UNIVERSALITY FOR MULTI-TERMINAL PROBLEMS VIA SPATIAL COUPLING

A Dissertation

by

ARVIND YEDLA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Electrical Engineering

UNIVERSALITY FOR MULTI-TERMINAL PROBLEMS VIA SPATIAL COUPLING

A Dissertation

by

ARVIND YEDLA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Henry Pfister
	Krishna Narayanan
Committee Members,	Thomas Schlumprecht
	Srinivas Shakkottai
Head of Department,	Costas Georghiades

August 2012

Major Subject: Electrical Engineering

ABSTRACT

Universality for Multi-terminal Problems

via Spatial Coupling. (August 2012)

Arvind Yedla, B.Tech., IIT Madras

Co–Chairs of Advisory Committee: Dr. Henry Pfister
Dr. Krishna Narayanan

Consider the problem of designing capacity-achieving codes for multi-terminal communication scenarios. For point-to-point communication problems, one can optimize a single code to approach capacity, but for multi-terminal problems this translates to optimizing a single code to perform well over the entire region of channel parameters. A coding scheme is called *universal* if it allows reliable communication over the entire achievable region promised by information theory.

It was recently shown that terminated low-density parity-check convolutional codes (also known as. spatially-coupled low-density parity-check ensembles) have belief-propagation thresholds that approach their maximum a-posteriori thresholds. This phenomenon, called "threshold saturation via spatial-coupling", was proven for binary erasure channels and then for binary memoryless symmetric channels. This approach provides us with a new paradigm for constructing capacity approaching codes. It was also conjectured that the principle of spatial coupling is very general and that the phenomenon of threshold saturation applies to a very broad class of graphical models.

In this work, we consider a noisy Slepian-Wolf problem (with erasure and binary symmetric channel correlation models) and the binary-input Gaussian multiple access channel, which deal with correlation between sources and interference at the receiver

respectively. We derive an area theorem for the joint decoder and empirically show that threshold saturation occurs for these multi-user scenarios. We also show that the outer bound derived using the area theorem is tight for the erasure Slepian-Wolf problem and that this bound is universal for regular LDPC codes with large left degrees. As a result, we demonstrate near-universal performance for these problems using spatially-coupled coding systems.

To my family

ACKNOWLEDGMENTS

I would like to thank my advisors Henry Pfister and Krishna Narayanan for their constant support and encouragement through the course of my graduate studies. I am also grateful to all my labmates and friends.

I would like to extend my thanks to the ECE staff for their excellent administrative support and to Thomas Schlumprecht and Srinivas Shakkottai for having served on my dissertation committee.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

Coding theory deals with the problem of reliable transmission of data from one point to another through an unreliable medium (known as the channel). This is accomplished by adding *redundancy* to the data in a systematic manner at the transmitter. The receiver uses this redundancy to recover the transmitted data, which has been distorted by the channel. Roughly speaking, the rate of transmission is the normalized *amount of information* transmitted using the coding scheme. In his seminal paper in 1948, Shannon showed the existence of a maximal rate of transmission, called the channel capacity, for a given channel [1]. Since then, much research has been focused on designing coding schemes in order to achieve the channel capacity. We review some preliminary concepts for point-to-point communication in Section A.

In this dissertation, we are interested in designing codes which perform well for multi-terminal problems. Multi-terminal communications involve the communication scenarios which have multiple transmitters and multiple receivers. The focus of this work is in the case when there are multiple transmitters and a single receiver. We consider two important multi-terminal problems which are described in Sections 1 and 2. A notion of *universality* naturally arises in this context, which is not present in point-to-point communication scenarios. This is discussed in Section C. The aim of this dissertation is to design practical coding schemes which are universal.

This dissertation follows the style of *IEEE Trans. on Information Theory*.

A. Point-to-Point Communication

The simplest communication model is that of point-to-point communication. This problem is modeled as shown in Fig. 1. The encoder output alphabet is the same

Fig. 1. A simple block diagram of a point-to-point communication system.

as the channel input alphabet (without loss of generality). We define the different blocks used in Fig. 1.

Definition I.1 (Source)**.** The source is modeled as a discrete time random process. As we are concerned with channel coding, we assume that the source outputs are independent and uniformly distributed over the source alphabet. The source output alphabet used throughout this work is $\mathcal{X} = \mathbb{F}_2$.[1]

Definition I.2 ((n, k) Binary Code)**.** An (n, k) binary (block) code $\mathcal{C}$ is a subset of $\mathcal{X}^n$, with $|\mathcal{C}| = 2^k$. The elements of the code are called codewords and n is called the block-length.

Definition I.3 (Rate)**.** The rate of an (n, k) binary code $\mathcal{C}$ is defined to be $R = \frac{k}{n}$.

Definition I.4 (Encoder)**.** An encoder is a bijective map from $\mathcal{X}^k$ to $\mathcal{C}$. The encoder maps k source bits into a codeword of length n, which is transmitted over the channel. Note that an encoder specifies the code completely and we say that the rate of an encoder is $R = \frac{k}{n}$.

[1]Sometimes it is convenient to let the alphabet be $\{\pm 1\}$ instead of $\mathbb{F}_2$, with the map $0 \mapsto +1$ and $1 \mapsto -1$, with addition over $\mathbb{F}_2$ replaced by standard multiplication on $\{\pm 1\}$. We shall use these alphabets interchangeably throughout this dissertation.

Definition I.5 (Channel)**.** A channel is a triple $(\mathcal{X}, \mathcal{Y}, p_{Y|X})$, where $\mathcal{X}$ is the input alphabet of the channel and $\mathcal{Y}$ is the output alphabet. Here, $p_{Y|X}$ is the conditional density of the channel output $Y \in \mathcal{Y}$ given the input $X \in \mathcal{X}$. The triple $(\mathcal{X}^n, \mathcal{Y}^n, p_{\mathbf{Y}|\mathbf{X}})$ is used to denote n uses of the channel. A channel is said to be memoryless, if

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) = \prod_{i=1}^{n} p_{Y|X}(Y_i|X_i),$$

where $\mathbf{X} = (X_1, \cdots, X_n) \in \mathcal{X}^n$ and $\mathbf{Y} = (Y_1, \cdots, Y_n) \in \mathcal{Y}^n$. A binary memoryless channel is said to be symmetric if there exists an involution $\imath : \mathcal{Y} \to \mathcal{Y}$ such that

$$p_{Y|X}(y| + 1) = p_{Y|X}(\imath(y)| - 1).$$

Definition I.6 (Log-Likelihood Ratio)**.** Consider a binary memoryless channel given by $(\mathcal{X}, \mathcal{Y}, p_{Y|X})$. The log-likelihood function is defined by

$$l(y) = \ln \frac{p_{Y|X}(y| + 1)}{p_{Y|X}(y| - 1)}.$$

Let L be the associated random variable, defined as $L = l(Y)$, and a be the conditional density of L, given $X = 1$ (called the L-density). Any binary memoryless symmetric channel (BMSC) can be equivalently represented by its L-density, denoted by a_{BMSC}.

Definition I.7 (Sufficient statistic)**.** Consider a channel $(\mathcal{X}, \mathcal{Y}, p_{Y|X})$ and a function $f(\cdot)$. We say $Z = f(Y)$ is a sufficient statistic for X given Y, if X is independent of Y given Z. For a binary memoryless channel, the log-likelihood ratio L is a sufficient statistic for decoding [2].

Definition I.8 (Decoder)**.** The decoder is a map $\hat{\boldsymbol{x}} : \mathcal{Y}^n \to \mathcal{X}^n$. The output of the decoder is used to make an estimate of the source output. The bit-wise maximum a

posteriori (MAP) decoder is given by $\hat{\boldsymbol{x}} = \left(\hat{x}_i^{\mathrm{MAP}}(\boldsymbol{y})\right)_{i=1}^{n}$, where

$$\hat{x}_i^{\mathrm{MAP}}(\boldsymbol{y}) = \underset{x_i=\pm 1}{\operatorname{argmax}} \sum_{\sim x_i} \left(\prod_j p_{Y|X}(y_j|x_j) \right) \mathbb{1}_{\{\boldsymbol{x}\in\mathcal{C}\}}. \tag{1.1}$$

We write $\sum_{\sim x_i}$ to indicate summation over all components of $\boldsymbol{x}$ except x_i. This decoder is optimal in terms of minimizing the probability of bit error at the receiver. Note that the output of this decoder need not be a codeword.

Definition I.9 (Entropy)**.** Let X be a discrete random variable with probability mass function $p_X(x)$. The entropy of X is a measure of uncertainty in the random variable and is given by

$$H(X) = -\sum_i p(x_i) \log p(x_i).$$

For continuous random variables with a PDF $p_X(x)$, the differential entropy is given by

$$h(X) = -\int_{-\infty}^{\infty} p(x) \log p(x) \mathrm{d}x.$$

We also define the binary entropy function $h_2(p) = -p \log p - (1-p)\log(1-p)$.

Definition I.10 (Mutual Information)**.** Let X and Y be two random variables. The mutual information between X and Y, denoted by $I(X;Y)$ is given by

$$I(X;Y) = H(X) - H(Y|X).$$

Definition I.11 (Channel Capacity)**.** The capacity of a channel $(\mathcal{X}, \mathcal{Y}, p_{Y|X})$ is the maximal rate at which information can be transmitted reliably through the channel. In other words there exists an encoder/decoder pair, using which reliable transmission

is possible at rates up-to the channel capacity. It is denoted by C and is given by

$$C = \sup_{p(X)} I(X;Y),$$

where $I(X;Y)$ is the mutual information between X and Y. For channels that depend on a single parameter, we denote the capacity by $C(\alpha)$.

Definition I.12 (Channel degradation)**.** Consider two memoryless channels specified by transition probabilities $p_{Y|X}$ and $p_{Z|X}$ respectively. We say that the $p_{Z|X}$ is degraded with respect to the first channel if

$$p_{Y,Z|X}(y,z|x) = p_{Y|X}(y|x)p_{Z|Y}(z|y).$$

A through discussion of channel degradation can be found in [2, p. 204].

Throughout the sequel, we consider families of channels which are characterized by a single parameter α. This implies that given a rate R there exists an encoder/decoder pair for which reliable transmission is possible for all channel parameters $\{\alpha|C(\alpha) \geq R\}$. If the channel is degraded with respect to α, we can define a threshold on the channel parameter, denoted by $\alpha^*(R) = C^{-1}(R)$, such that reliable communication is possible over all channels which are better than α^* using codes of rate R. This set is known as the set of achievable channel parameters (ACP). The goal of channel coding is to design low-complexity encoding/decoding schemes which enable transmission at channel parameters close to α^*, with an arbitrary low probability of error.

