condition such as stochastic degradation is not enough in general, a condition such as convexity is a better bet for this type of universality. In particular, we will conjecture that polar codes are universal over the set of BSCs.

These three sections will solely focus on the universality of Arıkan's original polar coding scheme. On the other hand, nothing prevents us from integrating different ideas from the coding theory literature into the polar coding framework to make polar codes more robust. In the subsequent two sections, we will propose two such modifications which will help us reach universal polar code designs over certain compound sets of channels:

- In Section 8.5, we will consider an approximation to the polar decoder's recursive computations of the likelihood ratios. Theorem 8.17 will show that polar codes using this approximation at the decoder side are robust over BSCs. Furthermore, we will provide simulation results that suggest the gap to capacity of the approximation is negligible. Combining the theoretical and experimental

analysis, we will conclude that mismatched polar codes can achieve rates very close to the true channel capacity for the class of BSCs.

- In the final section of this chapter, Theorem 8.23 will prove that universally attainable transmission rates can be achieved by polar coding over classes of B-DMCs satisfying certain mild conditions imposed on their structures by simply running the original polar decoder multiple times for the different channels until the generalized likelihood ratio test succeeds.

8.2 Universal Polar Coding with Channel Knowledge at the Decoder

Let $\mathcal{W}$ be a class of B-DMCs. In the problem of universal polar coding with channel knowledge at the decoder, the designer needs to find a single polar encoder, i.e., an appropriate information set, such that the transmitted codewords can be reliably communicated over any channel in the class when decoded at the receiver side by using the appropriate successive cancellation decoder adapted to the communication channel. Hence, this is equivalent to finding

$$\bigcap_{W \in \mathcal{W}} \mathcal{A}_N(W). \tag{8.6}$$

If the solution to (8.6) is given by the channel $V = \arg\min_{W \in \mathcal{W}} I(W)$, it is usually said that the polar code is universal over the class $\mathcal{W}$ for the considered communication scenario. An interesting question in that respect is to identify the conditions on the channel class structure which would lead to this type of universality.

The reader would remember that we did somewhat consider this problem before in Chapter 6. There, we proposed in Definition 6.2 the symmetric convex ordering as a novel partial ordering for B-DMCs and showed that the polar transform preserves this ordering. Thus, the symmetric convex ordering gives a structure leading to universality. Although it turned out that for symmetric channels this ordering coincides with stochastic degradation, by studying an example involving a Z-channel and BSCs, we proved that we obtain a strictly weaker partial ordering when at least one of the channels is asymmetric. In this example studied in Section 6.1, we saw that the information set of the polar code designed for the 'best' possible BSC (with the smallest crossover probability) which is smaller with respect to the symmetric convex ordering than the Z channel may be significantly larger than the set designed for the 'best' possible BSC which is stochastically degraded with respect to the Z-channel. The example also uncovered an advantage of the channel symmetrization operation for asymmetric channels before polarization; we illustrated that it matters during the design whether the channel is directly approximated by a degraded channel or the

channel is first symmetrized and then approximated by a degraded one. Beware that we do not try to achieve the capacity of the asymmetric channel, methods have been proposed in the literature to address this problem. We try to identify the right way to design polar codes over compound sets of mixed channels.

The symmetric convex ordering, despite being sufficient, is not a necessary condition for universality. Another recent work [46] which also studied this problem showed that polar codes are universal over sets of channels ordered in the less noisy ordering [47].

8.3 Universal Polar Coding with Channel Knowledge at the Encoder

The problem of universal polar decoding over a class of channels in scenarios where the encoder knows the actual communication channel can be seen as the complement of the problem of universal polar encoding with channel knowledge at the decoder. Here, we remove the assumption that the decoder uses the matched decoding metric during the decision procedure. The designer is required to design a single successive cancellation decoder which will be used for every channel in the class while the encoder and the decoder need to adapt the information set according to the actual communication channel and the possibly mismatched channel used in the design of the successive cancellation decoder. At this point, a comment is in order to avoid confusion. Note that for polar codes, the adjective 'universal' has been used in general to refer to the problem of universal polar encoding with channel knowledge at the decoder. Yet, 'universality' as studied by Blackwell et. al in [10], or as studied in [13], imposes stronger robustness than both of the described complementary problems. Now, be ready to meet the first low complexity $O(N \log N)$ decoder proved to be universal over one-sided sets of channels. (Hint: The answer should sound familiar!)

Theorem 8.3. *Given a class of one-sided B-DMCs $\mathcal{W}$, consider the polar successive cancellation decoder using the mismatched decoding rule for the channel*

$$V = \arg \min_{W \in \mathsf{cl}(\mathcal{W})} I(W),$$

and the class of polar codes with the information sets

$$\mathcal{A}_N^{\gamma}(W, V) = \big\{ i \in \{1, \ldots, N\} : \; D(W_N^{(i)}, V_N^{(i)}) \geq 1 - \gamma \big\},$$

where $W \in \mathcal{W}$, $N = 2^n$ with $n = 1, 2, \ldots$ is the block-length, and $\gamma \in (0, 1)$ is a desired threshold. Then, for any $R < I(V)$, one can select $\gamma \simeq 2^{-\sqrt{N}}$ and construct for all $W \in \mathcal{W}$ the information sets $\mathcal{A}_N^{\gamma}(W, V)$ of size at least as large as NR.

Moreover, the resulting decoding error probability $P_{\mathrm{e}}(W, V, \mathcal{A}_N^{\gamma}(W, V))$ over any channel $W \in \mathcal{W}$ of the corresponding polar code can be made arbitrarily small by taking $N \to \infty$. In that sense, the polar successive cancellation decoder is universally symmetric capacity achieving over one-sided sets of channels.

Proof. The one-sidedness of $\mathcal{W}$ ensures that for V chosen as in the hypothesis of the theorem, the relationship $I(W, V) \geq I(V)$ holds for any $W \in \mathcal{W}$. By Theorem 7.3 and Theorem 7.4, the claim on the construction of $\mathcal{A}_N^{\gamma}(W, V)$ follows. For a given N and $\gamma \in (0, 1)$, the mismatched decoding error probability over the channel $W \in \mathcal{W}$ of a polar code with a mismatched successive cancellation decoder operating with the parameters of the channel V and with information set $\mathcal{A}_N^{\gamma}(W, V)$ will be upper bounded by

$$P_{\mathrm{e}}(W, V, \mathcal{A}_N^{\gamma}(W, V)) \leq \sum_{i \in \mathcal{A}_N^{\gamma}(W,V)} P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}),$$

Taking $N \to \infty$, $\gamma \to 0$, and we get $P_{\mathrm{e}}(W, V, \mathcal{A}_N^{\gamma}(W, V)) \to 0$ as discussed in the proof of Theorem 7.18. We conclude that the described polar codes can achieve a rate of at least $I(V)$ over any $W \in \mathcal{W}$. □

8.4 Universal Polar Coding

So far, we did look to the problem form two opposite angles. We 'cheated' by allowing the decoder to know the channel and linked polar ordering to the notion of symmetric convex ordering. Then we allowed instead the encoder to know the channel, and using this assumption, we extended the results of Chapter 7 on the mismatched capacity of polar codes over the compound setting by using the notion of one-sidedness. Now, let us formulate the problem of designing a polar code for a compound set of channels as an optimization problem. In its most general form, the problem would require to identify the channel V^* which maximizes the number of common indices in the mismatched information sets constructed for every channel in the class:

$$V^* = \arg\max_{V \in \mathcal{W}} \Big| \bigcap_{W \in \mathcal{W}} \mathcal{A}_N(W, V) \Big|, \tag{8.7}$$

where $\mathcal{A}_N(W, V)$ is of the form (7.22).

In general, we know that polar codes are not universal over B-DMCs and do not expect to find a closed form solution to (8.7). However, for moderate block-lengths, solving this problem numerically (offline) seems plausible. Nevertheless, we will try to identify in this section some conditions on the channels for which this optimization problem has (or might have) a closed form solution. For instance, if the class of channels is one-sided or convex, the following theorem quantifies a

universally attainable transmission rate over the class.

Theorem 8.4. *Let $\mathcal{W}$ be a one-sided (convex) set of B-DMCs and*

$$V = \arg \min_{W \in \mathsf{cl}(\mathcal{W})} I(W).$$

Then, polar coding with information set $\bigcap_{W \in \mathcal{W}} \mathcal{A}_N(W, V)$ and with successive cancellation decoding with respect to V achieves the rate $\frac{1}{N} \left| \bigcap_{W \in \mathcal{W}} \mathcal{A}_N(W, V) \right|$ universally over $\mathcal{W}$.

To apply the preceding theorem to an arbitrary class of channels, one could extend the set of channels to a one-sided set if possible, or otherwise find its convex closure.

Let us consider an alternative for the simplest compound set consisting of two channels $\{W, V\}$. We would like to point a rather unexpected choice which would lead to a universal design (for both the encoder and the decoder) over sets of channels which satisfy the following polar ordering of the information sets: $\mathcal{A}_N(V) \subset \mathcal{A}_N(W)$. Instead of computing the solution in these cases, we could settle for a sub-optimal solution providing good indices for our class of two channels by finding the information set $\mathcal{A}_N(V, W)$ and decoding with the best channel and not the other way around. We could do so since $I(V^{s^n}, W^{s^n}) \leq I(V^{s^n})$ holds by Proposition 7.1, for all $s^n \in \{+, -\}^n$, and therefore we have $\mathcal{A}_N(V, W) \subset \mathcal{A}_N(V) \subset \mathcal{A}_N(W)$. By the results of Chapter 7, we know that with such a choice $|I(V, W)|^\dagger$ is an achievable rate. Certainly, this would be a very poor choice for most cases as the size of the set $\mathcal{A}_N(V, W)$ might be very small. Nevertheless, if the channels are very close, a universal polar code would be obtained at the price of a small rate loss from the symmetric capacity of the worst channel. To better explain this idea, we give a more concrete example.

Example 8.5. Suppose we have a model of the channel over which communication will take place, we do not know the exact transition probabilities, but we have some good estimates for their values. Say the model is a BSC with crossover probability $p = 0.1 \pm 0.02$. Modeling the class by $\mathcal{W} = \{BSC : p \in [0.08, 0, 12]\}$, we deduce that the polar code design with information set $\mathcal{A}_N(BSC(0.08), BSC(0, 12))$ achieves a rate 0.4569 over $\mathcal{W}$ using the polar decoder for the channel $BSC(0.08)$. Note that in this case $I(BSC) \in [0.4706, 0.5978]$, for all $BSC \in \mathcal{W}$.

8.4.1 Degradation is Not Sufficient

In this thesis we placed a special emphasis on the following subtlety: When defining the information set, the channel parameter used in the definition of the information set has to be adapted carefully to the context of communication keeping in mind that the transmitted data needs to be decoded reliably using the available decoding

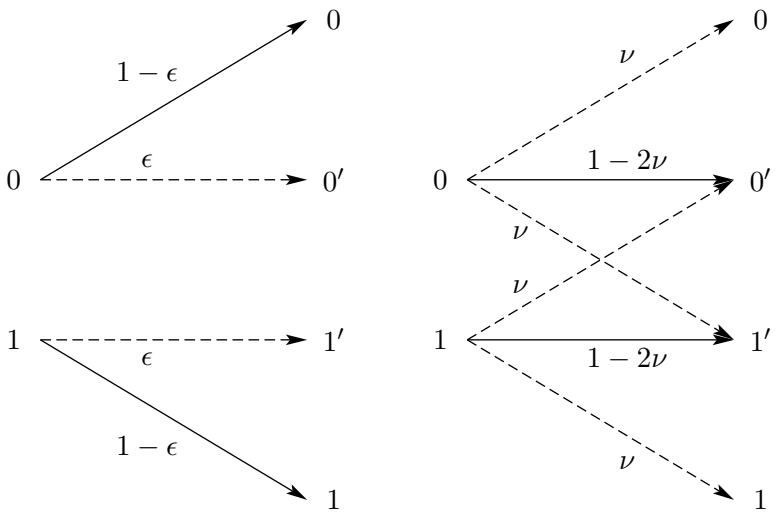

Figure 8.1: The channel W is shown on the left and the channel V on the right.

metric. In Chapter 7, this subtle requirement led to the definition of information sets of the form $\mathcal{A}_N(W, V)$. Moreover, we have just used these sets in Theorem 8.3 and Theorem 8.4. The question of practical interest we are interested in this section is to identify the classes of channels for which the relation $\mathcal{A}_N(V) \subset \mathcal{A}_N(W, V)$ holds, for all $W \in \mathcal{W}$. In such cases, a single polar code (encoder and decoder) designed for the 'worst' channel V can be used reliably over any channel in the class $\mathcal{W}$ and universality over these classes of channels can be achieved by the design channel.

Numerous works in the polar coding literature refer and make use of the universality of polar codes over stochastically degraded channels. Yet, this universality relies on the assumption that the decoder knows the communication channel, bypassing the need for the analysis carried in Chapter 7. One could start by hoping that polar codes are possibly universal over stochastically degraded channels even when the channel information is not available at the decoder. In this subsection, we give an (unfortunate) example of two channels W and V which are ordered by stochastic degradation, yet fails to satisfy the universality we hoped for, i.e., an example where $\mathcal{A}_N(W, V) \not\subset \mathcal{A}_N(V)$. Thus, although $\mathcal{A}_N(V) \subset \mathcal{A}_N(W)$ always holds for stochastically degraded channels [2], we conclude that $\mathcal{A}_N(V) \subset \mathcal{A}_N(W, V)$ does not necessarily hold.

Example 8.6. Let $W : \mathbb{F}_2 \to \mathcal{Y}$ and $V : \mathbb{F}_2 \to \mathcal{Y}$ be B-DMCs with $\mathcal{Y} : \{0, 0', 1', 1\}$. Suppose the channel transition probability matrices are of the form

$$W = \begin{bmatrix} 1-\epsilon & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & 1-\epsilon \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \nu & 1-2\nu & \nu & 0 \\ 0 & \nu & 1-2\nu & \nu \end{bmatrix},$$

for $\epsilon \in [0, 1]$ and $\nu \in [0, 1]$, see Figure 8.1. It is not difficult to see that the channel V can be stochastically degraded to obtain the channel W with a degrading channel

P. Moreover, basic derivations show that while both $I(W) = 1$ and $I(V) \approx 1$ hold, for $\epsilon = 0$ and $\nu \approx 0$, we still have $I(W, V) \approx 0$. As a last comment, note that the channels W and V are already 'trapped', i.e, the processes satisfy $I_\infty(W) = 1$, $I_\infty(V) = 1$, and $I_\infty(W, V) = 0$ a.s. As such, the polar transform will not be of use to improve the mismatched channel parameter.

8.4.2 Convex Sets May Be

The preceding example disclosed the fact that even the relatively strong stochastic degradation assumption on the structure of the compound class of channels do not accommodate in general the joint encoding and decoding universality of the polar code designed for the 'most noisy' member of the class. In the view of Theorem 8.3, this is not such an unexpected result; to ensure universality at the decoder, we imposed there the quite different notion of one-sided sets. These sets generalized the geometric properties of convex sets. This is why in this subsection we will play around convexity in place of degradation. To reduce the overhead, we will assume that the channels in the class satisfy the symmetry property of Definition 7.15.

In the previous chapter, we defined $P_{\mathrm{e}}(W, V, \mathcal{A}_N)$ as the best achievable block error probability over the channel W with mismatched polar decoding with respect to the channel V. In Proposition 7.16, we showed that

$$P_{\mathrm{e}}(W, V, \mathcal{A}_N) \leq \sum_{i \in \mathcal{A}_N} P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}),$$

Moreover, Proposition 7.21 simplified the computation of $P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)})$ for channels symmetrized by the same permutation: If W and V are B-DMCs symmetrized by the same permutation, then

$$P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}) = \sum_{y_1^N} W_N(y_1^N|0_1^N)\mathbf{H}\left(L_{V_N^{(i)}}(y_1^N, 0_1^{i-1})\right),$$

where

$$\mathbf{H}\left(L_{V_N^{(i)}}(y_1^N, 0_1^{i-1})\right) := \mathbf{1}\big\{L_{V_N^{(i)}}(y_1^N, 0_1^{i-1}) > 1\big\} + \frac{1}{2}\mathbf{1}\big\{L_{V_N^{(i)}}(y_1^N, 0_1^{i-1}) = 1\big\}.$$

For the sake of the analysis, we can thus assume the all zero sequence is transmitted over the channel. For shorthand notation, let $P_{\mathrm{e}N}^{(i)}(W, V) := P_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)})$ and we use once more $L_{V_N^{(i)}}(y_1^N) := L_{V_N^{(i)}}(y_1^N, 0_1^{i-1})$ (remark that, due to a possible mismatch, we keep in the subscript the channel with respect to which the likelihood ratio is computed). In this case, we know that the recursive likelihood ratio

computations indicated in (1.10) and (1.11) reduce to

$$L_{V_{2N}^{(2i-1)}}(y_1^{2N}) = \frac{L_{V_N^{(i)}}(y_1^N) + L_{V_N^{(i)}}(y_{N+1}^{2N})}{1 + L_{V_N^{(i)}}(y_1^N) L_{V_N^{(i)}}(y_{N+1}^{2N})},$$
$$L_{V_{2N}^{(2i)}}(y_1^{2N}) = L_{V_N^{(i)}}(y_1^N) L_{V_N^{(i)}}(y_{N+1}^{2N}).$$

Note that the likelihood ratios are of the form $f(L_{V_N^{(i)}}(y_1^N), L_{V_N^{(i)}}(y_{N+1}^{2N}))$, and so they are symmetric functions of their arguments. We will also use the notation

$$\mathbb{P}_W \left[L_{V_N^{(i)}}(Y_1^N) > 1 \right] := \sum_{y_1^N} W_N(y_1^N|0_1^N) \mathbf{1}\{ L_{V_N^{(i)}}(y_1^N) > 1 \},$$

and similar notations will hold for different sets within the indicator function.

The next theorem studies the one-step preservation properties for $P_{\mathrm{e}N}^{(i)}(W, V)$.

Theorem 8.7. *Let W and V be B-DMCs symmetrized by the same permutation. Suppose that the following conditions hold:*

A) $\mathbb{P}_V \left[L_{V_N^{(i)}}(Y_1^N) < 1 \right] \geq \mathbb{P}_V \left[L_{V_N^{(i)}}(Y_1^N) > 1 \right]$,

B) $P_{\mathrm{e}N}^{(i)}(W, V) - P_{\mathrm{e}N}^{(i)}(V) \leq 0$,

for a given $N = 2^n$ with $n = 0, 1, 2, \ldots$ and a given $i = 1, \ldots, N$. Then, the minus polar transform preserves these conditions. On the other hand, while the plus transform preserves condition A, condition B may not be preserved in general.

Before we prove the theorem, we introduce four propositions we will need in the proof. The proof of the propositions are given in Appendix 8.A.

Proposition 8.8. *For a symmetric B-DMC channel V such that the condition A of Theorem 8.7 holds for a given $i = 1, \ldots, N$, the polar transform preserves the inequality, i.e., for $j = 2i - 1, 2i$, we have*

$$\mathbb{P}_V \left[L_{V_{2N}^{(j)}}(Y_1^{2N}) < 1 \right] \geq \mathbb{P}_V \left[L_{V_{2N}^{(j)}}(Y_1^{2N}) > 1 \right].$$

Proposition 8.9. *For two B-DMCs W and V symmetrized by the same permutation, we have*

$$P_{\mathrm{e}N}^{(i)}(W, V) - P_{\mathrm{e}N}^{(i)}(V) = \sum_{y_1^N} \left[W_N(y_1^N|0_1^N) - V_N(y_1^N|0_1^N) \right] \times \sum_{y_{N+1}^{2N}} \left[W_N(y_{N+1}^{2N}|0_{N+1}^{2N}) + V_N(y_{N+1}^{2N}|0_{N+1}^{2N}) \right] \mathbf{H}\left(L_{V_{2N}^{(i)}}(y_1^{2N}) \right). \quad (8.8)$$

Proposition 8.10. *For two B-DMCs W and V symmetrized by the same permutation, the quantities $P_{\mathrm{e}2N}^{(i)}(W,V) - P_{\mathrm{e}2N}^{(i)}(V)$ can be recursively computed as*

$$P_{\mathrm{e}2N}^{(2i-1)}(W,V) - P_{\mathrm{e}2N}^{(2i-1)}(V) = \sum_{y_1^N} \left[W_N(y_1^N|0_1^N) - V_N(y_1^N|0_1^N)\right] \mathbf{H}\left(L_{V_N^{(i)}}(y_1^N)\right) J_N, \quad (8.9)$$

where

$$J_N := \left(\sum_{\substack{y_{N+1}^{2N}: \\ L_{V_N^{(i)}}(y_{N+1}^{2N})<1}} \left[W_N(y_{N+1}^{2N}|0_{N+1}^{2N}) + V_N(y_{N+1}^{2N}|0_{N+1}^{2N})\right] - \sum_{\substack{y_{N+1}^{2N}: \\ L_{V_N^{(i)}}(y_{N+1}^{2N})>1}} \left[W_N(y_{N+1}^{2N}|0_{N+1}^{2N}) + V_N(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \right), \quad (8.10)$$

and

$$P_{\mathrm{e}2N}^{(2i)}(W,V) - P_{\mathrm{e}2N}^{(2i)}(V) = \sum_{y_1^{2N}} \left[W_N(y_1^N|0_1^N) - V_N(y_1^N|0_1^N)\right] \times \left[W_N(y_{N+1}^{2N}|0_{N+1}^{2N}) + V_N(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \mathbf{H}\left(L_{V_N^{(i)}}(y_1^N) L_{V_N^{(i)}}(y_{N+1}^{2N})\right).$$

Proposition 8.11. *Assume W and V are B-DMCs such that the conditions A and B of Theorem 8.7 hold for a given $i = 1, \ldots, N$. Then,*

$$\mathbb{P}_W\left[L_{V_N^{(i)}}(Y_1^N) < 1\right] \geq \mathbb{P}_W\left[L_{V_N^{(i)}}(Y_1^N) > 1\right].$$

Proof of Theorem 8.7. We know the condition A) is preserved by Proposition 8.8. For the condition B$^-$), we have by Proposition 8.10:

$$P_{\mathrm{e}2N}^{(2i-1)}(W,V) - P_{\mathrm{e}2N}^{(2i-1)}(V) = \left[P_{\mathrm{e}N}^{(i)}(W,V) - P_{\mathrm{e}N}^{(i)}(V)\right] J_N. \quad (8.11)$$

Now, we claim that $J_N \geq 0$. Thus, $P_{\mathrm{e}2N}^{(2i-1)}(W,V) - P_{\mathrm{e}2N}^{(2i-1)}(V) \leq 0$ follows. To prove the claim, note that by equation (8.10), the constant J_N equals to

$$\mathbb{P}_W\left[L_{V_N^{(i)}}(Y_{N+1}^{2N}) < 1\right] + \mathbb{P}_V\left[L_{V_N^{(i)}}(Y_{N+1}^{2N}) < 1\right] - \mathbb{P}_W\left[L_{V_N^{(i)}}(Y_{N+1}^{2N}) > 1\right] - \mathbb{P}_V\left[L_{V_N^{(i)}}(Y_{N+1}^{2N}) > 1\right].$$

Then, the non-negativity of J_N follows by condition A and Proposition 8.11 which shows the conditions A and B imply

$$\mathbb{P}_W\left[L_{V_N^{(i)}}(Y_1^N) < 1\right] \geq \mathbb{P}_W\left[L_{V_N^{(i)}}(Y_1^N) > 1\right].$$

Finally, we give a counterexample for the condition B$^+$): Let W be a BSC of crossover probability 0.3 and V a symmetric B-DMC with $\mathcal{Y} = \{0, e, 1\}$ such that the likelihood ratios take the values $\{1/4, 1, 4\}$ with probabilities $V(y|0) = \{0.4, 0.5, 0.1\}$, respectively. One can check that although conditions A and B are satisfied for $N = 1$ and $i = 1$, condition B fails to hold after the plus transform ($N = 2$ and $i = 2$). □

Theorem 8.7 shows that we need to impose stronger constraints on the mismatched channel to be used if we want to ensure condition B is preserved under both the plus and minus polar transformations. Following this observation, we consider the mismatched Bhattacharyya distance $Z(W, V)$ we derived in Appendix 7.A from the 'mismatched E_0 function' in analogy to the Bhattacharyya distance $Z(W)$. For two B-DMCs W and V symmetrized by the same permutation, (7.27) reduces to $Z(W, V) = \sum_y W(y|0)\sqrt{L_V(y)}$. Next, we observe that $P_{\mathrm{e}}(W, V) := P_{\mathrm{e,\,ML}}(W, V) \leq Z(W, V)$. Moreover, similar to the matched case [2], one gets $Z_{2N}^{(2i)}(W, V) = Z_N^{(i)}(W, V)^2$ after applying the plus polar transform. From this latter, we get the following proposition.

Proposition 8.12.

$$Z_N^{(i)}(W, V) - Z_N^{(i)}(V) \leq 0 \quad \textit{if and only if} \quad Z_{2N}^{(2i)}(W, V) - Z_{2N}^{(2i)}(V) \leq 0.$$

In the next theorem, we explore the possible connection of such a result with Theorem 8.7.

Theorem 8.13. *Suppose that the channels W and V described in the hypothesis of Theorem 8.7 also satisfy the following conditions:*

$$P_{\mathrm{e}N}^{(i)}(W, V) - P_{\mathrm{e}N}^{(i)}(V) \leq 0 \quad \textit{if and only if} \quad Z_N^{(i)}(W, V) - Z_N^{(i)}(V) \leq 0, \quad (8.12)$$

Then, the condition B of Theorem 8.7 is preserved under both the plus and the minus polar transformations.

The theorem statement simply tells that if the Bhattacharyya upper bounds follow the same behavior as their $P_{\mathrm{e}N}^{(i)}$ counterparts; which can occur if for instance they are sufficiently tight for both the matched and mismatched error probabilities at any level, then as long as we design the polar code for a mismatched channel V such that $P_{\mathrm{e,\,ML}}(W, V) \leq P_{\mathrm{e,\,ML}}(V)$ is satisfied, we are safe to use the code over

the channel W. Although Theorem 8.13 provides a partial solution to the design problem, unfortunately it is non-constructive at this stage. We would need to study which channels could satisfy these type of constraints. The next proposition is presented as a first step in that direction.

Proposition 8.14. *Let $\mathcal{W}$ be a convex set of channels and $V = \arg\min_{W\in\mathcal{W}} I(W)$. Then,*

$$P_{\text{e, ML}}(W,V) \leq P_{\text{e, ML}}(V) \quad \textit{and} \quad Z(W,V) \leq Z(V). \tag{8.13}$$

Proof. The proposition will be proved by mimicking the proof of Proposition 8.1 given in [11, Proof of Corollary 6.10] and replacing the mutual information parameter first by $P_{\text{e, ML}}(W)$ and then by $Z(W)$. As the set $\{y \in \mathcal{Y} : \ L_W(y) > 1\}$ is convex, the indicator function of this set is log-concave. As a result, the convexity of the set of channels ensures $V = \arg\min_{W\in\mathcal{W}} P_{\text{e, ML}}(W)$. Hence, for every $W \in \mathcal{W}$ and for all $\alpha \in [0,1]$, we have $P_{\text{e, ML}}(\alpha W + (1-\alpha)V) \leq P_{\text{e}}(V)$. This implies

$$\lim_{\alpha\to 0} \frac{\partial}{\partial\alpha} P_{\text{e, ML}}(\alpha W + (1-\alpha)V) \leq 0.$$

Evaluating this derivative, $P_{\text{e, ML}}(W,V) \leq P_{\text{e, ML}}(V)$ is recovered. Similarly, by [2, Lemma 4], we know that $Z(W)$ is a concave function of the channel transition probabilities. Therefore, $V = \arg\min_{W\in\mathcal{W}} Z(W)$ holds, and for every $W \in \mathcal{W}$ and for all $\alpha \in [0,1]$, we have $Z(\alpha W + (1-\alpha)V) \leq Z(V)$. Hence,

$$\lim_{\alpha\to 0} \frac{\partial}{\partial\alpha} Z(\alpha W + (1-\alpha)V) \leq 0.$$

Evaluating this derivative, $Z(W,V) \leq Z(V)$ is recovered. □

8.4.3 Universal over BSCs?

Before proceeding onto the next section, let us see how the equations of the previous subsection 'simplify' for the class of BSCs.

Corollary 8.15. *For symmetric B-DMCs W and V, we have*

$$P_{\text{e}N}^{(i)}(W,V) - P_{\text{e}N}^{(i)}(V) = \sum_{y_1^N} \left[W(y_1|0) - V(y_1|0)\right]\left[W(y_2|0) + V(y_2|0)\right] \times$$
$$\prod_{i=2}^{n} \left(W(y_{2^{i-1}+1}|0)\dots W(y_{2^i}|0) + V(y_{2^{i-1}+1}|0)\dots V(y_{2^i}|0)\right) \mathbf{H}\left(L_{V_N^{(i)}}(y_1^N)\right).$$

Proof. The proof follows from the fact that $W_N(y_1^N|0_1^N) = \prod_{i=1}^N W(y_i|0)$ holds by (1.1) and (1.2) and by induction on Proposition 8.9. □

Let W and V be two BSCs with crossover probabilities $p_W \leq p_V \leq 0.5$. Let $L_V(1) := L \geq 1$. Then, $L_V(0) = 1/L$. Now, we define the following partial sums:

$$F_N^{(i)}(L_V(y_1)) := \sum_{y_2^N} \left[W(y_2|0) + V(y_2|0)\right] \left[W(y_3|0)W(y_4|0) + V(y_3|0)V(y_4|0)\right] \times \prod_{i=2}^{n} \left(W(y_{2^{i-1}+1}|0) \ldots W(y_{2^i}|0) + V(y_{2^{i-1}+1}|0) \ldots V(y_{2^i}|0)\right) \mathbf{H}\left(L_{V_N^{(i)}}(y_1^N)\right),$$

and observe by Corollary 8.15 that we have

$$P_{\mathrm{e}N}^{(i)}(W,V) - P_{\mathrm{e}N}^{(i)}(V) = \underbrace{(p_V - p_W)}_{\geq 0} \left[F_N^{(i)}(1/L) - F_N^{(i)}(L)\right].$$

Conjecture 8.16. We conjecture that for any $N = 2^n$ with $n = 1, 2, \ldots$ and for any $i \in \mathcal{A}_N(V)$,

$$F_N^{(i)}(1/L) \leq F_N^{(i)}(L).$$

Under this conjecture, polar codes are universal over BSCs with crossover probabilities in $[0, 0.5]$.

8.5 Practically Perfect in Every BSC

Suppose that upon reception of the channel outputs y_1^N, instead of proceeding with the exact computations using the recursive formulas described in (1.10) and (1.11), the polar decoder uses the following $\min$ *approximation* to the computations:

$$\widetilde{L}_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}) := \exp\left(-\mathrm{sign}(\ell_1 * \ell_2) \min\{|\ell_1|, |\ell_2|\}\right), \tag{8.14}$$

$$\widetilde{L}_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) := \begin{cases} \exp\left(\ell_1 + \ell_2\right), & \text{if } \hat{u}_{2i-1} = 0 \\ \exp\left(\ell_2 - \ell_1\right), & \text{if } \hat{u}_{2i-1} = 1 \end{cases}, \tag{8.15}$$

where $\ell_1 := \log \widetilde{L}_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})$ and $\ell_2 := \log \widetilde{L}_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2})$. The root elements are, as before, the initial likelihood ratios $\widetilde{L}_0(y_i) := L(y_i)$ computed for $i = 1, \ldots, N$ by using the law of the design channel. Thus, instead of the exact values $L_N^{(i)}(y_1^N, \hat{u}_1^{i-1})$, the polar decoder bases its decisions on the values of $\widetilde{L}_N^{(i)}(y_1^N, \hat{u}_1^{i-1})$, for all $i \in \mathcal{A}_N$.

The above approximation to the minus polar transform, also known as the *min-sum approximation* in the coding theory literature, was introduced in the context of polar coding in [48] for its suitability in hardware implementations. In this section, we have a completely different purpose in mind for introducing this approximation.

We take yet another approach and study the robustness of polar codes over compound BSCs with this 'approximate' polar decoder. In the next theorem, we show that replacing the minus polar transform with the $\min$ approximation, a certain ordering of the likelihood ratios of the synthesized channels $W_N^{(i)}$ and $V_N^{(i)}$ is obtained for each $i = 1, \ldots, N$.

Theorem 8.17. *Given a BSC of parameter $p < 0.5$, define $L = (1-p)/p > 1$. Assume y_1^N is observed at the channel output after transmission according to the polar encoding rule of an input. Let $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1})$ denote the approximate value obtained at the decoder for the likelihood ratio of the i-th synthetic channel by using* (8.14) *and* (8.15)*. Then,*

1. *$\tilde{f}(L, y_1^N, \hat{u}_1^{i-1})$ is increasing in L if and only if $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1}) > 1$.*
2. *$\tilde{f}(L, y_1^N, \hat{u}_1^{i-1})$ is decreasing in L if and only if $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1}) < 1$.*
3. *When $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1}) = 1$, either $L = 1$, or $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1})$ does not depend on L.*

Proof. The particular structure of $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1})$ leads straightforwardly to the result: As we assumed the minus polar transform is replaced with the $\min$ approximation, any likelihood ratio will be of the form $\tilde{f}(L, y_1^N, \hat{u}_1^{i-1}) = L^m$ for some $m \in \mathbb{Z}$ which depends on y_1^N, $\hat{u}_1^{i-1}$. This proves the claim. □

Remark 8.18. Lemma 8.24 in Appendix 8.B shows that if the statements 1), 2), and 3) of the theorem hold for a certain level n, then they continue to hold when the original minus polar transform given by (1.10) is applied to obtain a likelihood ratio at level $n+1$. Unfortunately, the proof by induction does not work as in this case the plus transform does not preserve properties 1), 2), and 3).

Let us now study the implications of the resulting likelihood ratio ordering on the robustness of the 'approximate' polar decoder. For the rest of this section, let W and V denote two BSCs of crossover probabilities p_W and p_V, respectively. We assume that $p_W, p_V < 0.5$ for simplicity. By Theorem 8.17, we deduce that the decoder estimate for a given output realization will be identical whether the computations are performed with respect to the likelihood ratios of the channel W or V, as long as $p_W, p_V < 0.5$. In particular, when $p_W \leq p_V$, beside the information set ordering $\mathcal{A}_N(V) \subset \mathcal{A}_N(W)$ due to channel degradation, the approximated likelihood ratios will also be ordered as follows:

$$1 \leq \tilde{f}(L_V, y_1^N, \hat{u}_1^{i-1}) \leq \tilde{f}(L_W, y_1^N, \hat{u}_1^{i-1}),$$
$$\text{or} \quad \tilde{f}(L_W, y_1^N, \hat{u}_1^{i-1}) \leq \tilde{f}(L_V, y_1^N, \hat{u}_1^{i-1}) \leq 1,$$

for any output realization y_1^N. Indeed, the approximate likelihood ratios of the BSC having a larger crossover probability will be closer to 1.