Example I.1 (Binary Erasure Channel - BEC(α))**.** This channel models the situation where the transmitted bits may be lost (erasures) but never corrupted. The channel model is shown in Fig. 2. Here α denotes the probability of erasure and $\mathcal{Y} = \{0, 1, ?\}$. The capacity of this channel is given by $C(\alpha) = 1 - \alpha$ and the L-density associated

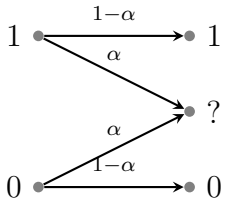

Fig. 2. A model of a BEC is shown above. The channel input is erased with probability α. This channel does not make any errors.

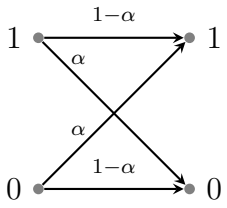

Fig. 3. A model of a BSC. The input bits are flipped with probability α.

with this channel is given by $\mathsf{a}_{\mathrm{BEC}(\alpha)}(x) = \alpha\Delta_0(x) + (1-\alpha)\Delta_{+\infty}(x)$, where $\Delta_a(x)$ is the Dirac delta function at $x = a$.

Example I.2 (Binary Symmetric Channel - BSC(α))**.** This channel is a generic model for binary-input memoryless channels where hard decisions are made at the receiver front end. The channel model is shown in Fig. 3. Here α denotes the crossover probability and $\mathcal{Y} = \{0, 1\}$. The capacity of this channel is given by $C(\alpha) = 1 - h_2(\alpha)$, where $h_2(\cdot)$ denotes the binary entropy function. The L-density associated with this channel is given by $\mathsf{a}_{\mathrm{BSC}(\alpha)}(x) = \alpha\Delta_{-\ln\frac{1-\alpha}{\alpha}}(x) + (1-\alpha)\Delta_{\ln\frac{1-\alpha}{\alpha}}(x)$.

Example I.3 (Additive White Gaussian Noise Channel - BIAWGNC(α))**.** This channel adds an additive noise to the transmitted data i.e., $Y = X + Z$, where Z is a Gaussian random variable with zero mean and variance σ^2. The L-density of this channel is given by $\mathcal{N}(2/\sigma^2, 4/\sigma^2)$. The noise variance is the unique value σ^2 such

that

$$\int_{-\infty}^{\infty} \mathsf{a}_{\mathrm{BAWGNC}}(x) \log_2(1 + \mathrm{e}^{-x}) \mathrm{d}x = \alpha.$$

Here $\mathcal{X} = \{\pm 1\}$ and $\mathcal{Y} = \mathbb{R}$. The channel model is shown in Fig. 4. The capacity of this channel cannot be expressed in elementary form.

B. Multi-user Communication

Many real world communication scenarios involve multi-user communication (wireless sensor networks, cellular systems, peer-to-peer networks etc.). Communication problems with more than one user require additional design considerations when compared to point-to-point communication strategies. This is due to the additional constraints of correlation between the sources and interference at the receiver. To better understand the additional design constraints, we consider the problems of sensor reachback and uplink in cellular systems, which deal with correlation between sources and interference at the receiver respectively.

1. Sensor Networks

Wireless sensor networks have become very popular in recent years and are being increasingly used in many commercial applications. A good survey of the problems involved with designing sensor networks can be found in [3, 4]. A sensor network

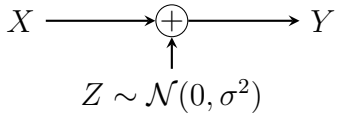

Fig. 4. The binary input AWGN channel is shown in the figure. The channel adds Gaussian noise to the input. The input alphabet is $\{\pm 1\}$.

typically has several transceivers (also called nodes), each of which has one or several sensors. The task of these sensor nodes is to collect measurements, encode them, and transmit them to some data collection points. The topology of sensor networks varies widely with the application, but typically the data from all the nodes is transmitted to a central node, also known as a gateway node, before further processing is done on the data. The implied communication problem is often referred to as the sensor reachback problem [5]. There are many constraints on the size and cost of the networks, so the nodes have limited computational capabilities, communication bandwidth etc. Hence the nodes have to perform distributed encoding, despite having to transmit correlated data. One of the main goals in the area of wireless sensor networks is to reduce the amount of transmitted data by taking advantage of the correlation between the sources. In many cases, there is generally a medium access control protocol in place, that eliminates interference between the different nodes. In this case, one can assume that each node transmits through an independent channel, from the same channel family. This problem is a noisy version of the celebrated Slepian-Wolf problem.

The SW problem was introduced and solved in the landmark paper [6] for noiseless channels, and shows that the optimal coding scheme suffers no loss in performance (in terms of rate) even in the absence of communication between the various encoders. The first practical SW coding scheme was introduced by Wyner and is based on linear error-correcting codes [7]. Chen et al. related the SW (distributed source coding) problem to channel coding via an equivalent channel describing the source correlation [8,9]. Using this observation they used density evolution to design LDPC coset codes that approach the SW bound. Distributed source coding using syndromes (DISCUS) also provides a practical method to transmit information for this problem when the encoding rates are restricted to the corner points of the rate region [10].

For transmision over noisy channels, separation between source and channel coding is known to be optimal when the channel state is known at the transmitter [5]. When the channel state is unknown, it is still desirable to take a joint source-channel coding (JSCC) approach (via direct channel coding and joint decoding at the receiver). The main reason is that separate source and channel coding requires compression of the sources to their joint entropy prior to channel encoding. After that, the variation in one channel's parameter cannot be offset by variation in the other channel. Further advantages of JSCC, over separated source coding and channel coding, are discussed further in [11]. The performance of concatenated LDGM codes has been studied in [12] and that of Turbo codes in [11]. Serially concatenated LDPC and convolutional codes were also considered in [13], where the outer LDPC code is used for distributed source coding.

Another interesting line of research in the area of sensor networks is the sensor location problem. The sensor locations are optimized in order to collect the most relevant data. A possibility of using moving sensors is present in a variety of applications, including air pollution estimation, traffic surveillance etc. [4]. A natural consequence of this is the variation in channel conditions as a result of sensor mobility. As a result, it may be unreasonable to assume that transmitters have detailed channel state information.

2. Cellular Systems

An important development in the past few decades in communications has been the evolution of cellular systems. There has been an abundance of scientific research in developing schemes for efficient bandwidth utilization and increasing the system capacity. A common multiple access scheme is direct sequence code-division multiple-access (DS-CDMA). Iterative multiuser detection (MUD) for CDMA systems using

forward error control coding has gained a lot of popularity in recent years (see [14–17] and the references therein). By studying the interaction between the MUD and the error-control code, transmission schemes that achieve a significant portion of the multiple-access channel capacity have been introduced [14, 18–22]. Many low complexity interference cancellation (IC) schemes have also been proposed as alternatives to the optimal MUD. A unified framework based on the factor graph representation was introduced in [16] to study the performance of the IC schemes using density evolution (DE) [23]. The DE analysis (using Gaussian approximation) was used to study the performance of IC schemes with convolutional codes, turbo codes and LDPC codes [16, 17]. The CDMA system load is shown to have a threshold and the system spectral efficiency is discussed in [24].

A crucial aspect of these systems is the design of the MUD to combat the interference at the receiver. To understand the aspects of code design for this problem, we look at the simple case of two users transmitting over a multiple access channel (MAC). When the received signal is corrupted by Gaussian noise, this channel is known as the Gaussian MAC. This channel can be characterized by a capacity region [25], and has been extensively studied in the literature. Many optimization schemes have been proposed to design good codes for this problem.

The corner points of the capacity region are known to be achievable by combining successive cancellation at the decoder with single-user codes [26]. This method can also be leveraged to achieve any point on the dominant face by time sharing or rate splitting [27]. The problem of designing good LDPC degree distributions was studied in [28] using density evolution (DE), where the authors design good LDPC codes for a few points in the achievable region (in terms of rate). Another approach was shown in [29] for the case when both users have the same transmit power, using EXIT charts.

These optimization procedures exploit knowledge of the channel gains to design

good codes. However, in practical scenarios the channel gains cannot be known non-causally at the transmitter (for example, a fading channel). So, it is desirable to fix the rate pair for transmission and view the capacity region in terms of the achievable channel gains for that rate pair.

C. Universality

Another interesting demarcation between multi-user communication and point-to-point communication is in the notion of channel degradation. Loosely speaking, a capacity achieving code designed for a particular channel condition is able to perform well for all channel conditions which are *better*. For point-to-point communications, this set is also the ACP set for that rate. This is no longer true for multi-user communication problems. This is seen in Fig. 5 (here we assume that a bigger α means that the channel is *better*). For fixed user code rates, reliable communication is theoretically possible over a wide range of channel conditions [25] and we note that the ACP set extends to a ACP region (ACPR) in this case. In the context of communication over parallel channels, the ACPR was called the reliable channel region [30].

Definition I.13 (Universal codes)**.** A code is called *universal* if it provides good performance for all system parameters that do not violate theoretical limits.

Remark I.1. This designation neglects the fact that the receiver is assumed to have channel state information and is based on the standard assumption that the receiver can estimate the channel state with negligible pilot overhead.

While irregular LDPC codes can be optimized to approach capacity for any particular channel condition, their performance can deteriorate markedly as the channel conditions change. So, codes which are robust to variation in channel conditions are

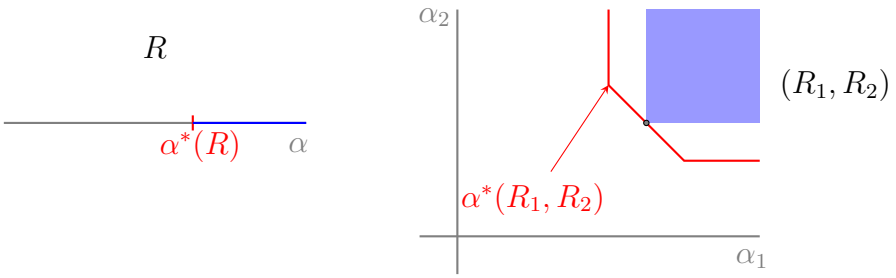

Fig. 5. An illustration of channel degradation in point-to-point communication and multi-user systems. Here we assume that bigger α implies a better channel. The ordering for this multi-user case is given by $(\alpha_1, \alpha_2) \succeq (\beta_1, \beta_2)$ iff $\alpha_1 \geq \beta_1$ and $\alpha_2 \geq \beta_2$.

desirable because they minimize the outage probability for quasi-static channels (e.g., when a probability distribution is assigned to the set of possible channel parameters).

Thus, universal codes can be expected to provide performance gains when used in sensor networks. Due to fading in wireless channels, the problem of unknown channel state at the transmitter naturally arises in this context, motivating the use of universal codes.

D. Outline

This dissertation is organized as follows. In Chapter II, we present a review of the background material for this dissertation. We first review linear codes which can be represented by sparse graphs and the belief-propagation algorithm, which is used for decoding. We then briefly review density evolution (DE), which presents a powerful analysis tool to characterize the performance of graph based codes. Generalized EXIT curves, which present an important connection between BP decoding and MAP decoding via an Area Theorem, are then introduced . The remainder of Chapter II introduces the multiple-terminal problem models used throughout this work and extends DE for these problems.

Chapter III discusses the MAP performance of LDPC codes for these channels and discusses a performance bound by extending the notion of GEXIT curves for these problems. Chapter IV reviews spatially-coupled codes and discusses the performance of spatially-coupled codes for multi-terminal problems in Chapter V. Some conclusions are provided in Chapter VI.

CHAPTER II

BACKGROUND

We describe the various terms and notation used in this work. We liberally use notation and definitions from [2].

A. LDPC Codes

LDPC codes are a class of linear codes introduced by Gallager in [31]. The message-passing rules which later became belief propagation was also introduced. Tanner generalized the notion of representing linear codes in terms of a bipartite graph [32]. Mackay rediscovered LDPC codes in [33] and noticed the advantages of sparse block codes. Since then many tools and analysis techniques have been introduced to understand and improve the performance of LDPC codes. Of particular note is the introduction of irregular LDPC codes by Luby et al. in [34, 35] and the density evolution (DE) algorithm introduced by Richardson et al. in [36].

EXIT charts were first introduced by ten Brink [37] as a visualization of BP decoding. For the BEC, these charts accurately represent the DE analysis. GEXIT functions were introduced in [38] as a natural generalization of EXIT charts to general channels. These functions fulfill the so-called Area Theorem, thereby giving an upper bound on the MAP performance of iterative decoding systems. This upper bound can be shown to be tight for the case of the BEC [2, Theorem 3.120]. We now proceed with some basic definitions and examples. In this chapter, all vectors are assumed to be column vectors.

Definition II.1 (Binary Linear Code)**.** A binary linear code of length n is a subspace

of $\mathbb{F}_2^n$, with dimension k. So, we can write

$$\mathcal{C} = \{\boldsymbol{x} | H\boldsymbol{x} = \mathbf{0}, \boldsymbol{x} \in \mathbb{F}_2^n\},$$

for some $H \in \mathbb{F}_2^{m \times n}$, such that $\text{rank}(H) = n - k$, $m \geq n - k$. Note that the rows of H need not be linearly independent. The matrix H is called the parity-check matrix of the code $\mathcal{C}$. Note that the parity-check matrix of a code is not unique. Henceforth, we shall refer to a binary linear code $\mathcal{C}$ in terms of an associated parity-check matrix H. Alternately the code $\mathcal{C}$ can be represented in terms of a generator matrix G as

$$\mathcal{C} = \{\boldsymbol{x}^T G, \boldsymbol{x} \in \mathbb{F}_2^k\}.$$

Definition II.2 (Tanner Graph)**.** The tanner graph associated with a parity-check matrix H is a bipartite graph. It has n variable nodes corresponding to the components of the codeword (which correspond to the columns of H) and m check nodes corresponding to the rows of H. A check node j is connected to variable node i if $H_{ji} = 1$. We use the notation ∂i to denote the set of neighbors of i. We can also associate a Tanner graph with a generator matrix G. The rows of G correspond to the information nodes and the columns correspond to the generator nodes.

Example II.1 (Tanner Graph)**.** The Tanner graph corresponding to a parity-check matrix H is shown in Fig. 6.

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

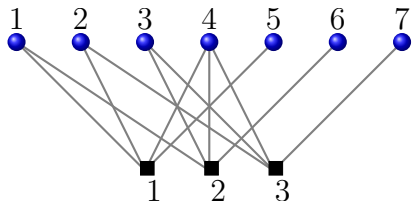

Fig. 6. The Tanner graph associated with the parity-check matrix of a Hamming code is shown above.

Definition II.3 (Degree Profile)**.** Consider the Tanner graph associated with a parity-check matrix H. The degree of a node is defined to be the number of edges connected to that node. Let L_i (R_i) denote the fraction of variable (check) nodes of degree i. The normalized degree distributions from the node perspective are defined by

$$L(x) \triangleq \sum_{i=2}^{\mathtt{l}_{\max}} L_i x^i, \qquad R(x) \triangleq \sum_{i=2}^{\mathtt{r}_{\max}} R_i x^i,$$

where $\mathtt{l}_{\max}$ and $\mathtt{r}_{\max}$ are the maximum variable and check node degrees respectively. Define the degree profiles from the edge perspective:

$$\lambda(x) = \sum_{i=2}^{\mathtt{l}_{\max}} \lambda_i x^{i-1} \triangleq \frac{L'(x)}{L'(1)}, \qquad \rho(x) = \sum_{i=2}^{\mathtt{r}_{\max}} \rho_i x^{i-1} \triangleq \frac{R'(x)}{R'(1)}.$$

Note that λ_i (ρ_i) is the fraction of edges that connect to a variable (check) node of degree i. The inverse relationship is given by

$$L(x) = \frac{\int_0^x \lambda(z)\mathrm{d}z}{\int_0^1 \lambda(z)\mathrm{d}z}, \qquad R(x) = \frac{\int_0^x \rho(z)\mathrm{d}z}{\int_0^1 \rho(z)\mathrm{d}z},$$

and the design rate is given by

$$\mathtt{R}(\lambda, \rho) = 1 - \frac{\int_0^1 \rho(z)\mathrm{d}z}{\int_0^1 \lambda(z)\mathrm{d}z} = 1 - \frac{L'(1)}{R'(1)}.$$

Definition II.4 (Low-Density Parity-Check Code)**.** A linear block code is called a low-density parity-check (LDPC) code if it admits a *sparse* parity-check matrix. A generic Tanner graph representation of LDPC codes is shown in Fig. 7.

Definition II.5 (Low-Density Generator-Matrix Code)**.** A linear block code is called a low-density generator-matrix (LDGM) code if it admits a *sparse* generator matrix. A generic Tanner graph representation of LDGM codes is shown in Fig. 8.

Definition II.6 (The Ensemble LDPC(n, λ, ρ))**.** The bipartite graph has n variable

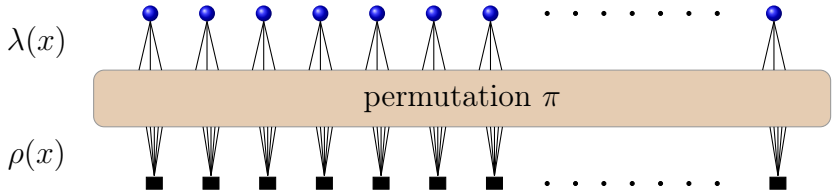

Fig. 7. A generic Tanner graph representation of regular LDPC codes is shown above.

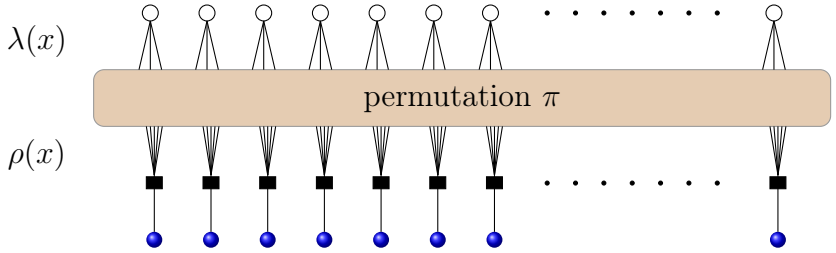

Fig. 8. A generic Tanner graph representation of regular LDGM codes. The white circles represent the punctured information bits of the LDGM code. The generator bits are represented by squares.

nodes. There are a total of $nL'(1)$ edges in the graph. So, the number of check nodes in the graph is $m = n\frac{L'(1)}{R'(1)}$. A node of degree i has i sockets from which the i edges emanate. Label the sockets on each side from the set $[nL'(1)]$ in an arbitrary but fixed way and let σ be a permutation on $[nL'(1)]$ (We use the notation $[n] \triangleq \{1, \cdots, n\}$). The i-th socket on the variable side is connected to the $\sigma(i)$-th socket on the check side. Define a probability distribution on the set of graphs generated this way by placing a uniform distribution on σ. We can associate a code with each such bipartite graph through the corresponding parity-check matrix H ($H_{ji} = 1$ if the i-th variable node is connected to the j-th check node an odd number of times). This ensemble of bipartite graphs is called LDPC(n, λ, ρ). The ensemble LDPC(λ, ρ) is the asymptotic version of LDPC(n, λ, ρ) (as n tends to infinity).

1. Belief Propagation Decoder

The belief propagation (BP) decoder is an iterative message-passing decoder. The decoder proceeds in rounds of message passing between the variable nodes and check nodes. The incoming messages at the check nodes are processed and forwarded to the variable nodes. The messages are then processed at the variable nodes and transmitted to the check nodes. This is one round of message passing. The incoming message to each node is the log-likelihood ratio of the conditional probability of that bit, and nodes process the messages as they were independent of all the other messages. The BP decoder employs a locally optimal processing rule at the variable and check nodes. This decoder is optimal (it performs the marginalization given in (1.1)) if the Tanner graph is a tree. It is sub-optimal when the Tanner graph has cycles. A more through discussion can be found in [2, Section 4.2]. The message passing rules at iteration ℓ for a check (variable) node of degree $\mathtt{l}$ ($\mathtt{r}$) are summarized below:

Check Node Update

$$\Phi^{(\ell)}(\mu_1, \cdots, \mu_{\mathtt{r}-1}) = \begin{cases} 0 & \ell = 0 \\ 2\tanh^{-1}\left(\prod_{i=1}^{\mathtt{r}-1} \tanh\frac{\mu_i}{2}\right) & \ell > 0 \end{cases}. \tag{2.1}$$

Variable Node Update

$$\Psi(\mu_0, \mu_1, \cdots, \mu_{\mathtt{l}-1}) = \mu_0 + \sum_{i=1}^{\mathtt{l}-1} \mu_i. \tag{2.2}$$

Following [2], μ_0 is used to represent the message from the channel, $\mu_1, \cdots, \mu_{\mathtt{l}-1}$ to denote the incoming messages for a variable node of degree $\mathtt{l}$ and $\mu_1, \cdots, \mu_{\mathtt{r}-1}$ to denote the incoming messages for a check node of degree $\mathtt{r}$.