The importance of this approximation is at least two-fold. One advantage brought is its practical relevance: it is proposed in [48] to facilitate hardware implementations from a computational standpoint. The second advantage we found is the provided robustness: for any given output realization the decoder estimate is identical independent of the channel law, as long as $p_W, p_V < 0.5$ (or by symmetry $p_W, p_V > 0.5$).

Similar to the exact case, we define $\widetilde{P}_{\mathrm{e}}(W, V, \mathcal{A}_N)$ as the resulting average error probability with the approximation at the decoder. Then, similar to Proposition 7.16 and Proposition 7.21, one can show that

$$\widetilde{P}_{\mathrm{e}}(W, V, \mathcal{A}_N) \leq \sum_{i \in \mathcal{A}_N} \widetilde{P}_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}),$$

where

$$\widetilde{P}_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)}) := \sum_{y_1^N} W_N(y_1^N|0_1^N)\mathbf{1}\{\tilde{f}(L, y_1^N, 0_1^{i-1}) > 1\} + \frac{1}{2}\sum_{y_1^N} W_N(y_1^N|0_1^N)\mathbf{1}\{\tilde{f}(L, y_1^N, 0_1^{i-1}) = 1\}. \quad (8.16)$$

Note that the exact and the approximate likelihood ratio random variables of the synthetic channels will be 'symmetrized by the same permutation'. The robustness shown in Theorem 8.17 ensures the following result holds.

Corollary 8.19. *By Theorem 8.17, we have the following ordering:*

$$\widetilde{P}_{\mathrm{e}}(W, \mathcal{A}_N) := \widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A}_N) = \widetilde{P}_{\mathrm{e}}(W, V, \mathcal{A}_N).$$

Figure 8.2 illustrates the result of Corollary 8.19. As a result, we do not need to worry about the robustness against channel parameter variations of the min-sum variant of the polar decoder. However, the question that needs to be addressed is whether the performance of the polar decoder implementing the approximation is sufficiently good and close to the original one.

Simulation results show that the matched and the mismatched successive cancellation decoders both using the original recursion of the polar transform can produce different estimates of the input sequence for an observed output realization y_1^N. Therefore, counter-examples to the ordering of the likelihood ratios implied by Theorem 8.17 exist when the original polar transform is used. This is also in accordance with Remark 8.18. Nevertheless, the performance degradation due to the approximation

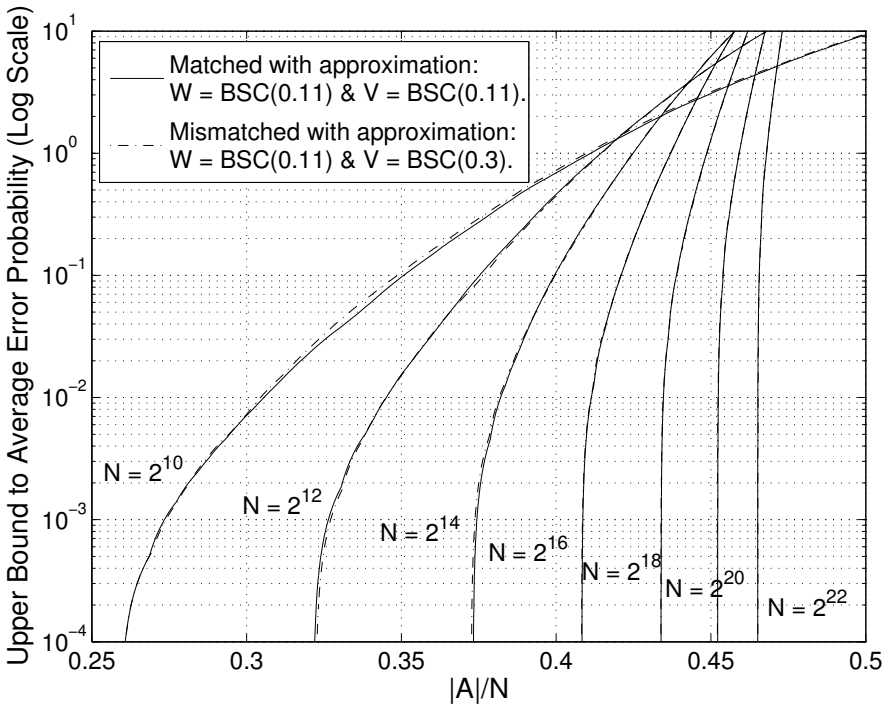

Figure 8.2: Rate vs. upper bound to $\widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A})$, $\widetilde{P}_{\mathrm{e}}(W, V, \mathcal{A})$.

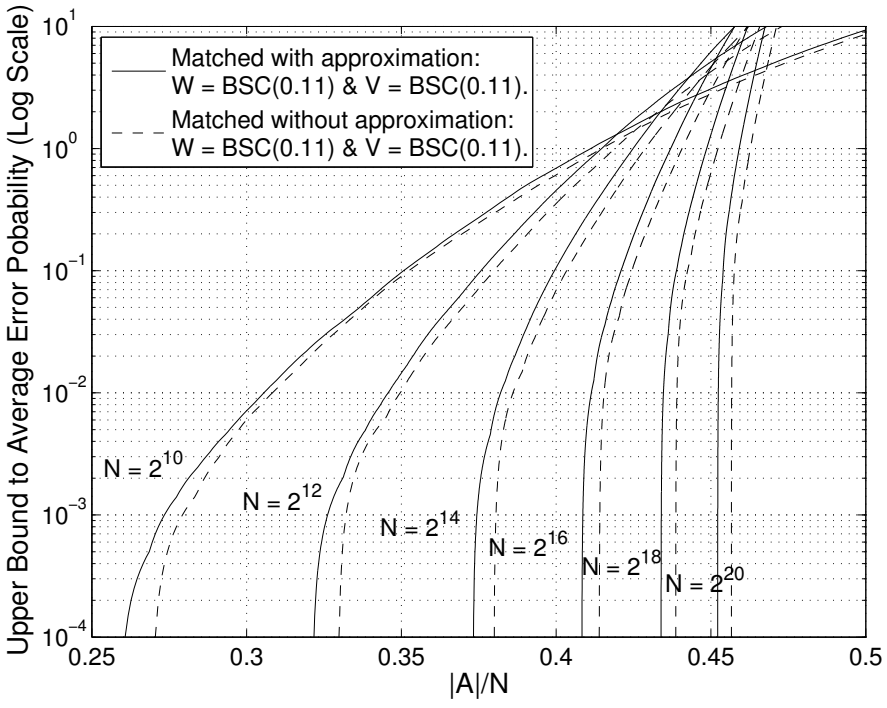

Figure 8.3: Rate vs. upper bound to $\widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A})$, $\widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A}')$.

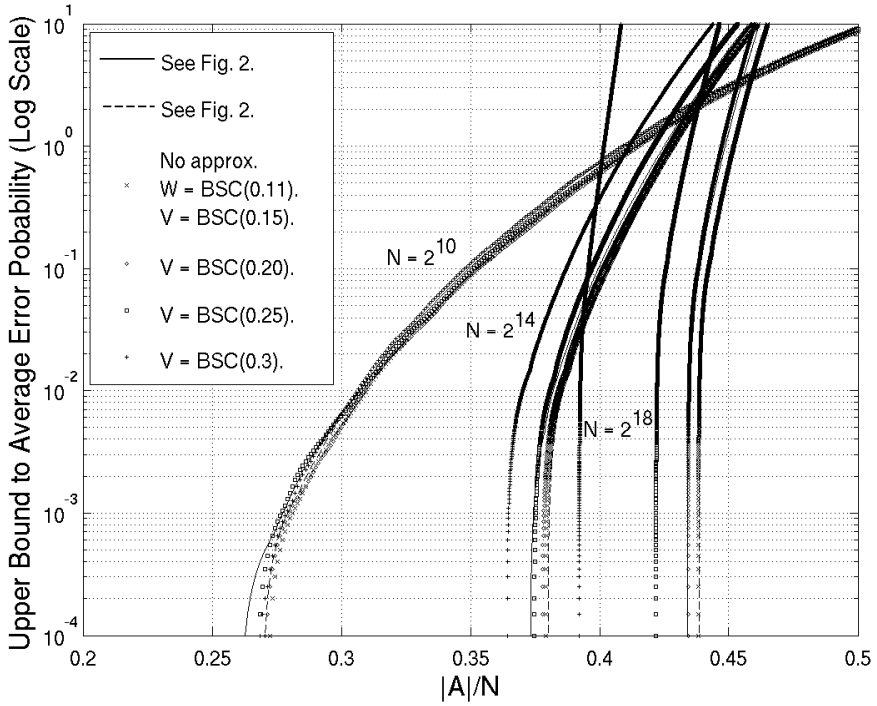

Figure 8.4: Rate vs. upper bound to $\widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A})$, $\widetilde{P}_{\mathrm{e}}(W, W, \mathcal{A}')$, and $\widetilde{P}_{\mathrm{e}}(W, V, \mathcal{A}'')$.

is claimed to be negligible in [48]. Our simulation results shown in Figure 8.3 also confirm this statement. We expect the error process with the approximation to have similar convergence properties as the matched process without the approximation. In Appendix 8.C, we address this problem in a theoretical framework, and obtain some partial results.

In addition, we carried simulations to analyze the performance degradation introduced by a mismatched design when the exact recursion of the polar transform is used. Figure 8.4 shows the results at three different block-lengths with multiple mismatched parameters. We observe that, as opposed to the previous case, the impact of the mismatch on the upper bounds is strongly dependent on the mismatched parameter. Furthermore, larger block-lengths seem to amplify this dependence. We thus refrain from drawing conclusions on the robustness of the original polar transform only based on simulations. (No contradiction to the achievability of $I(W, V)$.)

In the light of the above results, we propose for BSCs a polar coding scheme similar to the matched scenario [2]. For $N = 2^n$, we could design a polar code by setting the information set as

$$\widetilde{\mathcal{A}}_N^{\gamma}(W) := \big\{i \in \{1, \ldots, N\} : \ \widetilde{P}_{\mathrm{e,\,ML}}(W_N^{(i)}) \leq \gamma\big\},$$

for a desired threshold $\gamma \in (0, 1)$, and decoding according to the channel law of any

BSC V with crossover probability < 0.5 by using the approximation. We remark that, by Corollary 8.19, the information set can equivalently be constructed by estimating the parameters $\widetilde{P}_{\mathrm{e,\,ML}}(W_N^{(i)}, V_N^{(i)})$ using statistical methods (transmitting the all zero sequence and estimating via the polar decoder with the approximation using the metric of the channel V). Finally, selecting the threshold γ exponential in the square root of the block-length will ensure $\widetilde{P}_{\mathrm{e}}(W, \tilde{\mathcal{A}}_N^{\gamma}(W)) \to 0$ as $N \to \infty$ [30], and the simulations suggest that the number of indices that remain in $\tilde{\mathcal{A}}_N^{\gamma}(W)$ for such a choice of γ results in negligible capacity loss.

8.6 Generalized Likelihood Ratio Test

The research conducted so far revealed that strong assumptions must be imposed on the structure of the class of channels so that a member provides a universal polar code design for both the encoding and the decoding procedures. Hopefully, nothing prevents us from trying to modify and adapt the scheme in order to overcome the encountered difficulties.

In this section, we explore how the idea of the *generalized likelihood ratio test* (GRLT) can be applied for decoding polar codes over certain compound sets of channels. We will show that non-universality can be surmounted over these compound sets by allocating more resources to the decoding task. To begin with, we need two definitions from the theory of types/typicality.

Definition 8.20. [49] The empirical distribution $\widehat{P}$ of a sequence of length N is called the N*-type* of the sequence. The set of all sequences of N-type $\widehat{P}$ is denoted by $\mathcal{T}_{\widehat{P}}^N$ and can be described as:

$$\mathcal{T}_{\widehat{P}}^N := \left\{ x_1^N \in \mathcal{X}^N : \ \widehat{P}(x) = \frac{1}{N} \#\left\{ i : \ x_i = x \right\}, \forall x \in \mathcal{X} \right\}. \tag{8.17}$$

Definition 8.21. The *strongly typical set* $\mathcal{T}_P^{N,\varepsilon}$ of sequences $x_1^N \in \mathcal{X}^N$ with respect to the distribution $P(x)$ can be defined as

$$\mathcal{T}_P^{N,\varepsilon} = \left\{ \begin{array}{l} x_1^N \in \mathcal{X}^N : \\ \forall x \in \mathcal{X}, \quad \frac{1}{N} \#\left\{ i : \ x_i = x \right\} = \left\{ \begin{array}{ll} 0, & \text{if } \ P(x) = 0 \\ P(x) \pm \varepsilon, & \text{otherwise} \end{array} \right. \end{array} \right\},$$

where $\varepsilon > 0$ is an arbitrarily small number.

Note that these definitions can be directly applied to joint types. So, if P is the input distribution of a DMC $W : \mathcal{X} \to \mathcal{Y}$, then $T_{PW}^{N,\varepsilon}$ is the union of all type classes $T_{\widehat{Q}}^N$ where $\widehat{Q}(x, y)$ is in the ε-neighborhood of $PW(x, y)$. The next lemma identifies some mild conditions we will impose on the channels to establish the main result.

Lemma 8.22. *Let $\{W, V\}$ be a class of two DMCs with uniform input distribution P_{unif}. Suppose the channel W is better than the channel V in the sense that*

$$\sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\frac{1}{2}W(y|x)\log W(y|x) > \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\frac{1}{2}V(y|x)\log W(y|x), \tag{8.18}$$

$$\sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\frac{1}{2}W(y|x)\log V(y|x) > \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\frac{1}{2}V(y|x)\log V(y|x). \tag{8.19}$$

Then, for small enough ε, whenever $(\mathbf{x}, \mathbf{y}) \in T^{N,\varepsilon}_{P_{\text{unif}}W}$ and $(\mathbf{x}', \mathbf{y}) \in T^{N,\varepsilon}_{P_{\text{unif}}V}$, we have

$$W(\mathbf{y}|\mathbf{x}) > W(\mathbf{y}|\mathbf{x}') \quad \textit{and} \quad V(\mathbf{y}|\mathbf{x}) > V(\mathbf{y}|\mathbf{x}').$$

Proof. Under the conditions of the lemma

$$\frac{1}{N}\log W(\mathbf{y}|\mathbf{x}) = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\widehat{Q}_1(x,y)\log W(y|x), \tag{8.20}$$

$$\frac{1}{N}\log W(\mathbf{y}|\mathbf{x}') = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}}\widehat{Q}_2(x,y)\log W(y|x), \tag{8.21}$$

for some joint types $\widehat{Q}_1(x,y)$ in the ε-neighborhood of $PW(x,y)$ and $\widehat{Q}_2(x,y)$ in the ε-neighborhood of $PV(x,y)$. Note that for $\widehat{Q}_1 = P_{\text{unif}}W$ and $\widehat{Q}_2 = P_{\text{unif}}V$, the inequality in (8.18) ensures $W(\mathbf{y}|\mathbf{x}) > W(\mathbf{y}|\mathbf{x}')$; the continuity of (8.20) and (8.21) in $\widehat{Q}_1$ and $\widehat{Q}_2$ then lets us conclude that the inequality holds for small enough ε. Similarly (8.19) implies the claimed inequality for the channel V. □

Let W and V be two B-DMCs such that (8.18) and (8.19) hold. Suppose a polar code with information set $\mathcal{A}_N(V)\bigcap\mathcal{A}_N(V)$ is designed to ensure information is transmitted through the good synthetic channels regardless of the true channel. Assume the input message $\mathbf{u}$ is encoded with the polar encoding rule into the input sequence $\mathbf{x} = \mathbf{u}G_N$, which is then transmitted through the unknown channel. Upon reception of an output $\mathbf{y}$, an idea could be to run the polar successive cancellation decoder twice using the metrics of W and V. Let the estimates of the two runs be $\hat{\mathbf{u}}[W]$ and $\hat{\mathbf{u}}[V]$, respectively. Using these estimates, we can also estimate the channel inputs by applying

$$\hat{\mathbf{x}}[W] = \hat{\mathbf{u}}[W]G_N \quad \text{and} \quad \hat{\mathbf{x}}[V] = \hat{\mathbf{u}}[V]G_N.$$

Now, if the true channel was W, we would have $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}W}$ with high probability. As we also know by [2, Theorem 3] that the polar successive cancellation decoder would succeed with high probability in estimating the transmitted message correctly, we would expect as well $(\hat{\mathbf{x}}[W], \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}W}$ with high probability. As a result, the decoder can check whether $(\hat{\mathbf{x}}[W], \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}W}$, and if so, declare

$\hat{\mathbf{u}}[W]$ as the decoder's estimate of the transmitted message. However, the previous argument can also be rephrased for the channel V. Since the decoder does not know the true channel, we need to resolve the conflict which will arise in case both conditions $(\hat{\mathbf{x}}[W], \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}W}$ and $(\hat{\mathbf{x}}[V], \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}V}$ prevail or fail. Suppose first both relations hold. Then by Lemma 8.22, we have

$$W(\mathbf{y}|\hat{\mathbf{x}}[W]) \geq W(\mathbf{y}|\hat{\mathbf{x}}[V]) \quad \text{and} \quad V(\mathbf{y}|\hat{\mathbf{x}}[W]) \geq V(\mathbf{y}|\hat{\mathbf{x}}[V]).$$

This result tells us that, regardless of the true channel, the estimate $\hat{\mathbf{u}}[W]$ of the successive cancellation decoder using the metric of the channel W is always more likely than the estimate $\hat{\mathbf{u}}[V]$ of the decoder using the metric of the channel V. Therefore, to minimize the error probability, the decoder should declare that the message $\hat{\mathbf{u}}[W]$ was transmitted. Finally, if both conditions fail, the decoder can simply declare an erasure. The analysis in [2, Theorem 3] ensures this event has a vanishing error probability.