Assume that transmission takes place over a BMS channel using a code $\mathcal{C}$. Let the received vector be $\boldsymbol{y}$. The channel message for bit i is given by the log-likelihood

ratio $l(y_i)$ and all other messages are initialized to 0. The decoder iterates using (2.1) and (2.2) until a predefined stopping criteria is reached. The BP decoder then outputs the a-posteriori log-likelihood ratio for each bit, given by

$$l_i = l(y_i) + \sum_{j \in \partial i} \mu_{j \to i},$$

where $\mu_{j \to i}$ denotes the message from check node j to variable node i.

2. Density Evolution

The transformation of the densities of the incoming messages under the operations in (2.1) and (2.2) are denoted by $\boxplus$ and $\circledast$ respectively (see discussion in [2, p. 181]). Also, if a is an L-density, we denote

$$\mathsf{a}^{\boxplus n} \triangleq \underbrace{\mathsf{a} \boxplus \mathsf{a} \boxplus \cdots \boxplus \mathsf{a}}_{n},$$

and likewise for $\mathsf{a}^{\circledast n}$. To study the performance of the BP decoder, one can simplify the analysis by making the following key observations [2, Section 4.3]:

1. For any BMS channel, the error probability of the BP decoder is independent of the transmitted codeword. So, without loss of generality, we can assume that the all-zero codeword ($\mathbf{0} \in \mathbb{F}_2^n$) is transmitted. This implies that one needs only to track one density (conditioned on the all-zero codeword) to study the performance of an LDPC code with BP decoding.

2. The performance of any code chosen uniformly at random from the ensemble $\mathrm{LDPC}(n, \lambda, \rho)$ concentrates around the ensemble average for large block-lengths. This enables us to analyze the ensemble average (averaged over all possible channel realizations and over $\mathrm{LDPC}(n, \lambda, \rho)$).

Making these simplifications, in the limit of infinite block-lengths, the following equation captures the performance of the ensemble LDPC(λ, ρ), for transmission over a channel described by $\mathsf{a}_{\mathrm{BMSC}}$:

$$\mathsf{a}^{(\ell)} = \mathsf{a}_{\mathrm{BMSC}} \circledast \lambda(\rho(\mathsf{a}^{(\ell-1)})), \ \ell \geq 1, \tag{2.3}$$

with $\mathsf{a}_0 = \mathsf{a}_{\mathrm{BMSC}}$, $\lambda(\mathsf{a}) = \sum_i \lambda_i \mathsf{a}^{\circledast(i-1)}$ and $\rho(\mathsf{a}) = \sum_i \rho_i \mathsf{a}^{\boxtimes(i-1)}$. The density of the messages emanating from the variable nodes (assuming that the all-zero codeword was transmitted) at iteration ℓ is given by $\mathsf{a}^{(\ell)}$. The above equation is known as the density evolution (DE) equation.

Example II.2. The DE equation for the BEC(α) case simplifies to a one-dimensional recursion

$$a^{(\ell+1)} = \alpha\lambda(1 - \rho(1 - a^{(\ell)})),$$

where $\lambda(x) = \sum_i \lambda_i x^{i-1}$ and $\rho(x) = \sum_i \rho_i x^{i-1}$.

The error functional is defined by

$$\mathfrak{E}(\mathsf{a}) \triangleq \int_{-\infty}^{0^-} \mathsf{a}(x)\,\mathrm{d}x + \frac{1}{2}\int_{0^-}^{0^+} \mathsf{a}(x)\,\mathrm{d}x. \tag{2.4}$$

The expected residual error probability after ℓ iterations for the ensemble LDPC(λ, ρ) is given by $\mathfrak{E}(\mathsf{a}_{\mathrm{BMSC}} \circledast L(\rho(\mathsf{a}^{(\ell-1)}))$.

3. GEXIT Curves and MAP Performance

This section is an informal introduction to GEXIT curves and a bounding technique on the MAP threshold. A formal discussion can be found in [2, Section 4.12]. GEXIT curves were introduced as a generalization of EXIT curves. An upper bound on the MAP threshold can be computed using GEXIT curves, by means of an area theorem.

Consider transmission over a BMS channel $(\mathcal{X}^n, \mathcal{Y}^n, p_{\boldsymbol{Y}|\boldsymbol{X}})$ which is characterized by a parameter α. Throughout this section, we assume that a smaller α implies a *better* channel and that $\alpha \in [0,1]$. To emphasize the dependence on the channel parameter, we denote the output by $\boldsymbol{Y}(\alpha)$. The GEXIT function is defined as

$$\mathsf{g}(\alpha) = \lim_{n\to\infty} \frac{1}{n} \frac{\partial}{\partial \alpha} \mathbb{E}\left[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha))\right].$$

This function satisfies an area theorem by definition i.e.,

$$\int_0^1 \mathsf{g}(\alpha) \mathrm{d}\alpha = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha))\right] = \mathtt{R}.$$

The MAP threshold is defined as

$$\alpha^{\mathrm{MAP}} = \inf\left\{\alpha : \liminf_{n\to\infty} \mathbb{E}\left[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha))\right] > 0\right\},$$

from which one can obtain

$$\int_{\alpha^{\mathrm{MAP}}}^1 \mathsf{g}(\alpha) \mathrm{d}\alpha = \mathtt{R}.$$

The GEXIT function is hard to compute in general and hence one typically uses the BP-GEXIT function denoted by g^{BP}, which can be computed using the fixed points of density evolution [39].

Example II.3. For transmission over a BEC(α), the BP-GEXIT function is given in parametric form for regular ensembles ($\lambda(x) = x^{\mathtt{l}-1}$, $\rho(x) = x^{\mathtt{r}-1}$) by

$$\mathsf{g}^{\mathrm{BP}}(\alpha) = \begin{cases} (\alpha, 0) & \alpha \in [0, \alpha^{\mathrm{BP}}) \\ (\alpha(x), L(1-\rho(1-x))) & x \in (x^{\mathrm{BP}}, 1] \end{cases},$$

where $\alpha(x) = x/\lambda(1-\rho(1-x))$ and x^{BP} is the unique minimum of $\alpha(x)$ and $\alpha^{\mathrm{BP}} = \alpha(x^{\mathrm{BP}})$.

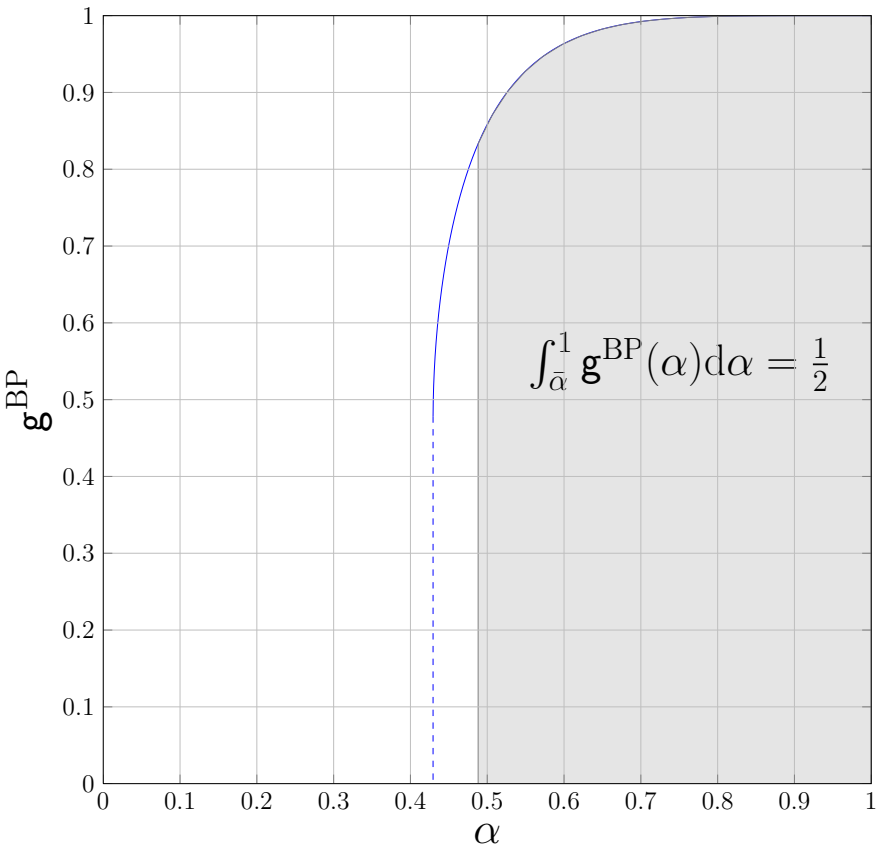

Fig. 9. The BP-GEXIT curve for the regular LDPC$(3, 6)$ ensemble for transmission over erasure channels. The MAP upper bound is given by $\bar{\alpha} \approx 0.4881$. This upper bound can be shown to be tight for erasure channels.

It can be shown that the BP-GEXIT function is a pointwise upper bound on the GEXIT function. So, we can upper bound the MAP threshold by $\bar{\alpha}$, where $\bar{\alpha}$ is the largest positive number such that

$$\int_{\bar{\alpha}}^{1} \mathrm{g}^{\mathrm{BP}}(\alpha)\mathrm{d}\alpha = \mathtt{R}.$$

The BP-GEXIT curve and the upper bound on the MAP threshold using this technique is shown in Fig. 9. These bounds can be shown to be tight for erasure-type channels.

B. The Noisy Slepian-Wolf Problem

Consider the problem of transmitting the outputs of two discrete memoryless correlated sources, $\left(U^{[1]}, U^{[2]}\right)$, to a central receiver through two independent discrete memoryless channels with capacities $C^{[1]}$ and $C^{[2]}$, respectively. The system model is shown in Fig. 10. We will assume that the channels belong to the same channel family, and that each channel can be parametrized by a single parameter α (e.g., the erasure probability for erasure channels). The two encoders are not allowed to communicate. Hence they must use independent encoding functions, which map k input symbols ($\boldsymbol{U}^{[1]}$ and $\boldsymbol{U}^{[2]}$) to n_1 and n_2 output symbols ($\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$), respectively. The rates of the encoders are given by $R_1 = k/n_1$ and $R_2 = k/n_2$. The decoder receives $(\boldsymbol{Y}^{[1]}, \boldsymbol{Y}^{[2]})$ and makes an estimate of $(\boldsymbol{U}^{[1]}, \boldsymbol{U}^{[2]})$.

The problem we consider is to design a graph-based code, for which a joint iterative decoder can successfully decode over a large set of channel parameters. For simplicity, we assume that both the encoders use identical codes of design rate R (i.e., $R = k/n, n_1 = n_2 = n$). Reliable transmission over a channel pair $(\alpha^{[1]}, \alpha^{[2]})$ is

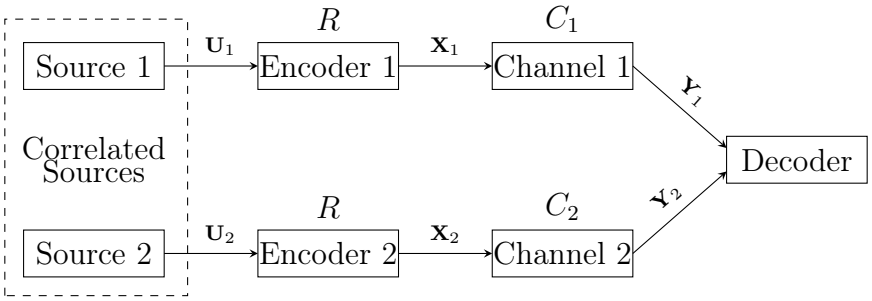

Fig. 10. System Model

possible as long as the SW conditions (2.5) are satisfied.

$$\begin{aligned} \frac{C^{[1]}(\alpha^{[1]})}{R} &\geq H\big(U^{[1]}\big|U^{[2]}\big) \\ \frac{C^{[2]}(\alpha^{[2]})}{R} &\geq H\big(U^{[2]}\big|U^{[1]}\big) \\ \frac{C^{[1]}(\alpha^{[1]})}{R} + \frac{C^{[2]}(\alpha^{[2]})}{R} &\geq H\big(U^{[1]}, U^{[2]}\big) \end{aligned} \tag{2.5}$$

For a given pair of rate-R encoding functions and a joint decoding algorithm, a pair of channel parameters $(\alpha^{[1]}, \alpha^{[2]})$ is *achievable* if the encoder/decoder combination can achieve an arbitrarily low error probability for the asymptotic limit as $k \to \infty$. We define the *achievable channel parameter region* (ACPR) as the set of all channel parameters which are achievable. Note that the ACPR is the set of all channel parameters for which successful recovery of the sources is possible for a fixed encoding rate pair (R, R). We also define the *SW-ACPR* as the set of all channel parameters $(\alpha^{[1]}, \alpha^{[2]})$ for which (2.5) is satisfied. The SW-ACPR for the erasure channel family is shown in Fig. 11.

1. Correlation Models

In this dissertation, we consider the following scenarios:

1. The channels are erasure channels and the source correlation is modeled through

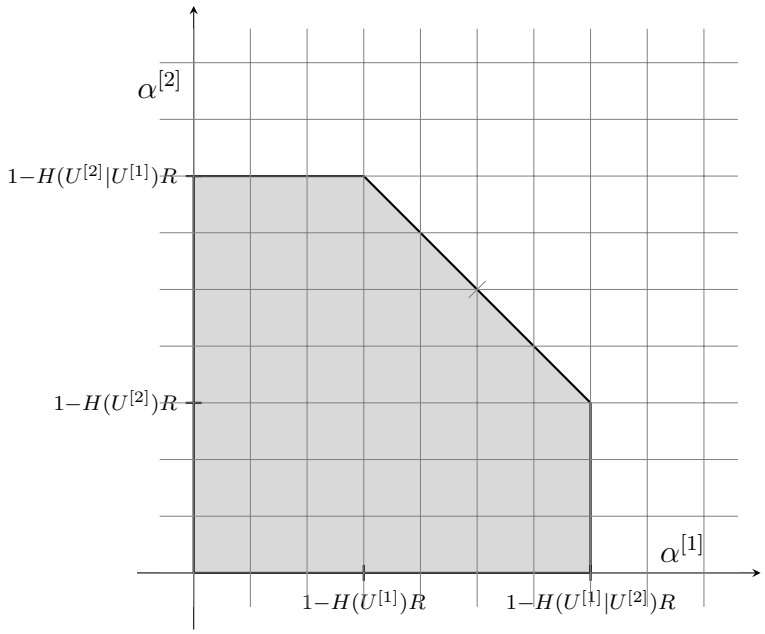

Fig. 11. The SW-ACPR for erasure channels, for a fixed rate pair (R, R)

erasures.

2. The channels are additive white Gaussian noise (AWGN) channels and the source correlation is modeled through a virtual correlation channel analogous to a binary symmetric channel (BSC).

These models might appear restrictive, but we believe they provide sufficient insight for the design of codes that perform well for arbitrary correlated sources and channels. Our analysis in Section 2 admits general correlation models and memoryless channels.

a. Erasure Correlation Model

The erasure system model is based on communication over binary erasure channels (BECs) and the source correlation is also modeled through erasures. Let Z be a Bernoulli-p random variable and X, X' be i.i.d. Bernoulli-$\frac{1}{2}$ random variables. The

sources $U^{[1]}$ and $U^{[2]}$ are defined by

$$\left(U^{[1]}, U^{[2]}\right) = \begin{cases} (X, X') & \text{if } Z = 0 \\ (X, X) & \text{if } Z = 1 \end{cases}.$$

We have $H\left(U^{[1]}|U^{[2]}\right) = H\left(U^{[2]}|U^{[1]}\right) = 1 - p$ and $H\left(U^{[1]}, U^{[2]}\right) = 2 - p$. This correlation model can be incorporated into the Tanner graph (see Fig. 12) at the decoder with the presence or absence of a check node between the source bits depending on the auxiliary random variable Z. Note that the decoder requires the realization of the random variable Z, for each source bit, as side information. Because of this requirement, one might consider this a toy model that is used mainly to gain a better understanding of the problem. Still, a very similar model was used recently to model internet file streaming from multiple sources [40].

This model can also be thought of as having two types of BSC correlation between the source bits (as described in the next section), one with parameter 0 and one with parameter 1. The correlation parameter p determines how many bits are correlated with parameter 1. The receiver knows which bits are correlated with parameter 1.

For a BEC correlation with probability p, there is a parity-check at the correlation node with probability p and with probability $1 - p$ there is no parity-check. Let $\zeta(\cdot)$ be the density transformation associated with these correlation nodes. Then $\zeta(\mathsf{a}) = (1 - p) + p\mathsf{a}$.

b. BSC Correlation Model

A more realistic model is the BSC/AWGN system model, where communication takes place over a binary-input additive white Gaussian-noise channel (BAWGNC) and the symmetric source correlation is defined in terms of a single parameter, namely $p = \Pr(U^{[1]} = U^{[2]})$. It is useful to visualize this correlation by the presence of an

auxiliary binary symmetric channel (BSC) with parameter $1-p$ between the sources. In other words, $U^{[2]}$ is the output of a BSC with input $U^{[1]}$ i.e., $U^{[2]} = U^{[1]} + Z$. Here Z is a Bernoulli-$(1-p)$ random variable and can be thought of as an *error*. Let $h_2(\cdot)$ denote the binary entropy function. Then, $H\big(U^{[1]}|U^{[2]}\big) = H\big(U^{[2]}|U^{[1]}\big) = h_2(p)$ and $H\big(U^{[1]}, U^{[2]}\big) = 1 + h_2(p)$.

This correlation model can be incorporated into the Tanner graph at the decoder (described in Section 2) as check nodes between the source bits, with a hidden node representing the auxiliary random variable Z (which outputs the constant log-likelihood ratio $\ln \frac{1-p}{p}$) attached to the check node. For this scenario, the decoder does not require any side information i.e., it does not need to know the realization of the auxiliary random variable Z.

The symmetry of the problem allows one to, without loss of generality, assume that user 1 transmits the all-zero codeword and the second user transmits a typical codeword (i.e., the fraction of ones equals the fraction of zeros as $k \to \infty$). Due to the constraints imposed by the correlation, the fraction of ones in the systematic part of the codeword is $1-p$. Density evolution proceeds with two types of messages (those connected to a variable node with transmitted value $+1$ and those connected to a variable node with transmitted value -1). By symmetry of the message passing rules [2, p. 210], we can factor out the sign for the messages connected to variable nodes with transmitted value -1. This sign can be factored into the correlation node (once again by the symmetry condition). The fraction of correlation nodes which are flipped is $1-p$. So, we introduce a parity-check at the correlation nodes which evaluates to a Bernoulli-p random variable. Then the density transformation operator associated with these correlation nodes is given by $\zeta(\mathsf{a}) = \mathsf{a}_{\mathrm{BSC}(p)} \circledast \mathsf{a}$. This simplification enables us to proceed with density evolution assuming the transmission of an all-zero codeword for both the users.

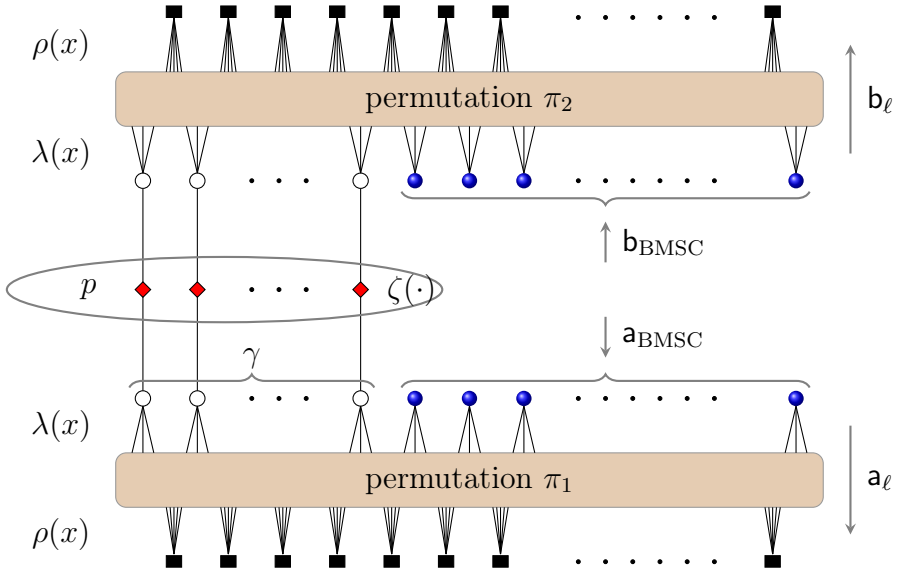

Fig. 12. The Tanner graph of a punctured systematic LDPC code with source correlation nodes is shown above.