The next theorem generalizes the idea presented above to finite classes of channels $\mathcal{W}$, where $L = |\mathcal{W}|$, and defines the GRLT polar decoder.

Theorem 8.23. *Let $\mathcal{W} = W_1 \dots W_L$ be a finite class of B-DMCs such that*

$$\sum_{y\in\mathcal{Y}} \sum_{x\in\mathbb{F}_2} \frac{1}{2} W_k(y|x) \log W_k(y|x) > \sum_{y\in\mathcal{Y}} \sum_{x\in\mathbb{F}_2} \frac{1}{2} W_{k+1}(y|x) \log W_k(y|x), \tag{8.22}$$

$$\sum_{y\in\mathcal{Y}} \sum_{x\in\mathbb{F}_2} \frac{1}{2} W_k(y|x) \log W_{k+1}(y|x) > \sum_{y\in\mathcal{Y}} \sum_{x\in\mathbb{F}_2} \frac{1}{2} W_{k+1}(y|x) \log W_{k+1}(y|x) \tag{8.23}$$

hold for all $k = 1, \dots, L-1$. Then, the information set of the polar code designed for the channel

$$\bigcap_k \mathcal{A}_N(W_k)$$

achieves the rate $\frac{1}{N}\left|\bigcap_k \mathcal{A}_N(W_k)\right|$ universally over $\mathcal{W}$ with the GRLT polar decoder implementing the following decoding algorithm:

1. *Set $k = 1$,*
2. *Decode the received output sequence $\mathbf{y}$ to form the estimate $\hat{\mathbf{u}}[W_k]$ by running the polar successive cancellation decoder using the metric of the channel W_k,*
3. *Construct $\hat{\mathbf{x}}[W_k] = \hat{\mathbf{u}}[W_k]G_N$,*
4. *If $(\hat{\mathbf{x}}[W_k], \mathbf{y}) \in \mathcal{T}^{N,\varepsilon}_{P_{\text{unif}}W_k}$ is true, declare $\hat{\mathbf{u}}[W_k]$ and exit. If $k = L$, declare an erasure and exit. Otherwise, increment k and go to step 2.*

This serial implementation requires at most $O(LN \log N)$ time complexity and $O(N)$ space complexity.

The GRLT polar decoder algorithm can also be implemented to run the successive cancellation decoders for the different channels in parallel. Such an implementation would require instead $O(N \log N)$ time complexity, but $O(LN)$ space complexity.

Appendix

This Appendix consists of three parts. In the first part, we prove Propositions 8.8, 8.9, 8.10, and 8.11. Then, we prove Lemma 8.24 in the second part. Finally, in the third part, we present an analysis addressing the gap to capacity of the min-sum approximation of the polar transform.

8.A Proofs of Propositions 8.8, 8.9, 8.10, and 8.11

Proof of Proposition 8.8. We let $L_{V_N^{(i)}}(y_1^N) = L_1$, $L_{V_N^{(i)}}(y_{N+1}^{2N}) = L_2$ and omit the subscript in $\mathbb{P}_V$ for simplicity. By symmetry in the construction of polar codes, we have $\mathbb{P}[L_1 < 1] = \mathbb{P}[L_2 < 1]$. Then, the claim follows by Proposition 5.7. □

Proof of Proposition 8.9. We develop the right hand side of (8.8):

$$\begin{aligned}
&\sum_{y_1^{2N}} \left[W(y_1^{2N}|0_1^{2N}) - V(y_1^{2N}|0_1^{2N})\right] \mathbf{H}\left(L_{V_{2N}^{(i)}}(y_1^{2N})\right) \\
&+\sum_{y_1^{2N}} W(y_1^N|0_1^N) V(y_{N+1}^{2N}|0_{N+1}^{2N}) \mathbf{H}\left(f(L_{V_N^{(i)}}(y_1^N), L_{V_N^{(i)}}(y_{N+1}^{2N}))\right) \\
&-\sum_{y_1^{2N}} W(y_{N+1}^{2N}|0_{N+1}^{2N}) V(y_1^N|0_1^N) \mathbf{H}\left(f(L_{V_N^{(i)}}(y_{N+1}^{2N}), L_{V_N^{(i)}}(y_1^N))\right) \\
= \quad & P_{\mathrm{e}N}^{(i)}(W,V) - P_{\mathrm{e}N}^{(i)}(V),
\end{aligned}$$

where we used the symmetry of the likelihood ratio functions in their arguments. □

Proposition 8.10. First we note that for B-DMCs W and V symmetrized by the same permutation, we have by Proposition (8.9)

$$\begin{aligned}
P_{\mathrm{e}2N}^{(i)}(W,V) - P_{\mathrm{e}2N}^{(i)}(V) = &\sum_{y_1^N} \left[W(y_1^N|0_1^N) - V(y_1^N|0_1^N)\right] \times \\
&\left(\sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N}|0_{N+1}^{2N}) + V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \mathbf{H}\left(L_{V_{2N}^{(i)}}(y_1^{2N})\right)\right), \quad (8.24)
\end{aligned}$$

For simplicity, we let $L_{V_N^{(i)}}(y_1^N) = L_1$ and $L_{V_N^{(i)}}(y_{N+1}^{2N}) = L_2$. We observe that the

recursion for the minus polar transform is given by

$$\mathbf{H}\left(\frac{L_1+L_2}{1+L_1L_2}\right)=\begin{cases}\frac{1}{2}, & \text{if} \quad L_1=1, & & \\ & \text{or} \quad L_2=1 & & \\ 1, & \text{if} \quad L_1<1 & \text{and} & L_2>1, \\ & \text{or} \quad L_1>1 & \text{and} & L_2<1 \\ 0, & \text{if} \quad L_1<1 & \text{and} & L_2<1, \\ & \text{or} \quad L_1>1 & \text{and} & L_2>1 \end{cases}.$$

We define the function $\overline{\mathbf{H}}\left(L_1\right):=\mathbf{1}\{L_1<1\}+\frac{1}{2}\mathbf{1}\{L_1=1\}$. Therefore, after we apply the minus polar transform, we have

$$\begin{aligned}
&P_{\mathrm{e}2N}^{(2i-1)}(W,V)-P_{\mathrm{e}2N}^{(2i-1)}(V) \\
=&\sum_{\substack{y_1^N: \\ L_1=1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times 1 \\
+&\sum_{\substack{y_1^N: \\ L_1>1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times \\
&\qquad\qquad\sum_{\substack{y_{N+1}^{2N}: \\ L_2\leq 1}}\left[W(y_{N+1}^{2N}|0_{N+1}^{2N})+V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \\
+&\sum_{\substack{y_1^N: \\ L_1<1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times \\
&\qquad\qquad\sum_{\substack{y_{N+1}^{2N}: \\ L_2\geq 1}}\left[W(y_{N+1}^{2N}|0_{N+1}^{2N})+V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \\
=&\sum_{\substack{y_1^N: \\ L_1=1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times 1 \\
+&\sum_{\substack{y_1^N: \\ L_1>1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times \\
&\qquad\qquad\sum_{y_{N+1}^{2N}}\left[W(y_{N+1}^{2N}|0_{N+1}^{2N})+V(y_{N+1}^{2N}|0_{N+1}^{2N})\right]\overline{\mathbf{H}}\left(L_2\right) \\
+&\sum_{\substack{y_1^N: \\ L_1<1}}\left[W(y_1^N|0_1^N)-V(y_1^N|0_1^N)\right]\times \\
&\qquad\qquad\sum_{y_{N+1}^{2N}}\left[W(y_{N+1}^{2N}|0_{N+1}^{2N})+V(y_{N+1}^{2N}|0_{N+1}^{2N})\right]\mathbf{H}\left(L_2\right).
\end{aligned}$$

By substituting $\overline{\mathbf{H}}(L_2) = 1 - \mathbf{H}(L_2)$ and regrouping the terms, we obtain

$$\begin{aligned}
&P_{\mathrm{e}2N}^{(2i-1)}(W,V) - P_{\mathrm{e}2N}^{(2i-1)}(V) \\
=&\sum_{y_1^N} \left[W(y_1^N|0_1^N) - V(y_1^N|0_1^N)\right] 2\mathbf{H}(L_1) \\
&+ \sum_{y_1^N} \left[W(y_1^N|0_1^N) - V(y_1^N|0_1^N)\right] [1 - 2\mathbf{H}(L_1)] \times \\
&\qquad\qquad \sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N}|0_{N+1}^{2N}) + V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] \mathbf{H}(L_2),
\end{aligned}$$

where we used the fact that $1 - 2\mathbf{H}(L_1) = \mathbf{1}\{L_1 < 1\} - \mathbf{1}\{L_1 > 1\}$. Now, note that the term in the second summation with the 1 sums to 0. Hence, we get

$$\begin{aligned}
&P_{\mathrm{e}2N}^{(2i-1)}(W,V) - P_{\mathrm{e}2N}^{(2i-1)}(V) \\
=&\sum_{y_1^N} \left[W(y_1^N|0_1^N) - V(y_1^N|0_1^N)\right] \mathbf{H}(L_1) \times \\
&\qquad\qquad \left[2 - \sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N}|0_{N+1}^{2N}) + V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] 2\mathbf{H}(L_2)\right] \\
=&\sum_{y_1^N} \left[W(y_1^N|0_1^N) - V(y_1^N|0_1^N)\right] \mathbf{H}(L_1) \times \\
&\qquad\qquad \sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N}|0_{N+1}^{2N}) + V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] [1 - 2\mathbf{H}(L_2)].
\end{aligned}$$

We recover (8.9) upon noticing J_N defined in (8.10) equals

$$\sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N}|0_{N+1}^{2N}) + V(y_{N+1}^{2N}|0_{N+1}^{2N})\right] [1 - 2\mathbf{H}(L_2)],$$

as $1 - 2\mathbf{H}(L_2) = \mathbf{1}\{L_2 < 1\} - \mathbf{1}\{L_2 > 1\}$. This proves the claim for the minus transform. The claim for the plus transform can be obtained directly by the expression given in (8.24). □

Proof of Proposition 8.11. We have

$$\begin{aligned}
&\mathbb{P}_W\left[L(y_1^N) > 1\right] + \frac{1}{2}\mathbb{P}_W\left[L(y_1^N) = 1\right] \\
&\qquad - \mathbb{P}_V\left[L(y_1^N) > 1\right] - \frac{1}{2}\mathbb{P}_V\left[L(y_1^N) = 1\right] \\
=&\mathbb{P}_V\left[L(y_1^N) < 1\right] + \frac{1}{2}\mathbb{P}_V\left[L(y_1^N) = 1\right] \\
&\qquad - \mathbb{P}_W\left[L(y_1^N) < 1\right] - \frac{1}{2}\mathbb{P}_W\left[L(y_1^N) = 1\right] \leq 0,
\end{aligned}$$

where the negativity follows by condition B. Therefore, adding both sides gives

$$\mathbb{P}_W\left[L(y_1^N) > 1\right] - \mathbb{P}_V\left[L(y_1^N) > 1\right] + \mathbb{P}_V\left[L(y_1^N) < 1\right] - \mathbb{P}_W\left[L(y_1^N) < 1\right] \leq 0.$$

$$\implies \mathbb{P}_W\left[L(y_1^N) < 1\right] - \mathbb{P}_W\left[L(y_1^N) > 1\right] \geq \mathbb{P}_V\left[L(y_1^N) < 1\right] - \mathbb{P}_V\left[L(y_1^N) > 1\right] \geq 0,$$

where the non-negativity follows by condition A. □

8.B Lemma 8.24

Lemma 8.24. *Assume the statements* 1), 2), *and* 3) *of Theorem 8.17 are valid for any* $L_N^{(i)}$, $i = 1, \ldots, N$, *at level* n. *Then, the statements of the theorem hold if the original minus polar transform is applied to obtain a likelihood ratio at level* $n + 1$.

Proof.

$$\begin{aligned} L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) =& \frac{L_N^{(i)}(y_1^N, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) + L_N^{(i)}(y_{N+1}^{2N}, \hat{u}_{1,e}^{2i-2})}{L_N^{(i)}(y_1^N, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) L_N^{(i)}(y_{N+1}^{2N}, \hat{u}_{1,e}^{2i-2}) + 1} \\ =& \frac{f_1 + f_2}{1 + f_1 f_2}, \end{aligned}$$

where $f_1 := L_N^{(i)}(y_1^N, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})$, and $f_2 := L_N^{(i)}(y_{N+1}^{2N}, \hat{u}_{1,e}^{2i-2})$. Then

$$\begin{aligned} \frac{\partial}{\partial L} L_{2N}^{(2i-1)}\left(y_1^{2N}, \hat{u}_1^{2i-2}\right) =& \frac{(f_1' + f_2')(1 + f_1 f_2) - (f_1 + f_2)(f_1' f_2 + f_1 f_2')}{(1 + f_1 f_2)^2} \\ =& \frac{f_2'(1 - f_1^2) + f_1'(1 - f_2^2)}{(1 + f_1 f_2)^2}. \end{aligned}$$

where $f_j' := \dfrac{\partial f_j}{\partial L}$ for $j = 1, 2$. As $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) > 1$ if and only if $f_1 > 1$ and $f_2 < 1$ (in which case $f_1' > 0$ and $f_2' < 0$ hold by assumption), or $f_1 < 1$ and $f_2 > 1$ (in which case $f_1' < 0$ and $f_2' > 0$ by assumption), we get

$$\frac{\partial}{\partial L} L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) > 0.$$

Hence, $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2})$ is increasing in L if and only if $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) > 1$. Similarly, one can show that $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2})$ is decreasing in L if and only if $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) < 1$. Finally, $L_{2N}^{(2i-1)}(y_1^{2N}, \hat{u}_1^{2i-2}) = 1$ if and only if $f_1 = 1$ or $f_2 = 1$, or $L = 1$. □

8.C Gap to Capacity of the Min-Sum Approximation of the Polar Transform

In this part, we discuss the properties of the min-sum approximation defined in (8.14). We will argue that some of the derivations carried out in Chapter 5 for the exact likelihood ratio process described in (5.2) extend to the following approximate likelihood ratio process:

$$\widetilde{L}_{n+1}(Y_1^{2N}) = \begin{cases} \widetilde{L}_n^-(Y_1^{2N}), & \text{if } B_{n+1} = 1 \\ \widetilde{L}_n^+(Y_1^{2N}), & \text{if } B_{n+1} = 0 \end{cases},$$

where $B_1, \ldots, B_n$ denote the sequence of Bernoulli random variables as before and

$$\widetilde{L}_n^+(Y_1^{2N}) = \widetilde{L}_n(Y_1^N)\widetilde{L}_n(Y_{N+1}^{2N}),$$

$$\widetilde{L}_n^-(Y_1^{2N}) = \exp\Big\{-\text{sign}\left(\log \widetilde{L}_n(Y_1^N) * \log \widetilde{L}_n(Y_{N+1}^{2N})\right) \times \min\Big\{\big|\log \widetilde{L}_n(Y_1^N)\big|, \big|\log \widetilde{L}_n(Y_{N+1}^{2N})\big|\Big\}\Big\}, \quad (8.25)$$

where Y_1^N and Y_{N+1}^{2N} are i.i.d. Observe that the approximate minus transform of the likelihood ratios satisfy, as in the exact case, the following properties:

$$\begin{aligned}
&1)\ \big\{\widetilde{L}_n^-(Y_1^{2N}) > 1\big\} \\
&\quad \iff \big\{\widetilde{L}_n(Y_1^N) > 1\big\} \bigcap \big\{\widetilde{L}_n(Y_{N+1}^{2N}) < 1\big\} \\
&\qquad \bigcup \big\{\widetilde{L}_n(Y_1^N) < 1\big\} \bigcap \big\{\widetilde{L}_n(Y_{N+1}^{2N}) > 1\big\} \\
&2)\ \big\{\widetilde{L}_n^-(Y_1^{2N}) = 1\big\}, \\
&\quad \iff \big\{\widetilde{L}_n(Y_1^N) = 1\big\} \bigcup \big\{\widetilde{L}_n(Y_{N+1}^{2N}) = 1\big\}, \\
&3)\ \big\{\widetilde{L}_n^-(Y_1^{2N}) < 1\big\} \\
&\quad \iff \big\{\widetilde{L}_n(Y_1^N) > 1\big\} \bigcap \big\{\widetilde{L}_n(Y_{N+1}^{2N}) > 1\big\} \\
&\qquad \bigcup \big\{\widetilde{L}_n(Y_1^N) < 1\big\} \bigcap \big\{\widetilde{L}_n(Y_{N+1}^{2N}) < 1\big\}.
\end{aligned}$$

Hence, the below counterparts to (5.3), (5.4), and (5.5) hold:

$$\mathbb{P}\left[\widetilde{L}_n^-(Y_1^{2N}) \lesseqgtr 1\right] = \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \lesseqgtr 1\right]^2 + \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \gtreqless 1\right]^2, \quad (8.26)$$

$$\mathbb{P}\left[\widetilde{L}_n^-(Y_1^{2N}) \gtreqless 1\right] = 2\mathbb{P}\left[\widetilde{L}_n(Y_1^N) \lesseqgtr 1\right]\mathbb{P}\left[\widetilde{L}_n(Y_1^N) \gtreqless 1\right], \quad (8.27)$$

$$\mathbb{P}\left[\widetilde{L}_n^-(Y_1^{2N}) = 1\right] = 2\mathbb{P}\left[\widetilde{L}_n(Y_1^N) = 1\right] - \mathbb{P}\left[\widetilde{L}_n(Y_1^N) = 1\right]^2. \quad (8.28)$$