2. Density Evolution

Assume that the sequences $\boldsymbol{U}^{[1]}$ and $\boldsymbol{U}^{[2]}$ are encoded using LDPC codes with a degree distribution pair (λ, ρ) and a punctured systematic encoder. Let the fraction of punctured (systematic) bits be $\gamma = \mathtt{R}(\lambda, \rho)$. The rate pair of the two codes after puncturing is (R, R), where

$$R = \frac{\mathtt{R}(\lambda, \rho)}{1 - \mathtt{R}(\lambda, \rho)}. \tag{2.6}$$

The Tanner graph [2] for the joint decoder is shown in Fig. 12. Codes 1 and 2 correspond to the bottom and top half of the graph. The codes are connected by correlation nodes attached to the punctured bits. The joint iterative decoder proceeds in rounds, by alternating one round of decoding for code 1 with one round of decoding for code 2. Let $\mathsf{a}^{(\ell)}$ and $\mathsf{b}^{(\ell)}$ denote the LLR density[1] of the messages

[1]Assuming that the transmission alphabet is $\{\pm 1\}$, the densities are conditioned on the transmission of a $+1$.

emanating from the variable nodes at iteration ℓ, corresponding to codes 1 and 2. The density evolution equations [2] can be written as follows

$$\begin{aligned}\mathsf{a}^{(\ell+1)} &= \left[\gamma\zeta\Big(L\left(\rho(\mathsf{b}^{(\ell)})\right)\Big) + (1-\gamma)\mathsf{a}_{\mathrm{BMSC}}\right] \circledast \lambda(\rho(\mathsf{a}^{(\ell)})) \\ \mathsf{b}^{(\ell+1)} &= \left[\gamma\zeta\Big(L\left(\rho(\mathsf{a}^{(\ell)})\right)\Big) + (1-\gamma)\mathsf{b}_{\mathrm{BMSC}}\right] \circledast \lambda(\rho(\mathsf{b}^{(\ell)})),\end{aligned} \tag{2.7}$$

where $\lambda(\mathsf{a}) = \sum_i \lambda_i \mathsf{a}^{\circledast(i-1)}$, $L(\mathsf{a}) = \sum_i L_i \mathsf{a}^{\circledast(i-1)}$, $\rho(\mathsf{a}) = \sum_i \rho_i \mathsf{a}^{\boxplus(i-1)}$, $\mathsf{a}_{\mathrm{BMSC}}$ and $\mathsf{b}_{\mathrm{BMSC}}$ are the densities of the log-likelihood ratios received from the channel. The function ζ at the correlation nodes depends on the equivalent channel corresponding to the correlation model, as described in [8]. Although one cannot assume that the all-zero codeword is sent simultaneously by both users, one can show that this DE recursion suffices for typical message pairs as defined previously. The fixed points of DE are the tuples $(\mathsf{a}_{\mathrm{BMSC}}, \mathsf{b}_{\mathrm{BMSC}}, \mathsf{a}, \mathsf{b})$ which satisfy

$$\begin{aligned}\mathsf{a} &= \left[\gamma\zeta\Big(L\left(\rho(\mathsf{b})\right)\Big) + (1-\gamma)\mathsf{a}_{\mathrm{BMSC}}\right] \circledast \lambda(\rho(\mathsf{a})) \\ \mathsf{b} &= \left[\gamma\zeta\Big(L\left(\rho(\mathsf{a})\right)\Big) + (1-\gamma)\mathsf{b}_{\mathrm{BMSC}}\right] \circledast \lambda(\rho(\mathsf{b})).\end{aligned} \tag{2.8}$$

The residual error probability at iteration ℓ, $(e_1^{(\ell)}, e_2^{(\ell)})$, is computed using the error functional $\mathfrak{E}(\cdot)$ defined in (2.4):

$$\begin{aligned}e_1^{(\ell)} &= \mathfrak{E}\left(\left[\gamma\zeta\Big(L\left(\rho(\mathsf{b}^{(\ell-1)})\right)\Big) + (1-\gamma)\mathsf{a}_{\mathrm{BMSC}}\right] \circledast L(\rho(\mathsf{a}^{(\ell-1)}))\right) \\ e_2^{(\ell)} &= \mathfrak{E}\left(\left[\gamma\zeta\Big(L\left(\rho(\mathsf{a}^{(\ell-1)})\right)\Big) + (1-\gamma)\mathsf{b}_{\mathrm{BMSC}}\right] \circledast L(\rho(\mathsf{b}^{(\ell-1)}))\right).\end{aligned}$$

For two residual error probabilities (e_1, e_2) and $(\tilde{e}_1, \tilde{e}_2)$, we define $(e_1, e_2) \preceq (\tilde{e}_1, \tilde{e}_2)$ iff $e_1 \leq \tilde{e}_1$ and $e_2 \leq \tilde{e}_2$.

Consider the line $\alpha^{[2]} = \theta\alpha^{[1]}$, for some $\theta \in [0, +\infty)$. The BP threshold along the

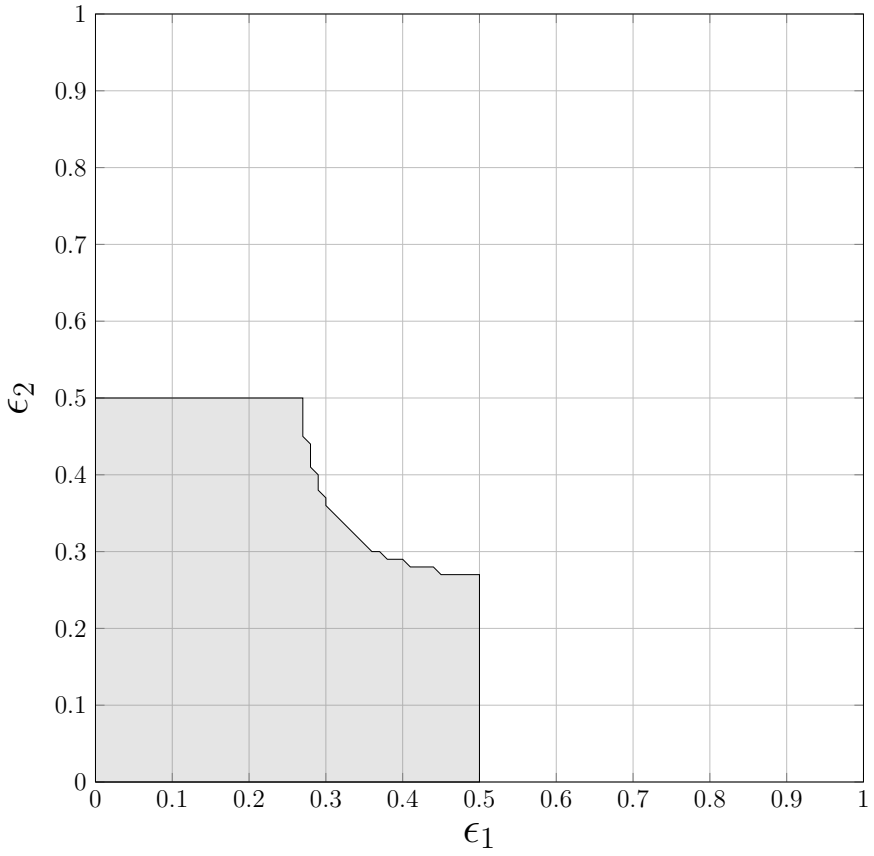

Fig. 13. The DE boundary for the punctured LDPC$(4, 6)$ ensemble is shown above for the SWE problem. The channel parameters are erasure probabilities.

line is defined by

$$\alpha^{\mathrm{BP}}(\lambda, \rho, \theta) = \inf\big\{\alpha : \text{ The fixed point equation (2.13) has a solution } (\mathsf{a}, \mathsf{b}) \neq (\Delta_{+\infty}, \Delta_{+\infty})\big\}.$$

The set of all points $(\alpha, \theta\alpha)$ such that $\alpha \leq \alpha^{\mathrm{BP}}(\lambda, \rho, \theta)$ is called the BP-ACPR and its boundary is called the DE boundary. The DE boundary for the punctured LDPC$(4, 6)$ ensemble is shown in Fig.s 13 and 14 for the SWE problem and the BSC correlation model respectively.

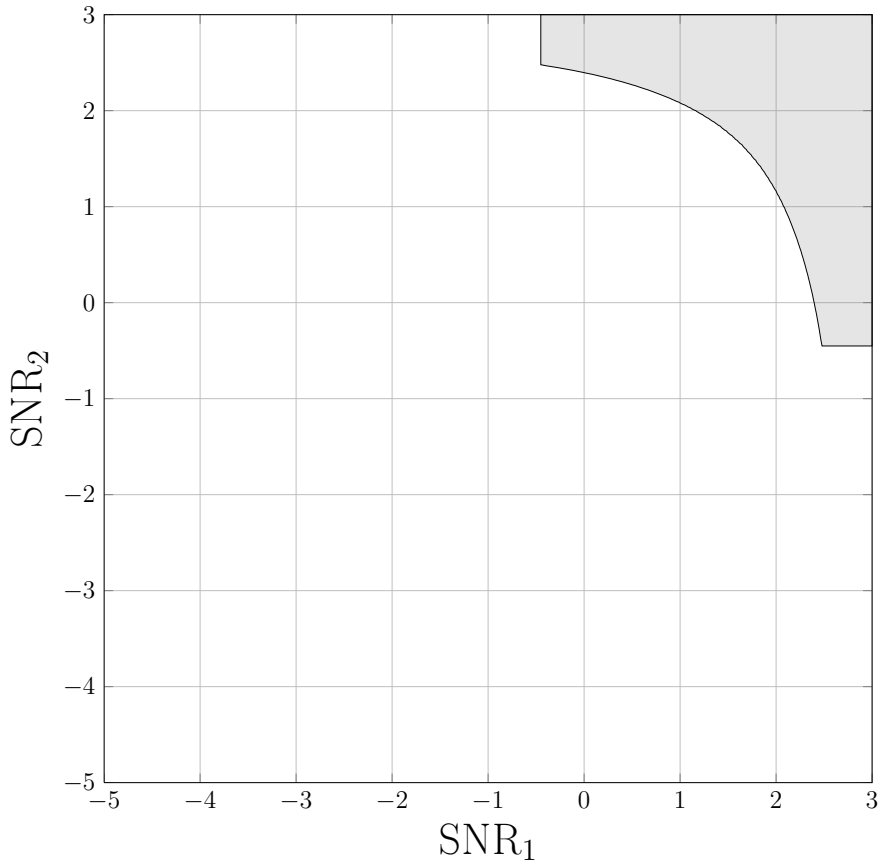

Fig. 14. The DE boundary for the punctured LDPC$(4, 6)$ ensemble for the BSC correlation model is shown above. The figure is plotted with respect to the signal-to-noise ratio and not the parameter α.

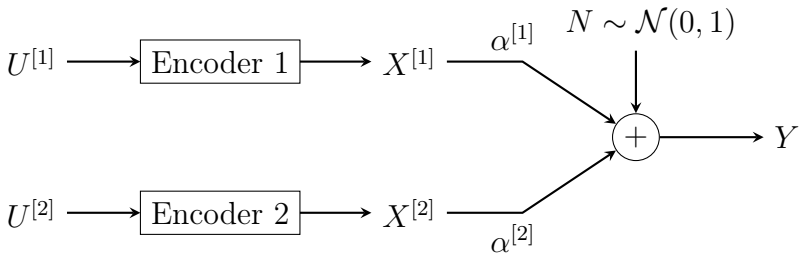

Fig. 15. The Gaussian MAC

C. The Gaussian Multiple-Access Channel

Consider the problem of transmitting the outputs of two independent discrete memoryless sources, $\left(U^{[1]}, U^{[2]}\right)$, to a central receiver through a multiple access channel (MAC). One of the simplest models is the binary-input Gaussian MAC. The 2-user binary-input Gaussian MAC has been extensively studied in the literature and is defined by

$$Y = \alpha^{[1]} X^{[1]} + \alpha^{[2]} X^{[2]} + N. \tag{2.9}$$

The system model is shown in Fig. 15. The channel inputs are binary i.e., $X^{[1]}, X^{[2]} \in \{\pm 1\}$ and the variation in channel gains $\alpha^{[1]}, \alpha^{[2]} \in [0, \infty)$ can be explained either by fading or by different power constraints for the two users. The noise N is a zero-mean Gaussian random variable, with fixed variance of 1. The capacity region is defined as the set of all achievable rate tuples $(R^{[1]}, R^{[2]})$, given by the equations

$$\begin{aligned} R^{[1]} &\leq I\left(X^{[1]}; Y | X^{[2]}\right) \\ R^{[2]} &\leq I\left(X^{[2]}; Y | X^{[1]}\right) \\ R^{[1]} + R^{[2]} &\leq I\left(X^{[1]}, X^{[2]}; Y\right). \end{aligned} \tag{2.10}$$

In this work, we fix the rate pair for transmission and view the capacity region

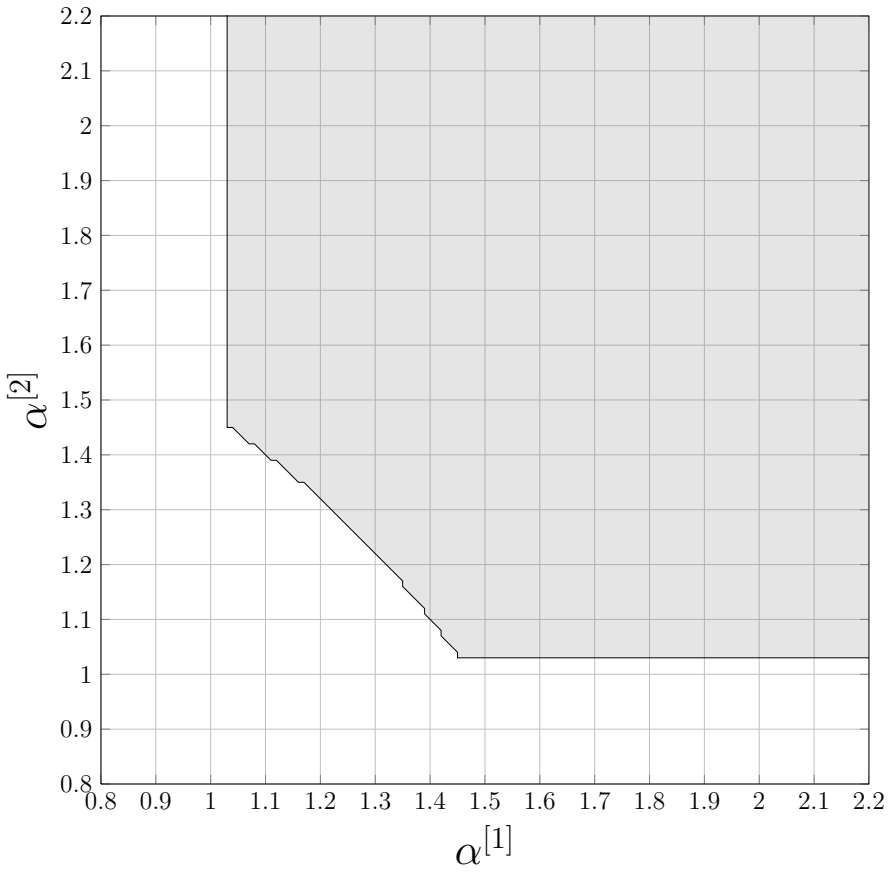

Fig. 16. The MAC-ACPR for the rate pair $(0.5, 0.5)$ is shown above.

in terms of the achievable channel gains for that rate pair.[2] In other words, the capacity region is the set of all channel gains $(\alpha^{[1]}, \alpha^{[2]})$ that are achievable, i.e., satisfy (2.10). We call this region as the MAC achievable channel-parameter region (MAC-ACPR), to illustrate that the capacity region is defined in terms of achievable channel parameters. The MAC-ACPR for the rate pair $(0.5, 0.5)$ is shown in Fig. 16.

[2]To simplify notation, we assume that both users employ codes with design rate R, chosen independently from the same ensemble.

1. Density Evolution

To simplify notation, we assume that the transmission is bit-aligned. The factor graph of the joint decoder (see Fig. 17) consists of two single user Tanner graphs, whose variable nodes are connected through a function node [2, p. 308]. Codes 1 and 2 correspond to the bottom and top half of the Tanner graph. The variable nodes that are connected via the function node are chosen at random.[3] Let $X_i = \left(X_i^{[1]}, X_i^{[2]}\right)$ and $\boldsymbol{X} = \left(\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]}\right)$. Without loss of generality, we can label the elements of $\{\pm 1\}^2$ by integers $\mathcal{X} \triangleq \{0, 1, 2, 3\}$ using the map $\pi : \mathcal{X} \to \{\pm 1\}^2$, defined by

$$0 \mapsto (+1, +1), 1 \mapsto (+1, -1), 2 \mapsto (-1, +1) \text{ and } 3 \mapsto (-1, -1).$$

Let $\pi_1, \pi_2 : \mathcal{X} \to \{\pm 1\}$ be the projections onto the first and second coordinate respectively. Then, the canonical representation of the channel output is given by

$$\begin{aligned} \nu_{x_i}(y_i) &= p_{Y|X^{[1]},X^{[2]}}(y_i|\pi_1(x_i), \pi_2(x_i)) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \alpha^{[1]}\pi_1(x_i) - \alpha^{[2]}\pi_2(x_i))^2}{2\sigma^2}\right]. \end{aligned}$$

Let $\mu_{i,v\to f}^{[j]}$ and $\mu_{i,f\to v}^{[j]}$ denote the "variable node to function node" and "function node to variable node" messages[4], respectively, for variable node i of the jth user. Here $j \in \{1, 2\}$ and $i \in \{1, 2, \ldots, n\}$. The message passing rules at the function node are given by

$$\mu_{i,f\to v}^{[1]} = \log \frac{\nu_0(y_i)\mathrm{e}^{\mu_{i,v\to f}^{[2]}} + \nu_1(y_i)}{\nu_2(y_i)\mathrm{e}^{\mu_{i,v\to f}^{[2]}} + \nu_3(y_i)}, \tag{2.11}$$

$$\mu_{i,f\to v}^{[2]} = \log \frac{\nu_0(y_i)\mathrm{e}^{\mu_{i,v\to f}^{[1]}} + \nu_2(y_i)}{\nu_1(y_i)\mathrm{e}^{\mu_{i,v\to f}^{[1]}} + \nu_3(y_i)}. \tag{2.12}$$

[3] Other matching rules result in a different performance in general.

[4] Here, the messages are in the log-likelihood domain.

In general, this function node operation is not symmetric with respect to the users. The operation is symmetric only for the case of the same fading coefficients i.e., when $\alpha^{[1]} = \alpha^{[2]}$.

One cannot use the all-zero codeword assumption for this problem. Instead, one may assume that both users transmit codewords of type one-half, which occurs with high probability (a more thorough discussion can be found in [2, p. 296]). We use the notation $\mathsf{a}_{\mathrm{BAWGNMA}} \triangleq \mathsf{a}_{\mathrm{BAWGNMA}(\alpha^{[1]},\alpha^{[2]})}$ to denote the density of the received random variable Y. Let $\zeta_{1\to 2}(\cdot, \mathsf{a}_{\mathrm{BAWGNMA}})$ (resp. $\zeta_{2\to 1}(\cdot, \mathsf{a}_{\mathrm{BAWGNMA}})$) be the density transformation operator corresponding to a message from user 1 to user 2 (resp. user 2 to user 1) via the function node. More precisely,

$$\zeta_{1\to 2}(\mathsf{a}, \mathsf{a}_{\mathrm{BAWGNMA}}) \triangleq \sum_{x\in\mathcal{X}} p_X(x)\zeta_{12}(\mathsf{a}(\pi_1(x)u), \nu_x(u))$$
$$\zeta_{2\to 1}(\mathsf{b}, \mathsf{a}_{\mathrm{BAWGNMA}}) \triangleq \sum_{x\in\mathcal{X}} p_X(x)\zeta_{21}(\mathsf{b}(\pi_2(x)u), \nu_x(u)),$$

where $\zeta_{12}(\cdot,\cdot)$ and $\zeta_{21}(\cdot,\cdot)$ are density transformation operators corresponding to (2.11) and (2.12). In this case, $p_X(x) = 1/4, \forall x \in \mathcal{X}$. Here, $\mathsf{a}(u)$ (respectively $\mathsf{b}(u)$) is the density of the messages $m^{[1]}_{i,v\to f}$ ($m^{[2]}_{i,v\to f}$). These operators can be computed numerically for discretized densities following the procedure outlined in [41]. Using the notation described in Section A, the DE equations for the joint decoder are given by

$$\mathsf{a}_{\ell+1} = \zeta_{2\to 1}\Big(L\left(\rho(\mathsf{b}_\ell)\right), \mathsf{a}_{\mathrm{BAWGNMA}}\Big) \circledast \lambda(\rho(\mathsf{a}_\ell))$$
$$\mathsf{b}_{\ell+1} = \zeta_{1\to 2}\Big(L\left(\rho(\mathsf{a}_\ell)\right), \mathsf{a}_{\mathrm{BAWGNMA}}\Big) \circledast \lambda(\rho(\mathsf{b}_\ell)).$$

These equations accurately represent the evolution of densities at the decoder due to the symmetry of the variable and check node operations. The fixed points of density