For the plus transform, as the symmetry in the likelihood ratios is preserved by the approximation, one can use similar to (5.13) the fact that $\sum_{\widetilde{L}(y)=\tilde{\ell}} W(y|0) = \sum_{\widetilde{L}(y)=1/\tilde{\ell}} W(y|0)L(y)$ to derive the below counterparts to (5.6) and (5.7):

$$\mathbb{P}\left[\widetilde{L}_n^+(Y_1^{2N}) \lesssim 1\right] = \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \lesssim 1\right]^2 + \frac{1}{N}\sum_{i=1}^{N} \sum_{\substack{y_1^{2N}: \\ \widetilde{L}_N^{(i)}(y_1^N)\gtrsim 1 \\ \widetilde{L}_N^{(i)}(y_{N+1}^{2N})\gtrsim 1}} W_N(y_1^{2N}|0_1^{2N}) \max\{L_N^{(i)}(y_1^N), L_N^{(i)}(y_{N+1}^{2N})\}, \quad (8.29)$$

and

$$\mathbb{P}\left[\widetilde{L}_n^+(Y_1^{2N}) \gtrsim 1\right] = \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \gtrsim 1\right]^2 + \frac{1}{N}\sum_{i=1}^{N} \sum_{\substack{y_1^{2N}: \\ \widetilde{L}_N^{(i)}(y_1^N)\gtrsim 1 \\ \widetilde{L}_N^{(i)}(y_{N+1}^{2N})\gtrsim 1}} W_N(y_1^{2N}|0_1^{2N}) \min\{L_N^{(i)}(y_1^N), L_N^{(i)}(y_{N+1}^{2N})\}. \quad (8.30)$$

As a result, one can carry the proofs of Propositions 5.5 and 5.7 in exactly the same way by replacing the uses of (5.3), (5.4),(5.5), (5.6), and (5.7) by (8.26), (8.27), (8.28), (8.29), and (8.30), respectively. Moreover, we observe that

$$\begin{aligned}
&\sum_{\substack{y_1^{2N}: \\ \widetilde{L}_N^{(i)}(y_1^N)\gtrsim 1 \\ \widetilde{L}_N^{(i)}(y_{N+1}^{2N})\gtrsim 1}} W_N(y_1^N|0_1^N) W_N(y_{N+1}^{2N}|0_{N+1}^{2N}) \\
&\qquad \left[\max\{L_N^{(i)}(y_1^N), L_N^{(i)}(y_{N+1}^{2N})\} + \min\{L_N^{(i)}(y_1^N), L_N^{(i)}(y_{N+1}^{2N})\}\right] \\
&= \sum_{\substack{y_1^{2N}: \\ \widetilde{L}_N^{(i)}(y_1^N)\gtrsim 1 \\ \widetilde{L}_N^{(i)}(y_{N+1}^{2N})\gtrsim 1}} W_N(y_1^N|0_1^N) W_N(y_{N+1}^{2N}|0_1^N) \left[L_N^{(i)}(y_1^N) + L_N^{(i)}(y_{N+1}^{2N})\right] \\
&\qquad = 2\mathbb{P}\left[\widetilde{L}_N^{(i)}(Y_1^N) \gtrsim 1\right] \sum_{\substack{y_1^N: \\ \widetilde{L}_N^{(i)}(Y_1^N)\gtrsim 1}} W_N(y_1^N|0_1^N) L_N^{(i)}(y_1^N) \\
&\qquad\qquad = 2\mathbb{P}\left[\widetilde{L}_n(Y_1^N) \gtrsim 1\right] \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \lesssim 1\right],
\end{aligned}$$

where in the last line we used the fact that Y_1^N and Y_{N+1}^{2N} are i.i.d. and the exact and approximated likelihood ratios are both symmetrized by the same permutation.

We also make the following observation: If we limit the analysis of Section 5.3 to channels symmetrized by the same permutation, we get $T(W_n, V_n) = T^0(W_n, V_n) =$

$T^1(W_n, V_n)$, where $T^0(W_n, V_n)$ is given by (5.15) and $T^1(W_n, V_n)$ by (5.16). In this case, the following monotonicity properties can be derived by extending the analysis carried for $T_n(W)$ in the proof of Proposition 5.1 to $T_n(W, V)$.

Lemma 8.25. $T_n(W^-, V^-) = T_n(W, V)^2 \leq T_n(W, V) \leq T_n(W^+, V^+)$.

Letting $\widetilde{T}_n = \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \lesssim 1\right] - \mathbb{P}\left[\widetilde{L}_n(Y_1^N) \gtrsim 1\right]$, one can show using the previous derivations the following counterpart to Proposition 5.1.

Proposition 8.26. *The process $\widetilde{T}_n$ is a bounded supermartingale in $[0, 1]$ and converges a.s. to a $\{0, 1\}$ valued random variable since it satisfies*

$$\widetilde{T}_{n+1} = \begin{cases} \widetilde{T}_n^-, & \text{if } B_{n+1} = 1 \\ \widetilde{T}_n^+, & \text{if } B_{n+1} = 0 \end{cases},$$

with the successive choices taken independently, and where

$$\begin{aligned} \widetilde{T}_n^- &= \widetilde{T}_n^2, \\ \widetilde{T}_n^+ &\in \left[\widetilde{T}_n, 2\widetilde{T}_n - \widetilde{T}_n^2\right]. \end{aligned}$$

We managed to prove the convergence of a process associated with the $\min$ approximation of the polar transform. We note that for a given $L(Y_{N+1}^{2N}) \neq 1$, while the exact minus transformation is strictly monotone (increasing or decreasing) in $L(Y_1^N)$, the approximate one is no longer strictly but simply monotone. So, one particular difference caused by the approximation to the minus polar transform is identical likelihood ratios obtained for some outputs which would otherwise be different from each others. Hence, following the approximation a plus transformation at the next level will result in more outputs having likelihood ratios equal to one. In order to characterize the gap to capacity of the approximation, one needs to analyze the fraction of synthetic channels for which the approximation $\widetilde{T}_n$ converges to 1. This appears to be a difficult analysis. Whether ultimately the approximation would cause loss in the performance is an open problem.

A sufficient condition for any approximation to be robust is the following:

$$\text{If } \left\{ \begin{array}{l} \{L_n(Y_1^N) < 1\} = \{\widetilde{L}_n(Y_1^N) < 1\}, \\ \{L_n(Y_1^N) > 1\} = \{\widetilde{L}_n(Y_1^N) > 1\} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \{L_{n+1}(Y_1^{2N}) < 1\} = \{\widetilde{L}_{n+1}(Y_1^{2N}) < 1\}, \\ \{L_{n+1}(Y_1^{2N}) > 1\} = \{\widetilde{L}_{n+1}(Y_1^{2N}) > 1\} \end{array} \right\}. \quad (8.31)$$

An idea could be to slightly perturb the identical likelihood ratios forced by the approximation to distinct values while keeping the symmetry such that this new version of the approximation would satisfy (8.31).

Chapter 9

Channel Polarization over Non-Stationary B-DMCs

In this chapter, we extend the original theory of polar coding to *non-stationary* B-DMCs. In this model, the channel law is no longer assumed to be stationary during the transmission of a codeword and is allowed to vary over each use of the channel in a memoryless fashion. The model is quite useful to capture the effects of time-varying noise present in real communication systems. Note, however, that the nature of this fluctuation is benign and not malicious as in the arbitrarily varying channel model. So, do not let your guard down if malicious opponents are around!

What's Coming, Doc?

With $(U_1, U_2, X_1, X_2, Y_1, Y_2)$ denoting the ensemble where U_1, U_2 are i.i.d. and uniform in $\mathbb{F}_2$, $(X_1, X_2) = (U_1 \oplus U_2, U_2)$ and $\mathbb{P}[y_1, y_1|x_1, x_2] = W(y_1|x_1)W(y_2|x_2)$, the synthetic channels W^- and W^+ describe $U_1 \to Y_1Y_2$ and $U_2 \to Y_1Y_2U_1$, respectively. Looking at this chain of channel combining and splitting operations, recall that Arıkan [2] observes that the polar transform preserves symmetric capacity:

$$\begin{aligned} I(W^-) + I(W^+) &= I(U_1; Y_1Y_2) + I(U_2; Y_1Y_2U_1) \\ &= I(U_1U_2; Y_1Y_2) = I(X_1X_2; Y_1Y_2) = I(W) + I(W). \end{aligned} \tag{9.1}$$

This may be as good a place as any to note that 'conservation of $I(\cdot)$' does not require the channel that connects X_1 to Y_1 and the channel that connects X_2 to Y_2 to be identical; as we remarked in Chapter 4, the same computation as in (9.1) will yield

$$I(\langle W_1, W_2\rangle^-) + I(\langle W_1, W_2\rangle^+) = I(W_1) + I(W_2) \tag{9.2}$$

if W_1 and W_2 are independent channels describing these connections and $\langle W_1, W_2\rangle^-$ and $\langle W_1, W_2\rangle^+$ defined in (4.13) and (4.14) are the synthetic channels that link U_1 to Y_1Y_2 and U_2 to $Y_1Y_2U_1$, respectively.

Arıkan proves that a phenomenon of polarization takes place when the polar transform is repeatedly applied: the n-fold application of the polar transform to synthesize 2^n channels yields, asymptotically, only extremal channels (those with symmetric capacity close to zero or one) and in the 'right proportion' of each of the two kinds. Formally:

Theorem 9.1. *For any binary input channel W, and any $0 < a < b < 1$,*

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{2^n} \#\big\{s \in \{+,-\}^n :\ I(W^s) \in [0,a)\big\} &= 1 - I(W),\\
\lim_{n\to\infty} \frac{1}{2^n} \#\big\{s \in \{+,-\}^n :\ I(W^s) \in [a,b]\big\} &= 0,\\
\lim_{n\to\infty} \frac{1}{2^n} \#\big\{s \in \{+,-\}^n :\ I(W^s) \in (b,1]\big\} &= I(W).
\end{aligned}$$

Arıkan's polar construction for a stationary memoryless channel can be trivially generalized to the case when we need to communicate over a binary input, memoryless, but not necessarily stationary channel. Specifically, suppose W_t, is the channel law at time instant $t \in \mathbb{N}$. Arıkan's polar construction will successively transform this collection of channels into, first, a collection $\{W_{1,t} :\ t \in \mathbb{N}\}$ of channels where

$$\begin{aligned}
W_{1,2m} &= \langle W_{2m}, W_{2m+1}\rangle^-,\\
W_{1,2m+1} &= \langle W_{2m}, W_{2m+1}\rangle^+,
\end{aligned}$$

next, a collection $\{W_{2,t} :\ t \in \mathbb{N}\}$ of channels where

$$\begin{aligned}
W_{2,4m} &= \langle W_{1,4m}, W_{1,4m+2}\rangle^-\\
W_{2,4m+1} &= \langle W_{1,4m+1}, W_{1,4m+3}\rangle^-\\
W_{2,4m+2} &= \langle W_{1,4m}, W_{1,4m+2}\rangle^+\\
W_{2,4m+3} &= \langle W_{1,4m+1}, W_{1,4m+3}\rangle^+,
\end{aligned}$$

and further, with $n \geq 1$, $N = 2^n$ and $0 \leq j < N/2$, to successive collections $\{W_{n,t}\,;\, t \in \mathbb{N}\}$ given by

$$\begin{aligned}
W_{n,Nm+j} &= \langle W_{n-1,Nm+j}, W_{n-1,Nm+N/2+j}\rangle^-\\
W_{n,Nm+N/2+j} &= \langle W_{n-1,Nm+j}, W_{n-1,Nm+N/2+j}\rangle^+.
\end{aligned}$$

(where we have defined $W_{0,t} = W_t$). Figure 9.1 helps visualize the construction that leads to this transformation.

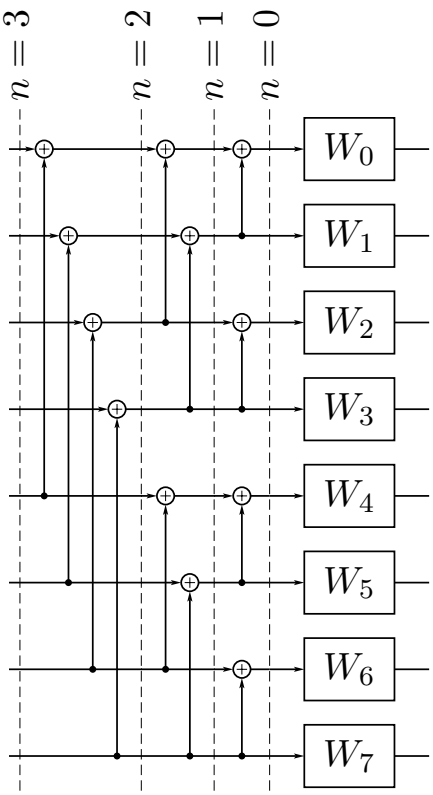

Figure 9.1: Arıkan construction after three stages; the dashed lines are the input planes to the synthetic channels of successive stages.

A natural question to ask at this point is if the successive collections eventually polarize; a possible formalization of what is meant by this could be "for every $0 < a < b < 1$, $\lim_n \theta_n(a, b) = 0$," where

$$\theta_n(a,b) := \liminf_{\tau\to\infty} \frac{1}{\tau} \#\big\{0 \le t < \tau : \ I(W_{n,t}) \in [a,b]\big\}.$$

We shall abide by this particular formalism.

In Theorem 9.7, we will prove that the method of channel combining and splitting via the polar transform polarizes as well non-stationary B-DMCs. The technique we employ for the proof of the theorem will be far from being a trivial extension of the proof technique we have discussed so far to prove channel polarization. This is explained by the fact that while the standard proof that polarization does take place over stationary B-DMCs requires one to construct an artificial stochastic process (the random walk on the 'tree of channels' synthesized by the polar transform defined in (4.4)) and the concept of martingales, this construction cannot be extended to the above non-stationary setting we introduced. As a result, we will first give a simpler proof of the polarization phenomenon over stationary B-DMCs that relies only on elementary concepts. Subsequently, we will show that this simple technique

allows to prove that Arıkan's construction also polarizes non-stationary memoryless channels.

9.1 A simple Proof of Polarization

As we just remarked, Arıkan's proof of polarization makes use of an elegantly constructed martingale and appealing to the martingale convergence theorem. We will give a proof of this theorem by making use of (9.1) and the following restatement[1] of Mrs. Gerber's lemma (see e.g., [50, p. 19]).

Lemma 9.2. *If $I(W) = 1 - h_2(p)$, then $I(W^-) \leq 1 - h_2(p * p)$, where $h_2(p)$ is the binary entropy function and $p * q = p(1-q) + (1-p)q$. Consequently $I(W^+) \geq 1 - 2h_2(p) + h_2(p * p)$ and $\big[I(W^+) - I(W^-)\big]/2 \geq h_2(p * p) - h_2(p)$.*

Observe that for p in the interval $[0, \frac{1}{2}]$, $p * p$ is also in this interval and $p * p \geq p$ with equality only when $p = 0$ or $p = \frac{1}{2}$. Consequently $h_2(p * p) - h_2(p) \geq 0$ with equality only under the same conditions. Since $p \to h_2(p * p) - h_2(p)$ is continuous, it follows that for any $0 < p_0 < p_1 < \frac{1}{2}$,

$$\inf\{h_2(p * p) - h_2(p) : \ p \in [p_0, p_1]\} \\ = \min\{h_2(p * p) - h_2(p) : \ p \in [p_0, p_1]\}, \quad (9.3)$$

and the right hand side is a strictly positive quantity whose value depends on p_0 and p_1. It further follows, that there exists a non-negative function $\kappa(a, b)$ such that (i), $I(W) \in [a, b]$ implies

$$K(W) := \frac{1}{2}\big[I(W^+) - I(W^-)\big] \geq \kappa(a, b),$$

and (ii), $\kappa(a, b) > 0$ whenever $0 < a < b < 1$. Indeed, $\kappa(a, b)$ may be taken to be the right hand side of (9.3) with $p_0 = h_2^{-1}(1 - b)$ and $p_1 = h_2^{-1}(1 - a)$.

With the above technical facts at our disposal we are now ready to prove Theorem 9.1.

Proof of Theorem. Given W and $0 < a < b < 1$, define

$$\theta_n(a, b) := \frac{1}{2^n}\#\big\{s \in \{+, -\}^n : \ I(W^s) \in [a, b]\big\}$$

[1] the equivalence with the standard form is immediate upon noting that $I(W) = 1 - H(X_1|Y_1) = 1 - H(X_2|Y_2)$ and $I(W^-) = 1 - H(X_1 \oplus X_2|Y_1Y_2)$ and further noting that (X_1, Y_1) and (X_2, Y_2) are independent.

as the fraction of synthetic channels that are not yet polarized after n-fold polar transform. Similarly define

$$\alpha_n(a) := \frac{1}{2^n}\#\big\{s \in \{+,-\}^n : \ I(W^s) < a\big\}$$
$$\beta_n(b) := \frac{1}{2^n}\#\big\{s \in \{+,-\}^n : \ I(W^s) > b\big\}.$$

Our task is to show that α_n, β_n and θ_n converge to $1-I(W)$, $I(W)$ and 0, respectively. To that end, let

$$\mu_n = \frac{1}{2^n}\sum_{s\in\{+,-\}^n} I(W^s) \ \text{ and } \ \nu_n = \frac{1}{2^n}\sum_{s\in\{+,-\}^n} \big[I(W^s)\big]^2.$$

Observe that both of these quantities are in $[0,1]$ and note that, thanks to (9.1),

$$\begin{aligned}\mu_{n+1} &= \frac{1}{2^{n+1}}\sum_{s\in\{+,-\}^{n+1}} I(W^s)\\ &= \frac{1}{2^n}\sum_{t\in\{+,-\}^n}\frac{1}{2}\big[I(W^{t+}) + I(W^{t-})\big]\\ &= \frac{1}{2^n}\sum_{t\in\{+,-\}^n} I(W^t) = \mu_n.\end{aligned}$$

Owing to the identity $\frac{1}{2}(u^2+v^2) = \big[\frac{1}{2}(u+v)\big]^2 + \big[\frac{1}{2}(u-v)\big]^2$, we also have

$$\begin{aligned}\nu_{n+1} &= \frac{1}{2^{n+1}}\sum_{s\in\{+,-\}^{n+1}} \big[I(W^s)\big]^2\\ &= \frac{1}{2^n}\sum_{t\in\{+,-\}^n}\frac{1}{2}\big[I(W^{t+})^2 + I(W^{t-})^2\big]\\ &= \frac{1}{2^n}\sum_{t\in\{+,-\}^n}\big[I(W^t)\big]^2 + \big[K(W^t)\big]^2\\ &\geq \nu_n + \theta_n(a,b)\kappa(a,b)^2.\end{aligned}$$

Thus we see that $I(W) = \mu_0 = \mu_1 = \ldots$, and $I(W)^2 = \nu_0 \leq \nu_1 \leq \nu_2 \leq \cdots \leq 1$. The sequence ν_n is thus bounded and monotone and consequently convergent; in particular $\nu_{n+1} - \nu_n$ converges to zero. As θ_n is sandwiched by

$$0 \leq \theta_n(a,b) \leq \frac{\nu_{n+1}-\nu_n}{\kappa(a,b)^2},$$

i.e., between two quantities both convergent to zero, we conclude that $\lim_n \theta_n(a,b) =$

0. Lastly, observe that

$$I(W) = \mu_n \leq a\alpha_n(a) + b\theta_n(a,b) + \beta_n(b) \\ = a + (b-a)\theta_n(a,b) + (1-a)\beta_n(b),$$

thus $I(W) \leq a + (1-a)\liminf_n \beta_n(b)$. Since this last inequality is valid for all a in $(0,b)$, we see (by taking a infinitesimally small) that

$$\liminf_n \beta_n(b) \geq I(W).$$

An analogous calculation, starting with $1-\mu_n \leq \alpha_n(a)+(1-a)\theta_n(a,b)+(1-b)\beta_n(b)$ yields

$$\liminf_n \alpha_n(a) \geq 1 - I(W).$$

But as we have $\alpha_n(a) + \beta_n(b) \leq 1$, it follows that $\lim_n \alpha_n(a) = 1 - I(W)$ and $\lim_n \beta_n(b) = I(W)$. □

9.2 Extensions to Non-Stationary B-DMCs

Encouraged by the simple proof of polarization given in the previous section, we proceed by defining

$$\mu_n = \lim_{\tau\to\infty} \frac{1}{\tau}\sum_{t=0}^{\tau-1} I(W_{n,t}),$$

$$\nu_n = \liminf_{\tau\to\infty} \frac{1}{\tau}\sum_{t=0}^{\tau-1} [I(W_{n,t})]^2,$$

and follow the program of showing (i) $\mu_{n+1} = \mu_n$, (ii) $\nu_{n+1} \geq \nu_n$, and (iii) relate the differences in the ν sequence to the θ sequence.

Remark 9.3. One needs to justify that the limit that defines μ_n exists. For this purpose we assume that μ_0 is well defined, namely, that $\lim_\tau \frac{1}{\tau}\sum_{t=0}^{\tau-1} I(W_t)$ exists. This is sufficient to guarantee the existence of all μ_n. In the following two paragraph we will show the existence of μ_1, the general case follows by the same reasoning.

We begin with a general fact: if a sequence $\{s_\tau : \tau \geq 1\}$ is a 'running average' of a bounded sequence $\{a_t : t \geq 0\}$; i.e., $s_\tau = (1/\tau)\sum_{t=0}^{\tau-1} a_t$, then for any $k \geq 1$, the 'arithmetic' subsequence $s_k, s_{2k}, s_{3k} \ldots$ has the same limit points as $\{s_\tau\}$. For if $s_{n_1}, s_{n_2}, \ldots$ is a convergent subsequence of $\{s_\tau\}$, set $m_i = k\lceil n_i/k \rceil$ by rounding up each n_i to the nearest multiple of k, and consider the sequence s_{m_i} (which is a subsequence of the arithmetic subsequence). Note that s_{m_i} and s_{n_i} differ at most by Ak/m_i with $A = \sup_t |a_t|$. Consequently $\{s_{m_i}\}$ and $\{s_{n_i}\}$ have the same limit.

Thanks to (9.2), for even values of τ, we have

$$\frac{1}{\tau}\sum_{t=0}^{\tau-1} I(W_{1,t}) = \frac{1}{\tau}\sum_{t=0}^{\tau-1} I(W_t),$$

and thus, the arithmetic subsequence (with $k = 2$) of the sequence that defines μ_1 has a single limit point, μ_0. By the previous paragraph the sequence that defines μ_1 then also has a single limit point, and its limit is well defined.

The first step (i) of the program, that $\mu_{n+1} = \mu_n$ is thus a simple consequence of (9.2). For step (ii), we again appeal to a restatement of the general form of Mrs. Gerber's lemma

Lemma 9.4. *If W_1 and W_2 are independent binary input channels with $I(W_1) = 1 - h_2(p_1)$ and $I(W_2) = 1 - h_2(p_2)$, then $I(\langle W_1, W_2\rangle^-) \leq 1 - h_2(p_1 * p_2)$.*

Upon noting that for $0 \leq p_1, p_2 \leq 1/2$, one has $p_1 * p_2 \geq \max\{p_1, p_2\}$, with equality if and only if either p_1 or p_2 takes an extremal value, we see

$$I(\langle W_1, W_2\rangle^-) \leq \min\{I(W_1), I(W_2)\} \leq \max\{I(W_1), I(W_2)\} \leq I(\langle W_1, W_2\rangle^+),$$

where the first and last equalities are strict unless W_1 or W_2 is extremal. Indeed, the same reasoning as in the paragraph after Lemma 9.2 shows that there is a non-negative function $\eta(a, b)$ such that (i) $I(W_1), I(W_2) \in [a, b]$ implies

$$I(\langle W_1, W_2\rangle^+) - I(\langle W_1, W_2\rangle^-) \geq |I(W_1) - I(W_2)| + \eta(a, b),$$

and (ii), $\eta(a, b) > 0$ whenever $0 < a < b < 1$. Equivalently, there is a non-negative function $\zeta(a, b)$ such that

$$K^2(W_1, W_2) := \tfrac{1}{2}\big[I(\langle W_1, W_2\rangle^-)^2 + I(\langle W_1, W_2\rangle^+)^2\big] - \tfrac{1}{2}\big[I(W_1)^2 + I(W_2)^2\big] \geq \zeta(a, b),$$

with $\zeta(a, b) > 0$ whenever $0 < a < b < 1$. In words "when we combine two mediocre channels we create variance."

The non-negativity of ζ (and an appeal to the fact on arithmetic subsequences mentioned in Remark 9.3 above) suffices to conclude that $\nu_{n+1} \geq \nu_n$. Since ν_n is bounded by 1, we see that ν_n is a convergent sequence. The final task is to relate the difference $\nu_{n+1} - \nu_n$ to θ_n. One quickly realizes however, that this is in general not possible: consider, for example that our sequence of channels W_t is such that W_t is extremal when t is even, and is mediocre (say with $I(W_t) = 1/2$) when t is odd. In this case half of our channels are unpolarized, i.e., $\theta = 1/2$, but in the first polarization stage all channel combinations involve one extremal channel, so no

variance is created. Indeed, after the first polarization stage, half of our channels are still mediocre, only their locations may have changed: if the original sequence of channels were

$$W_0, W_1, W_2, W_4, \cdots = \mathtt{G}, \mathtt{M}, \mathtt{B}, \mathtt{M}, \ldots$$

(with G, M, B denoting good, mediocre, and bad channels), and this "GMBM" pattern repeats ad infinitum, after the first stage, we will obtain the (repeating) sequence

$$\mathtt{M}, \mathtt{G}, \mathtt{B}, \mathtt{M}, \ldots$$

In this particular case no mediocre channels are combined even at the the second stage, which yields

$$\mathtt{B}, \mathtt{M}, \mathtt{M}, \mathtt{G}, \ldots$$

It is only at the third stage that mediocre channels are combined.

The example above shows that when $\theta_n \leq 1/2$ there is no strictly positive lower bound to $\nu_{n+1} - \nu_n$, nor even to $\nu_{n+2} - \nu_n$. Nevertheless it is easy to convince oneself that starting with an arrangement where half channels are mediocre, while no extremalization may take place in the first two stages, by the third stage mediocre channels have to be combined and thus creating extremalization. To make use of this insight we need to make a digression and prove a combinatorial fact.

Consider the set $\mathcal{A} = \{\mathtt{B}, \mathtt{M}, \mathtt{G}\}$ with the ordering $\mathtt{B} < \mathtt{M} < \mathtt{G}$. Given two sequences $\mathbf{u} = (u_0, \ldots, u_{K-1})$, $\mathbf{v} = (v_0, \ldots, v_{K-1})$ in $\mathcal{A}^K$, let $\mathbf{u} \vee \mathbf{v}$ and $\mathbf{u} \wedge \mathbf{v}$ denote the component-wise maximum and minimum of the two sequences. Define now, recursively, maps $\pi_k : \mathcal{A}^K \to \mathcal{A}^K$, with $K = 2^k$ and $k \geq 0$ as follows: π_0 is the identity map from $\mathcal{A}$ to $\mathcal{A}$. To find $\pi_{k+1}(u_0, \ldots, u_{2K-1})$ first compute $\mathbf{x} = \pi_k(u_0, \ldots, u_{K-1})$ and $\mathbf{y} = \pi_k(u_K, \ldots, u_{2K-1})$, and set $\pi_{k+1}(u_0, \ldots, u_{2K-1}) = (\mathbf{x} \wedge \mathbf{y})(\mathbf{x} \vee \mathbf{y})$. E.g., to compute $\pi_2(\mathtt{GMMB})$, we need to compute $\pi_1(\mathtt{GM}) = \mathtt{MG}$ and $\pi_1(\mathtt{MB}) = \mathtt{BM}$, and find the result as BMMG. Note that the permutation π_k mimics the rearrangement of the sequence of channels $W_0, \ldots, W_K$ (each classified as good, bad, mediocre) after k polarization stages. For $K = 2^k$, call a sequence $\mathbf{u} \in \mathcal{A}^K$ *extremal* if in the computation of $\pi_k(\mathbf{u})$ we encounter no instance of $\mathtt{M} \wedge \mathtt{M}$ or $\mathtt{M} \vee \mathtt{M}$. For example, the sequences GMMB and BMMG are extremal, whereas MBBM is not. Note that extremal sequences of length $K = 2^k$ are precisely those channel sequences for which during the first k polarization stages no two mediocre channels are combined. The key fact about extremal sequences is that an extremal sequence may not contain too many M's:

Lemma 9.5. *If $K = 2^k$ and $\mathbf{u} = (u_0, \ldots, u_{K-1})$ is extremal, then, at most $\binom{k}{\lfloor k/2 \rfloor}$ of the u_i's may be M.*

Proof. See appendix. □

Corollary 9.6. *If* $\theta_n(a,b) > \binom{k}{\lfloor k/2 \rfloor}/2^k := \epsilon_k$*, then*

$$\nu_{n+k} \geq \nu_n + \delta$$

where $\delta > 0$ *is a quantity that depends only on* k*,* θ_n*,* a*,* b*.*

Proof. For simplicity of notation take $n = 0$. Set $K = 2^k$, and group the channels $W_0, W_1, \ldots$ into blocks of size K: $(W_0, \ldots, W_{K-1})$, $(W_K, \ldots, W_{2K-1})$, ... At the same time, designate each channel W_i to be of type `B`, `M` or `G` according to $I(W_i)$ being less than a, between a and b, or larger than b. This designation will assign to each block a 'pattern' in $\mathcal{A}^K$.

Call a block to be extremal if its pattern is extremal. Since extremal patterns contain at most $\binom{k}{\lfloor k/2 \rfloor}$ `M`'s, we have

$$\theta_0(a,b) \leq (1-\psi)\epsilon_k + \psi$$

where ψ is the fraction (defined as a $\liminf$) of non-extremal blocks. Consequently, ψ satisfies $\psi \geq (\theta_0 - \epsilon_k)/(1-\epsilon_k) > 0$ and we see that a positive fraction of blocks is non-extremal. In the first k stages of polarization, in each of the non-extremal blocks, at least two mediocre channels will be combined, contributing at least $\zeta(a,b)/2^K$ to the variance within that block. Thus,

$$\nu_k \geq \nu_0 + \psi\zeta(a,b)/2^k. \qquad \square$$

Theorem 9.7. *For any sequence of channels* $W_0, W_1, \ldots$*, and any* $0 < a < b < 1$*,*

$$\lim_n \theta_n(a,b) = 0.$$

Proof. Since ν_n is a convergent sequence, for any fixed k, the difference $\nu_{n+k} - \nu_n$ approaches zero as n gets large. By Corollary 9.6, we conclude that $\theta_n \leq \epsilon_k = \binom{k}{\lfloor k/2 \rfloor}/2^k$ for sufficiently large n. Since $\lim_k \epsilon_k = 0$, the conclusion follows. $\square$

Corollary 9.8. *For any sequence of channels* $W_0, W_1, \ldots$ *for which*

$$\mu = \lim_\tau \frac{1}{\tau} \sum_{t<\tau} I(W_t) \tag{9.4}$$

is well defined, Arıkan's construction polarizes the sequence and for any $b < 1$ *the fraction of 'good' synthetic channels*

$$\beta_n = \limsup_\tau \frac{1}{\tau} \#\{0 \leq t < \tau : \ I(W_{n,t}) > b\}$$

approaches μ *as* n *gets large.*

Proof. Follows from Theorem 9.7 above and the same reasoning as in the proof of Theorem 9.1 that established $\beta_n(b) = I(W)$. □

9.2.1 Universal Polar Coding with Channel Knowledge at the Decoder

To study the universality of polar codes over non-stationary B-DMCs, here we apply the order preserving property of the polar transform that we derived in Chapter 6.

Corollary 9.9. *Let $\mathcal{W}$ be a set of B-DMCs and V be a B-DMC such that*

$$|\Delta_V| \prec_{icx} |\Delta_W|,$$

for all $W \in \mathcal{W}$. Then, the polar code designed for the channel V is universal for $\mathcal{W}$ in the sense that, if $W_{0,t} \in \mathcal{W}$, for any $t \in \mathbb{N}$, the following orderings hold:

$$\mathcal{A}_N^{f_s,\gamma}(V) \subset \mathcal{A}_N^{f_s,\gamma}\left(\{W_{n,t} : t \in \mathbb{N}\}\right),$$

for any $\gamma \in (0,1)$ and $f_s \in \mathcal{F}_{s,cx}$, where $\mathcal{F}_{s,cx}$ is defined in (6.2)*.*

Proof. This result follows as a corollary to Theorem 6.5. Note that the order preserving property is already shown there for the more general polar transform $\langle W_1, W_2\rangle^{\pm}$. For notational consistency, we denote by $\{V_{n,t} : t \in \mathbb{N}\}$ the set of synthetic channels obtained from the n-fold application of the polar transform to the copies of the channel V, i.e, we have $V_{0,t} = V$, for any $t \in \mathbb{N}$. By the recursive construction procedure, we conclude that

$$\left|\Delta_{V_{n,t}}\right| \prec_{icx} \left|\Delta_{W_{n,t}}\right|,$$

for all $i = 1, \ldots, 2^n$. □

Assuming that the decoder knows the sequence of realizations of the non-stationary channel, the corollary reveals that the universality emerging from the symmetric convex ordering, and hence also from channel degradation, extends form the stationary setting to the non-stationary one.

Appendix

9.A Proof of Lemma 9.5

We prove Lemma 9.5, to wit, if $K = 2^k$, and $\mathbf{u} \in \mathcal{A}^K$ is extremal than at most $\binom{k}{\lfloor k/2 \rfloor}$ coordinates of $\mathbf{u}$ can be $\mathtt{M}$.

Observe that a sequence of length $2K$ $(u_0, \ldots, u_{2K-1})$ is extremal if and only if $(u_0, \ldots, u_{K-1})$ and $(u_K, \ldots, u_{2K-1})$ are both extremal, and the two sequences $\mathbf{x} = \pi_k(u_0, \ldots, u_{K-1})$ and $\mathbf{y} = \pi_k(u_K, \ldots, u_{2K-1})$ do not have a common $\mathtt{M}$, that is, for no $0 \leq j < K$ we have $x_j = y_j = \mathtt{M}$.

For $k = 0, 1, \ldots,$ and $K = 2^k$ define

$$R_k = \{\pi_k(\mathbf{u}) : \mathbf{u} \in \mathcal{A}^K \text{ is extremal}\}.$$

Clearly $R_0 = \mathcal{A}$, and by the observation in the previous paragraph,

$$R_{k+1} = \{\mathbf{uv} : \mathbf{u} = \mathbf{x} \wedge \mathbf{y}, \mathbf{v} = \mathbf{x} \vee \mathbf{y}, \mathbf{x}, \mathbf{y} \in R_k \text{ with no common } \mathtt{M}\}. \quad (9.5)$$

The first few R_k's can be readily found as:

$$\begin{aligned} R_0 &= \{\mathtt{B}, \mathtt{M}, \mathtt{G}\} \\ R_1 &= \{\mathtt{BB}, \mathtt{BM}, \mathtt{MG}, \mathtt{GG}\} \\ R_2 &= \{\mathtt{BBBB}, \mathtt{BBBM}, \mathtt{BBBG}, \mathtt{BBMG}, \mathtt{BBGG}, \mathtt{BMBG}, \mathtt{BMMG}, \\ &\qquad \mathtt{BMGG}, \mathtt{BGBG}, \mathtt{BGMG}, \mathtt{BGGG}, \mathtt{MGGG}, \mathtt{GGGG}\} \end{aligned}$$

As $\pi_k(\mathbf{u})$ is a permutation of $\mathbf{u}$, to prove Lemma 9.5 it suffices to prove:

Lemma 9.10. *If $K = 2^k$ and $\mathbf{u} = (u_0, \ldots, u_{K-1}) \in R_k$, then at most $\binom{k}{\lfloor k/2 \rfloor}$ of the u_i may be $\mathtt{M}$.*

To prove Lemma 9.10 we first show that the recursion (9.5) may be written in a simpler form.

Lemma 9.11. *If $\mathbf{x}, \mathbf{y} \in R_k$, then (i) $\mathbf{x} \wedge \mathbf{y} \in R_k$, (ii) $\mathbf{x} \vee \mathbf{y} \in R_k$, moreover, (iii)*

$$R_{k+1} = \{\mathbf{xy} : \mathbf{x} \leq \mathbf{y}, \mathbf{x}, \mathbf{y} \in R_k \textit{ with no common } \mathtt{M}\} \quad (9.6)$$

($\mathbf{x} \leq \mathbf{y}$ indicates component-wise inequality).

Proof. The assertions (i) and (ii) for $k = 0$ are trivially true. The proof follows by induction on k: The truth of (i), (ii) for a given value of k and (9.5) establish the truth of (iii) for k, and also the truth of (i) and (ii) for $k + 1$. □

With $K = 2^k$, an element $\mathbf{u} = (u_0, \ldots, u_{K-1})$ of $\mathcal{A}^K$ can be viewed as a function from the subsets of $\{1, \ldots, k\}$ to $\mathcal{A}$ by the association $u(S) = u_j$ where $j = \sum_{i \in S} 2^{i-1}$. E.g., $\mathbf{u} = \mathtt{BMMG}$ is associated with the function $u(\cdot)$ with $u(\{\}) = \mathtt{B}$,

$u(\{1\}) = u(\{2\}) = \mathtt{M}$, $u(\{1,2\}) = \mathtt{G}$. Now, if $k \geq 1$ and $\mathbf{u} \in R_k$, by (9.6), we see

$$(u_0 \dots u_{K/2-1}) \leq (u_{K/2}, \dots, u_{K-1})$$

or equivalently $u(S) \leq u(S \cup \{k\})$ (and not both equal to $\mathtt{M}$) for all S not including k. The recursive structure of (9.6) further implies that $u(S) \leq u(S \cup \{i\})$ (and not both equal to $\mathtt{M}$) for all S not including i, $(1 \leq i \leq k)$. We thus see that $\mathbf{u} \in R_k$ if and only if $u(S) \leq u(T)$ (and not both equal to $\mathtt{M}$) whenever $S \subsetneq T$.

Proof of Lemma 9.10. Given $\mathbf{u} \in R_k$, by the observation in the previous paragraph, the collection

$$\mathcal{F} = \{S : \ u(S) = \mathtt{M}\}$$

forms a Sperner system on $\{1, \dots, k\}$, i.e., no member of $\mathcal{F}$ includes any other. But no Sperner system on $\{1, \dots, k\}$ may have more than $\binom{k}{\lfloor k/2 \rfloor}$ members (see, e.g., [51, Theorem §3.1 on p. 10]). □

Chapter 10

Conclusions

We hope you did not directly jump into this last chapter. Otherwise, it would be wiser to continue your reading with the dialogues of 'Mr. Minus and Mrs. Plus' when the clock strikes.

Is This the End, Doc?

It is time to recapitulate what we accomplished in this thesis and point out possible directions for future research.

10.1 Overview of Thesis Contributions

The overview proceeds without necessarily following the order of the presentation.

10.1.1 Variations on the Polar Transform ♫♫♩♪♫.♫♫♩♪♫.

In this thesis, the original polar coding framework of Arıkan [2] was extended along various paths. We used the polar transform $\langle W_1, W_2 \rangle^{\pm}$ to combine two independent channels that are not necessarily identical. We showed that the one-step application of this generalized polar transform does improve general quality measures of the system, and the improvement is bound to happen within the limits dictated by the extremal channels— the BECs and the BSCs. In particular, Theorem 4.7 certified that polarization 'creates' E_0. We saw that the improvement in E_0 translates to an improvement in the complexity–error-probability trade-off. This observation gave yet another justification as to why the polar transform yields capacity achieving and low complexity codes.

Chapter 9 went further ahead by combining arbitrary B-DMCs via the recursive application of the polar transform. With this application discussed for the first time in this thesis, we extended the polar coding framework to non-stationary channels by proving that the polar transform polarizes non-stationary channels in the same way as it polarizes stationary ones.

Another previously 'untouched' problem that we considered was: Polar coding in the presence of a decoding mismatch. Motivated by the fact that in practice almost all decoders are mismatched, we studied in Chapter 7 the performance of polar coding with mismatched polar decoding. We showed that the synthetic channels 'seen' by a mismatched polar decoder end up being either 'perfect' or 'completely noisy'. Thus, we found that the transmission capacity with mismatched polar decoding is equal to the fraction of the synthetic channels which are still perfect from the mismatched decoder's perspective. Moreover, we proved that this fraction is lower bounded by a sequence of tighter bounds whose first element is given by the generalized mutual information parameter, which we denoted by $I(W, V)$ and defined in (7.4). Furthermore, we showed that the speed of polarization of the processes is not affected by the presence of mismatch. Polar coding is thus also possible in mismatched communication scenarios, and the mismatched polar decoder still operates in $O(N \log N)$ complexity in the block-length N. Surprisingly, it turned out that it is even possible to achieve better rates than the classical mismatched capacity with the help of the polar transform.

These contributions strongly support the view point that "polarization is a fairly general phenomenon". As illustrated by this thesis and by many other works, the idea of channel polarization can be applied to different communication scenarios, and polar coding can be welcomed with all its unique properties.

10.1.2 Extrema, Extremal, Extremality

The extremality of the BEC and the BSC was shown for the polar transform in Chapter 4. The result can be interpreted in the context of information combining: Theorem 4.8 shows that even if we change the measure of information from the customary mutual information to Gallager's E_0 evaluated under the uniform input distribution, the BEC and BSC still remain extremal. The extremal channels are of particular importance for the theory of polarization as they provide bounds to the evolution of information measures associated with the synthetic channels whose output alphabets keep growing in general arbitrarily large under the sequence of polar transformations. In fact, these bounds proved already to be useful technical tools for the study of the channel polarization phenomenon. For instance, they can be used to identify to which values the synthetic channels' symmetric capacities or E_0 parameters converge. Even the derivation of the rate of polarization carried out

in [30] uses the fact that the BEC model constitutes an extremal channel model in the one-step evolution of the Bhattacharyya parameter while being preserved under the polar transform [2] (so the BEC is extremal during the entire evolution).

Besides the extremality results for the polar transform, we described in Theorem 3.1 certain extremality properties for B-DMCs when the information measure is once more Gallager's E_0 evaluated under the uniform input distribution. These properties yielded in straightforward fashion extremal properties for the Rényi entropies and recent results by Guillén i Fàbregas et al. [22], [23] showing that amongst all symmetric B-DMCs of the same capacity, the BEC and the BSC are $E_0(\rho)$ extremal for all $\rho > -1$. It is worth emphasizing that all the conclusions of the chapter are valid for arbitrary binary input channels as long as one evaluates all the quantities under the uniform input distribution.

Although it is not very often explicitly formulated, the extremal channel bounds are powerful information-theoretic tools. It would thus be appropriate to add a couple of words about the proof technique used in the proofs of these extremality results. The technique is not new and the use of similar techniques in information theory can be traced back to the proof of Mrs. Gerber's Lemma in [52]. Similar techniques have been used to show that the BEC and the BSC are extremal channels in the context of information combining in [25] and [53]. The result of [22] we have just mentioned uses as well this technique. Obtaining upper and lower bounds on $E[f(A)]$, for a bounded random variable A, when the function $f(.)$ is convex or concave is a straightforward application of Jensen's inequalities. However, expressing the quantity of interest in a convenient form and proving convexity may not be trivial at all.

10.1.3 Now in 3D! Performance vs. Complexity vs. Universality

In Chapter 8, we took a designer's perspective to identify channel conditions under which universal polar codes can be designed. Partial results were obtained for the cases where either the encoder or the decoder know the actual channel which occurred during the transmission of the codeword. We saw that the conditions for the universal encoding and the universal decoding (with the original polar successive cancellation decoder) of polar codes are different in nature: while conditions such as stochastic degradation, symmetric convex ordering, or even less noisy ordering enable the design of universal encoders, convexity or one-sidedness must be imposed on the sets of channels to design universal polar decoders. Unfortunately, these conditions are not necessarily compatible with each other, and for this reason, it was argued that the original polar code construction would lead to universal codes for both the encoding and the decoding operations only under strong assumptions on the set of channels.

Following this observation, the chapter studied the performance of two alternative decoders which slightly modified the original polar decoding procedure in order to obtain universal polar code designs for certain classes of channels. The first modification introduced the min-sum approximation to the recursive computations of the likelihood ratios of the synthetic channels. We showed that mismatched polar codes using this approximation at the decoder are robust over the class of BSCs whose crossover probabilities satisfy $p < 0.5$ (or $p > 0.5$). Moreover, we observed in simulations that no significant loss is incurred by the approximation. Hence when the encoder knows the channel, regardless of the decoding metric, we expect that rates very close to the actual channel capacity can be achieved by polar codes which are designed using this approximation. In scenarios where hardware implementation dictates such an approximation must be introduced as stated by [48], unavoidably no better rates can be achieved. In this case, polar codes designed with possibly mismatched parameters provide at the least a 'practical universality' for BSCs.

The second modification that we considered incorporated the idea of the generalized likelihood ratio test into the decoding procedure of polar codes. In Theorem 8.23, we proved that information set designs for polar coding can be made universal over some finite class of channels satisfying certain mild conditions by granting the decoder more resources to summon multiple runs of the polar successive cancellation decoder with the metrics of the different channels in the class. No surprise good performance + low-complexity + universality comes with a price. The time complexity of a serial implementation of this decoder, we called the GRLT polar decoder, scales at most like $O(LN \log N)$, where L is the number of channels. Thus, when the value of L is not too large, the GRLT polar decoder will still be a low complexity decoder, and the price for universality will be affordable. What about arbitrary class of channels? This result tells us that for classes of channels which can be represented with, or quantized into, a rather small number of representative channels satisfying the necessary mild conditions, information set designs for polar coding can achieve universal rates over these compound sets by using a low complexity GRLT polar decoder implementation.

10.1.4 Mission Ispossible: The Undergrad Experience

Arıkan's polar codes, besides their various merits, are natural candidates to be taught in a course in information theory. Unlike classical codes, polar codes can be introduced without any reference to algebraic structures (no disrespect to algebra intended); their economy of concepts allow them to play an analogous role in teaching channel coding as Huffman codes do in teaching source coding; they also explicitly show that affine codes achieve the symmetric capacity of binary input channels. Indeed, polarization can be described as soon as the concept of a 'stationary memoryless channel' and mutual information is explained. From a pedagogical point

of view, the numerical study of the case of the binary erasure channel is a very useful exercise: the students can see with their own eyes that as the number of polarization levels increase the fraction of 'unpolarized channels' vanishes.

Yet, the standard proof that polarization does take place requires one to construct an artificial stochastic process (the random walk on the 'tree of channels' synthesized by the polar transform) and the concept of martingales. In a senior undergraduate or introductory graduate level course, it is unwise to assume familiarity with martingales; consequently, any instructor wishing to teach polarization in such a course either has to resort to hand-waving arguments, or spend time to teach a fair bit of background. Chapter 9 described a simpler proof of the polarization phenomenon that relies only on elementary concepts familiar to senior undergraduates.

The method described in Section 9.1 hopefully makes clear that polarization and polar codes can be taught without requiring any particularly advanced background in stochastic processes. The method is also versatile enough to establish polarization in the more general context of non-stationary channels (but one probably should not attempt the proof in an undergraduate course). We should note that the method, just as the original method of Arıkan, is not powerful enough to conclude anything about the speed of polarization. In particular, to show that polarization happens fast enough to arrest error propagation, or that the error probability of polar codes decays exponentially in the square root of the block-length still require non-elementary techniques.

10.2 Open Problems

We close the curtain with some directions for future research.

1. **Efficient construction of mismatched information sets:** To put into practice the coding theorem of Section 7.3, we proposed in the same section to construct mismatched information sets of the form (7.22) via statistical methods. On the other hand, we believe that low complexity algorithms as proposed in [3] for the matched information sets can also be found for the mismatched sets, (this would be useful in the scenario we know the true channel, but do mismatched polar decoding due to feasibility requirements). In that respect, the symmetric convex ordering could be a useful tool to approximate the channels, but preserving the ordering under the polar transform in the presence of a mismatch could be more tricky.

2. **Gap to capacity of the min-sum approximation:** An approach to study the universality of polar codes over BSCs would be to analyze the performance of the min approximation described in (8.14) in a theoretical framework. In

Appendix 8.C, we developed some ideas in that direction, but left the topic as an open problem.

3. **À la mode, list decoding of polar codes:** Another problem of practical interest which was not considered in this thesis is the finite block-length performance of polar codes. List decoding was proposed in [54] and shown to significantly improve this performance. Thus, it would be interesting to study the effects of a decoding mismatch on the finite-length performance of polar codes with list decoding.

4. **Extensions to non-binary inputs:** This thesis focused on polar coding over binary input channels. It would certainly be valuable to extend our results to non-binary input alphabets.

5. **Playing over arbitrarily varying channels:** The arbitrary varying channel model is quite similar to the non-stationary channel model in its description, yet quite different in its nature; the channel variations might in fact be caused by malicious attackers trying to jam the communication. As a result, the code designer would need to figure out a polar coding strategy to protect the communication system against these attacks. We believe the material presented in this thesis should prove useful in the study of this setting.

Appendix A

A.1 Linear Codes Achieve the Symmetric Compound Capacity: A Proof by Strong Typicality

Definition A.1. Given a field $(\mathcal{X}, +, .)$, we say that a code $\mathcal{C} \subset \mathcal{X}^N$ is linear if it is a vector space. Therefore, $\forall a, b \in \mathcal{X}$, $\forall \mathbf{x}, \mathbf{y} \in \mathcal{C}$, we have $a\mathbf{x} + b\mathbf{y} \in \mathcal{C}$.

The channel coding theorem for linear codes states that capacity achieving linear codes exist. The proof simply relies on Shannon's proof of the random coding theorem. Here, we follow a similar approach. Using a random coding argument, we first show in Theorem A.2 that $I_{\text{comp}}(\mathcal{W}) = \min_{W \in \mathcal{W}} I(W)$ can be achieved in the compound case. Then, we extend the result to linear codes in Theorem A.4.

Theorem A.2. *Let $\mathcal{W}$ be a finite set of channels with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, uniform input distribution and transition probabilities $W(y|x)$ where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ for each $W \in \mathcal{W}$. Then, there exists a block code of block-length N and rate $R \geq 0$ such that $R \leq I_{\text{comp}}(\mathcal{W})$ and, for any $\xi > 0$, the average block decoding error probability $P_{\text{e, avg}} < \xi$ with $\xi \to 0$ as $N \to \infty$.*

The proof of the theorem will use the notion of strong typicality introduced in Definition 8.21. Before we start the proof, we give a useful lemma related to strong typicality.

Lemma A.3. *Let Z_1^N be a sequence of random variables drawn identically independently according to the product distribution $\prod_i P(z_i)$. Given another product distribution $\prod_i Q(z_i)$, we have*

$$\mathbb{P}\left[z_1^N \in \mathcal{T}_Q^{N,\varepsilon}\right] = \exp_2\left\{-N\left[\text{Div}(Q\|P) \pm O(\varepsilon)\right]\right\},$$

where $\text{Div}(Q\|P)$ is the Kullback-Leibler divergence.

Proof of Lemma A.3. We denote $\mathbf{z} = z_1^N$. Then

$$\begin{aligned}
\mathbb{P}\left[\mathbf{z} \in \mathcal{T}_Q^{N,\varepsilon}\right] &= \sum_{\mathbf{z}\in\mathcal{T}_Q^{N,\varepsilon}} P(z_1)\dots P(z_N) \\
&= \sum_{\mathbf{z}\in\mathcal{T}_Q^{N,\varepsilon}} \prod_{z\in\mathcal{Z}} P(z)^{N(Q(z)\pm\varepsilon)} \\
&= \sum_{\mathbf{z}\in\mathcal{T}_Q^{N,\varepsilon}} \exp_2\left\{N\left[\sum_{z\in\mathcal{Z}} Q(z)\log P(z) \pm O(\varepsilon)\right]\right\} \\
&\overset{(1)}{=} \exp_2\left\{N\left[H_Q(Z)\pm\varepsilon\right]\right\} \exp_2\left\{N\left[\sum_{z\in\mathcal{Z}} Q(z)\log P(z) \pm O(\varepsilon)\right]\right\} \\
&= \exp_2\left\{-N\left[\mathrm{Div}(Q\|P)\pm O(\varepsilon)\right]\right\}
\end{aligned}$$

where (1) follows by the strong asymptotic equipartition property and $H_Q(Z)$ is the entropy of the random variable $Z \sim Q(z)$. □

Proof of Theorem A.2. We assume that for each message $m = 1, \dots, M$ the encoder generates the codewords using the function $\mathcal{E}nc\colon \{1,\dots,M\} \to \mathcal{X}^N$:

$$\mathcal{E}nc(m) = \mathbf{x}_m = \{x_1^N\}_m,$$

where each element of the m-th codeword $\{x_i\}_m$ are chosen i.i.d. from $P(x)$. Given the output sequence $\mathbf{y} = y_1^N$, the decoder makes a decision using the function $Dec\colon \mathcal{Y}^N \to \{1,\dots,M\} \cup 0$:

$$Dec(\mathbf{y}) = \begin{cases} m, & \text{if m is the unique message such that} \\ & \qquad (\mathcal{E}nc(m), \mathbf{y}) \in \bigcup\limits_W \mathcal{T}_{PW}^{N,\varepsilon} \\ 0, & \text{otherwise} \end{cases},$$

where PW stands for the joint distribution $P(x)W(y|x)$. Let the block decoding error probability of a message m be $P_{\mathrm{e},m}$. We note that

$$P_{\text{e, avg}} = \sum_{m=1}^{M} \frac{1}{M}\mathbb{E}[P_{\mathrm{e},m}] = \mathbb{E}[P_{\mathrm{e},m}],$$

as by symmetry $\mathbb{E}[P_{\mathrm{e},1}] = \dots = \mathbb{E}[P_{\mathrm{e},M}]$. The decoder makes an error if and only if

- $(\mathbf{x}_m, \mathbf{y}) \notin \bigcup\limits_W \mathcal{T}_{PW}^{N,\varepsilon}$, or
- For some $m' \neq m$, $(\mathbf{x}_{m'}, \mathbf{y}) \in \bigcup\limits_W \mathcal{T}_{PW}^{N,\varepsilon}$.

Hence,

$$\mathbb{E}[P_{\mathrm{e},m}] \leq \mathbb{P}\left[(\mathbf{x}_m, \mathbf{y}) \notin \bigcup_W \mathcal{T}_{PW}^{N,\varepsilon}\right] + \sum_{m' \neq m} \mathbb{P}\left[(\mathbf{x}_{m'}, \mathbf{y}) \in \bigcup_W \mathcal{T}_{PW}^{N,\varepsilon}\right].$$

As the codewords generated by the encoder are independent, we can deduce that, for any $m' \neq m$, the pair $(\mathbf{x}_{m'}, \mathbf{y})$ has probability $\prod_i P(\{x_i\}_{m'})PW_T(y_i)$, where W_T is the true channel. Therefore,

$$\begin{aligned}
&\mathbb{P}\left[(\mathbf{x}_{m'}, \mathbf{y}) \in \mathcal{T}_{PW}^{N,\varepsilon}\right] = \exp_2\left\{-N\left[\mathrm{Div}\left(P(x)W(y|x) \| P(x)PW_T(y)\right) \pm O(\varepsilon)\right]\right\} \\
&= \exp_2\left\{-N\left[\sum_{x,y} P(x)W(y|x)\log\frac{W(y|x)}{PW(y)} + \sum_y PW(y)\log\frac{PW(y)}{PW_T(y)} \pm O(\varepsilon)\right]\right\} \\
&\leq \exp_2\left\{-N\left[I(W) - O(\varepsilon)\right]\right\}.
\end{aligned}$$

Due to the strong law of large numbers, $\mathbb{P}\left[(\mathbf{x}_m, \mathbf{y}) \notin \bigcup_W \mathcal{T}_{PW}^{N,\varepsilon}\right] \to 0$. Hence,

$$\begin{aligned}
P_{\mathrm{e,\,avg}} &\leq \sum_{m' \neq m} \mathbb{P}\left[(\mathbf{x}_{m'}, \mathbf{y}) \in \bigcup_W \mathcal{T}_{PW}^{N,\varepsilon}\right] \\
&\leq \sum_{m' \neq m} \sum_W \mathbb{P}\left[(\mathbf{x}_{m'}, \mathbf{y}) \in \mathcal{T}_{PW}^{N,\varepsilon}\right] \\
&\leq |\mathcal{W}| \exp_2\left\{-N\left[I_{\mathrm{comp}}(\mathcal{W}) - R - O(\varepsilon)\right]\right\},
\end{aligned}$$

where we assumed $M = \lceil 2^{NR} \rceil$ and $|\mathcal{W}|$ is finite. □

Theorem A.4. *Let $\mathcal{W}$ be a finite set of channels with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, uniform input distribution and transition probabilities $W(y|x)$ for each $W \in \mathcal{W}$. Then, for any rate $R \geq 0$ such that $R \leq I_{\mathrm{comp}}(\mathcal{W})$ and any $\xi > 0$ there exists a linear code with block decoding error probability $P_{\mathrm{e}} < \xi$.*

Proof of Theorem A.4. An affine code with block-length N generates its codewords as follows: $\mathbf{x} = \mathbf{u}G_N + \mathbf{v}$, where $\mathbf{u}$ is the message, G_N is the generator matrix of the affine code, and $\mathbf{v}$ is an arbitrarily fixed sequence in $\mathcal{X}^N$. From [5, Chapter 6], we know that the codewords of an affine code are pairwise independent. Hence, the proof of Theorem A.2 can be applied to random affine codes. In addition, we recognize that the same performance as an affine code can be obtained with a communication system using a linear code with codewords $\mathbf{x} = \mathbf{u}G_N$. Hence, we also expect random linear codes to achieve the symmetric compound capacity $I_{\mathrm{comp}}(\mathcal{W})$. Moreover, since the average error probability can be made arbitrarily small, we know that there exists at least one linear code which will have an error probability smaller than or equal to the average error probability. □

Bibliography

[1] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[2] E. Arıkan. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Trans. Inf. Theory*, 55(7):3051–3073, 2009.

[3] I. Tal and A. Vardy. How to construct polar codes. *IEEE Trans. Inf. Theory*, 59(10):6562–6582, 2013.

[4] E. Arıkan. Channel combining and splitting for cutoff rate improvement. *IEEE Trans. Inf. Theory*, 52(2):628–639, 2006.

[5] R. G. Gallager. *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., New York, NY, USA, 1968.

[6] D. J. Costello and Jr. Forney, G. D. Channel coding: The road to channel capacity. *Proc. of the IEEE*, 95(6):1150–1177, 2007.

[7] R. Fano. A heuristic discussion of probabilistic decoding. *IEEE Trans. Inf. Theory*, 9(2):64–74, 1963.

[8] P. Elias. Coding for noisy channels. *IRE Conv. Rec.*, pages 37–46, 1955.

[9] E. Arıkan. An inequality on guessing and its application to sequential decoding. *IEEE Trans. Inf. Theory*, 42(1):99–105, 1996.

[10] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacity of a class of channels. *The Annals of Mathematical Statistics*, 3(4):1229–1241, 1959.

[11] I. Csiszár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, Inc., Orlando, FL, USA, 1982.

[12] I Csiszár and P. Narayan. Channel capacity for a given decoding metric. *IEEE Trans. Inf. Theory*, 41(1):35–43, 1995.

[13] I. Csiszár. The method of types [information theory]. *IEEE Trans. Inf. Theory*, 44(6):2505–2523, 1998.

[14] S. Arimoto. Information measures and capacity of order α for discrete memoryless channels. In I. Csiszár and P. Elias, editors, *Topics in information theory*, volume 16, pages 41–52, The Netherlands, 1977. North-Holland Publishing Co.

[15] I. Csiszár. Generalized cutoff rates and Renyi's information measures. *IEEE Trans. Inf. Theory*, 41(1):26–34, 1995.

[16] A. Rényi. On measures of entropy and information. *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, 1:547–561, 1961.

[17] R. G. Gallager. A simple derivation of the coding theorem and some applications. *IEEE Trans. Inf. Theory*, 11(1):3–18, 1965.

[18] S. Arimoto. On the converse to the coding theorem for discrete memoryless channels (corresp.). *IEEE Trans. Inf. Theory*, 19(3):357–359, 1973.

[19] J. L. Massey. Guessing and entropy. In *Proc. of the IEEE Int. Symposium on Inf. Theory*, page 204, 1994.

[20] E. Arıkan and E. Telatar. BEC and BSC are E_0 extremal. Unpublished note.

[21] *The American Heritage Dictionary of the English Language,*. Fifth Edition copyright 2014 by Houghton Mifflin Harcourt Publishing Company.

[22] A. Guillén i Fàbregas, I. Land, and A. Martinez. Extremes of random coding error exponents. In *Proc. of the IEEE Int. Symposium on Inf. Theory*, pages 2896–2898, 2011.

[23] A. Guillén i Fàbregas, I. Land, and A. Martinez. Extremes of error exponents. *IEEE Trans. Inf. Theory*, 59(4):2201–2207, 2013.

[24] E. Arıkan and E. Telatar. On the rate of channel polarization. *eprint arXiv:0807.3806*, 2008.

[25] I. Sutskover, S. Shamai, and J. Ziv. Extremes of information combining. *IEEE Trans. Inf. Theory*, 51(4):1313–1325, 2005.

[26] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Inf. Theory*, 47(2):498–519, 2001.

[27] D. Williams. *Probability with Martingales*. Cambridge mathematical textbooks. Cambridge University Press, 1991.

[28] Jr. Forney, G. D. 1995 Shannon Lecture—Performance and complexity. *IEEE. Inf. Theory Soc. Newslett.*, 46:3–4, 1996.

[29] A. J. Viterbi and J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, New York, NY, USA, 1979.

[30] E. Arıkan and I. E. Telatar. On the rate of channel polarization. In *Proc. of the IEEE Int. Symposium on Inf. Theory,*, pages 1493–1495, 2009.

[31] R. Mori and T. Tanaka. Performance of polar codes with the construction using density evolution. *IEEE Communications Letters*, 13(7):519–521, July 2009.

[32] S. Karlin and A. Novikoff. *Generalized Convex Inequalities*. Pacific J. Math, 1963.

[33] R. Szekli. *Stochastic ordering and dependence in applied probability*. Lecture notes in statistics. Springer-Verlag, 1995.

[34] S. B. Korada. *Polar codes for channel and source coding*. PhD thesis, Lausanne, 2009.

[35] D. Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

[36] W. Hürlimann. Extremal moment methods and stochastic orders. *Boletín de la Asociación Matemática Venezolana*, 15(2):153–301, 2008.

[37] R. Pedarsani, S. H. Hassani, I. Tal, and I. E. Telatar. On the construction of polar codes. In *Proc. of the IEEE Int. Symposium on Inf. Theory*, pages 11–15, 2011.

[38] V. D. Goppa. Nonprobabilistic mutual information without memory. *Probl. Contr. Inf. Theory*, 4:97–102, 1975.

[39] C. E. Shannon. The zero error capacity of a noisy channel. *IRE Trans. on Inf. Theory*, 2(3):8–19, 1956.

[40] V. B. Balakirsky. Coding theorem for discrete memoryless channels with given decision rule. *Proc. of the First French-Soviet Workshop on Algebraic Coding*, pages 142–150, Jul. 1991.

[41] V. B. Balakirsky. A converse coding theorem for mismatched decoding at the output of binary-input memoryless channels. *IEEE Trans. Inf. Theory*, 41(6):1889–1902, 1995.

[42] T. R. M. Fischer. Some remarks on the role of inaccuracy in Shannon's theory of information transmission. In *Trans. of the Eighth Prague Conference*, volume 8A of *Czechoslovak Academy of Sciences*, pages 211–226. Springer Netherlands, 1978.

[43] G. Kaplan and S. Shamai. Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment. *AËU*, 47(4):228–239, 1993.

[44] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai. On information rates for mismatched decoders. *IEEE Trans. Inf. Theory*, 40(6):1953–1967, 1994.

[45] E. Abbe and L. Zheng. Linear universal decoding for compound channels. *IEEE Trans. Inf. Theory*, 56(12):5999–6013, 2010.

[46] D. Sutter and J. M. Renes. Universal polar codes for more capable and less noisy channels and sources. In *Proc. of the IEEE Int. Symposium on Inf. Theory*, pages 1461–1465, 2014.

[47] J. Körner and K. Marton. A source network problem involving the comparison of two channels. *Trans. Colloq. Inf. Theory*, 1975.

[48] C. Leroux, I. Tal, A. Vardy, and W. J. Gross. Hardware architectures for successive cancellation decoding of polar codes. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 1665–1668, 2011.

[49] I. Csiszár and P. C. Shields. Information theory and statistics: A tutorial. *Foundations and Trends in Communications and Information Theory*, 1(4):417–528, 2004.

[50] A. E. Gamal and Y. H. Kim. *Network Information Theory*. Cambridge University Press, New York, NY, USA, 2011.

[51] B. Bollobás. *Combinatorics*. Cambridge University Press, Cambridge, 1986.

[52] A. D. Wyner and J. Ziv. A theorem on the entropy of certain binary sequences and applications–i. *IEEE Trans. Inf. Theory*, 19(6):769–772, 1973.

[53] I. Land, S. Huettinger, P. A. Hoeher, and J. B. Huber. Bounds on information combining. *IEEE Trans. Inf. Theory*, 51(2):612–619, 2005.

[54] I. Tal and A. Vardy. List decoding of polar codes. In *Proc. of the IEEE Int. Symposium on Inf. Theory*, pages 1–5, 2011.

[55] C. E. Shannon. A note on a partial ordering for communication channels. *Information and Control*, 1(4):390 – 397, 1958.

[56] E. Telatar. Private communications.

Index of Terms

Index of Symbols

Curriculum Vitae

Education

École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Ph.D., Communication Systems — Apr. 2010 - Sept. 2014

M.Sc., Communication Systems — Sept. 2007 - Feb. 2010

- Specialization: Wireless Communications

Middle East Technical University (METU), Ankara, Turkey

B.Sc., Electrical and Electronics Engineering — Sept. 2003 - Jun. 2007

- Specialization: Telecommunications

Professional Experience

École Polytechnique Fédérale de Lausanne, Switzerland

Research Assistant, Information Theory Laboratory — Apr. 2010 - Oct. 2014

Teaching Assistant

- Information Theory and Coding — Fall 2011, Fall 2012
- Probability and Statistics for IN, SC — Spring 2012

Nokia Research Center, Lausanne, Switzerland

Intern, Pervasive Communications Laboratory — Feb. - Aug. 2009

ASELSAN Military Electronic Inc., Ankara, Turkey

Intern — Jul. - Aug. 2006

Languages

Turkish (native), **English** (fluent), **French** (fluent), **German** (basic)

List of Publications

Journal Papers

[J.1] M. Alsan and E. Telatar, Conditions for robustness of polar codes in the presence of channel mismatch, in preparation.

[J.2] M. Alsan and E. Telatar, The mismatched capacity of polar codes, submitted to *IEEE Journal on Selected Areas in Communications: Recent Advances In Capacity Approaching Codes*, Oct. 2014.

[J.3] M. Alsan, The symmetric convex ordering: A novel partial order for B-DMCs ordering the information sets of polar codes, submitted to *IEEE Transactions on Information Theory*, Aug. 2014.

[J.4] M. Alsan and E. Telatar, A simple proof of polarization and polarization for non-stationary channels, submitted to *IEEE Transactions on Information Theory*, Aug. 2014.

[J.5] M. Alsan, Extremality for Gallager's reliability function E_0, submitted to *IEEE Transactions on Information Theory*, Jan. 2014.

[J.6] M. Alsan and E. Telatar, Polarization improves E_0, *IEEE Transactions on Information Theory*, 60(5):2714–2719, May 2014.

[J.7] M. Alsan, Extremal channels of Gallager's E_0 under the basic polarization transformations, *IEEE Transactions on Information Theory*, 60(3):1582–1591, Mar. 2014.

Conference Papers

[C.1] M. Alsan and E. Telatar, Polarization as a novel architecture to boost the classical mismatched capacity of B-DMCs, *Proceedings of the IEEE Information Theory Workshop*, Hobart, Nov. 2014.

[C.2] M. Alsan, Universal polar decoding with channel knowledge at the encoder, *Proceedings of the IEEE Information Theory Workshop*, Hobart, Nov. 2014.

[C.3] M. Alsan and E. Telatar, A simple proof of polarization and polarization for non-stationary channels, *Proceedings of the IEEE International Symposium on Information Theory*, Honolulu, Jul. 2014.

[C.4] M. Alsan, A novel partial order for the information sets of polar codes over B-DMCs, *Proceedings of the IEEE International Symposium on Information Theory*, Honolulu, Jul. 2014.

[C.5] M. Alsan, A lower bound on achievable rates by polar codes with mismatch polar decoding, *Proceedings of the IEEE Information Theory Workshop*, Sevilla, Sept. 2013.

[C.6] M. Alsan and E. Telatar, Polarization improves E_0, *Proceedings of the IEEE International Symposium on Information Theory*, Istanbul, Jul. 2013.

[C.7] M. Alsan, Properties of the polarization transformations for the likelihood ratios of symmetric B-DMCs, *13th Canadian Workshop on Information Theory*, Toronto, Jun. 2013.

[C.8] M. Alsan, Performance of mismatched polar codes over BSCs, *International Symposium on Information Theory and its Applications*, Honolulu, Oct. 2012.

[C.9] M. Alsan, Extremality properties for Gallager's random coding exponent. *Proceedings of the IEEE International Symposium on Information Theory*, Boston, Jul. 2012.