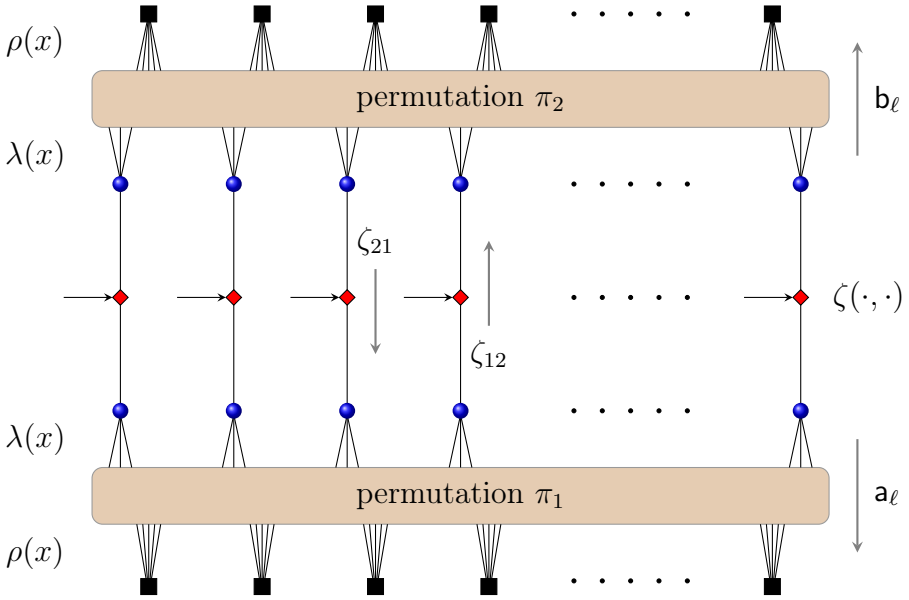

Fig. 17. Tanner graph of the joint decoder. The variable nodes of each code are connected through function nodes, which receives the channel outputs. The joint decoder iterates by passing messages between the component decoders.

evolution are the triples $(\mathsf{a}_{\mathrm{BAWGNMA}}, \mathsf{a}, \mathsf{b})$ that satisfy

$$\begin{aligned} \mathsf{a} &= \zeta_{2\to 1}\Big(L\left(\rho(\mathsf{b})\right), \mathsf{a}_{\mathrm{BAWGNMA}}\Big) \circledast \lambda(\rho(\mathsf{a})) \\ \mathsf{b} &= \zeta_{1\to 2}\Big(L\left(\rho(\mathsf{a})\right), \mathsf{a}_{\mathrm{BAWGNMA}}\Big) \circledast \lambda(\rho(\mathsf{b})). \end{aligned} \tag{2.13}$$

Consider the line $\alpha^{[2]} = \theta\alpha^{[1]}$, for some $\theta \in [0, +\infty]$. The BP threshold along the line is defined by

$$\alpha^{\mathrm{BP}}(\lambda, \rho, \theta) = \sup\big\{\alpha : \text{ The fixed point equation (2.13) has a solution } (\mathsf{a}, \mathsf{b}) \neq (\Delta_{+\infty}, \Delta_{+\infty})\big\}.$$

The set of all points $(\alpha, \theta\alpha)$ such that $\alpha \geq \alpha^{\mathrm{BP}}(\lambda, \rho, \theta)$ is called the BP-ACPR and its boundary is called the DE boundary.

CHAPTER III

THE MAP DECODING THRESHOLD

Analogous to point-to-point communication, we can define a GEXIT function for multi-terminal problems by taking the gradient of the *residual entropy* with respect to the channel parameters. There is an area theorem associated with the line integral of the GEXIT function. In this chapter we introduce GEXIT functions for the noisy SW problem and the Gaussian MAC and discuss the associated area theorem in Sections 1 and 1. Using this area theorem, we can construct an upper bound on the MAP performance of LDPC codes for these channels. This upper bound is discussed in Sections 2 and 2. In Section 3, we show that this upper bound is tight under some conditions, for the SWE problem, by considering an analytic extension of the BP-GEXIT curve.

A. The Noisy Slepian-Wolf Problem

Consider the noisy SW problem with channel parameter $(\alpha^{[1]}, \alpha^{[2]})$. Suppose we use different codes from the ensemble LDPC(n, λ, ρ) for each user, using a punctured systematic encoder. The following discussion applies to the two users using different LDPC ensembles. Let $X_i^{[1]}$ and $X_i^{[2]}$ denote the ith bit of user 1 and 2 respectively. Let $X_i = \left(X_i^{[1]}, X_i^{[2]}\right)$, $\boldsymbol{X} = \left(\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]}\right)$, $Y_i = \left(Y_i^{[1]}, Y_i^{[2]}\right)$ and $\boldsymbol{Y} = \left(\boldsymbol{Y}^{[1]}, \boldsymbol{Y}^{[2]}\right)$. Also, we denote $\boldsymbol{Y}(\alpha^{[1]}, \alpha^{[2]}) = \left(\boldsymbol{Y}^{[1]}(\alpha^{[1]}), \boldsymbol{Y}^{[2]}(\alpha^{[2]})\right)$ to emphasize the dependence on the channel parameter. We use the notation $[X]_{k_1}^{k_2}$ to denote the sub-vector $(X_{k_1}, X_{k_1+1}, \cdots, X_{k_2})$. The bits $[X]_{n-k+1}^{n}$ are systematic bits and are not transmitted, and hence $[Y]_{n-k+1}^{n}$ are considered to be erasures and do not depend on the channel.

We define the GEXIT function

$$\mathsf{g}(\alpha^{[1]}, \alpha^{[2]}) \triangleq \frac{1}{2(n-k)} \nabla \mathbb{E}\left[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha^{[1]}, \alpha^{[2]}))\right].$$

By definition, the line integral of the GEXIT function is path independent. As seen in the next section, it is instructive to consider line integrals along *monotonic* curves. The projection of the GEXIT function along such curves satisfies a natural area theorem, enabling us to obtain a bound on the MAP performance, along the lines of Section 3.

1. GEXIT Curves

In this section we consider projections of the GEXIT function along monotonic curves in parameter space, to find a simple characterization of the GEXIT function. Here monotonicity is defined with respect to the partial order implied by channel degradation. Suppose that X_i is transmitted via a channel with parameter $(\alpha_i^{[1]}, \alpha_i^{[2]})$. Consider a curve $\mathcal{C}$ in $[0,1]^{2(n-k)}$, parametrized by α i.e., $\alpha \mapsto \mathsf{c}(\alpha) \triangleq \left[(\alpha_i^{[1]}, \alpha_i^{[2]})\right]_1^{n-k}$. We assume that $\mathcal{C}$ is smooth and that the channel is degraded with respect to α. Further, $\alpha_i^{[j]}(0) = 0$ and $\alpha_i^{[j]}(1) = 1$ for $j = 1, 2$ and $i = 1, \cdots, n-k$. The projection of the GEXIT function along the curve $\mathcal{C}$ is defined by

$$\begin{aligned} \mathsf{g}_{\mathcal{C}}(\alpha) &\triangleq \frac{1}{2(n-k)} \nabla H(\boldsymbol{X}|\boldsymbol{Y}) \cdot \nabla \mathsf{c}(\alpha) \qquad (3.1) \\ &= \frac{1}{2(n-k)} \sum_{i=1}^{n-k} \underbrace{\frac{\partial}{\partial \alpha_i^{[1]}} H(\boldsymbol{X}|\boldsymbol{Y}) \frac{\partial \alpha_i^{[1]}}{\partial \alpha} + \frac{\partial}{\partial \alpha_i^{[2]}} H(\boldsymbol{X}|\boldsymbol{Y}) \frac{\partial \alpha_i^{[2]}}{\partial \alpha}}_{\triangleq \mathsf{g}_{\mathcal{C},i}(\alpha_i^{[1]}, \alpha_i^{[2]})}. \end{aligned}$$

Let $\mathbf{y}_{\sim i}^{[j]} = \mathbf{y} \backslash y_i^{[j]}$,

$$\phi_i^{[j]}(\mathbf{y}_{\sim i}^{[j]}) = \log p_{X_i^{[j]}|\boldsymbol{Y}_{\sim i}^{[j]}}(+1|\mathbf{y}_{\sim i}^{[j]}) - \log p_{X_i^{[j]}|\boldsymbol{Y}_{\sim i}^{[j]}}(-1|\mathbf{y}_{\sim i}^{[j]})$$

and $\Phi_i^{[j]} \triangleq \phi_i^{[j]}(\boldsymbol{Y}_{\sim i}^{[j]})$ be the corresponding random variable, for $j = 1, 2$ and $i = 1, \cdots, n$. Note that $\phi_i^{[j]}(\cdot)$ is the extrinsic MAP estimator of $X_i^{[j]}$, for $j = 1, 2$, and $i = 1, \cdots, n$.

Lemma III.1. *Suppose that all bits are transmitted through the same channel i.e., $\alpha_i^{[j]}(\alpha) = \alpha^{[j]}(\alpha)$, for $j = 1, 2$ and $i = 1, \cdots, n-k$. Then, the ith GEXIT function is given by*

$$\begin{aligned}\mathsf{g}_{\mathcal{C},i}(\alpha^{[1]}, \alpha^{[2]}) = \frac{\partial \alpha^{[1]}}{\partial \alpha} \int_u \mathsf{a}_i(u) \kappa(\mathsf{a}_{BMSC(\alpha^{[1]})}, u)\, du \\ + \frac{\partial \alpha^{[2]}}{\partial \alpha} \int_u \mathsf{b}_i(u) \kappa(\mathsf{a}_{BMSC(\alpha^{[2]})}, u) du,\end{aligned}$$

where $\mathsf{a}_i(u)$ (resp. $\mathsf{b}_i(u)$) is the distribution of $\Phi_i^{[1]}$ (resp. $\Phi_i^{[2]}$) given $X_i^{[1]} = +1$ (resp. $X_i^{[2]} = +1$) and the GEXIT kernel is given by

$$\kappa(\mathsf{a}_{BMSC(\alpha)}, u) = \int_v \frac{\partial}{\partial \alpha} \mathsf{a}_{BMSC(\alpha)}(v) \log_2(1 + e^{-u-v}) dv. \tag{3.2}$$

Proof. The proof is given in Appendix A. ■

The GEXIT function is hard to compute and hence we use the BP-GEXIT function instead. The BP-GEXIT function is obtained by replacing the MAP extrinsic estimator with the corresponding BP estimator. Let $\Phi_i^{\mathrm{BP},\ell,n}$ denote the BP extrinsic estimate of X_i after ℓ iterations of the joint decoder. The BP extrinsic estimate is computed using the computation graph of depth ℓ for function node i. Define the BP-GEXIT function at the ℓth iteration $\mathsf{g}_{\mathcal{C}}^{\mathrm{BP},\ell,n}(\alpha)$ in a similar manner to [39] (taking an expectation over all possible computation graphs) and the asymptotic BP-GEXIT function is defined as $\mathsf{g}_{\mathcal{C}}^{\mathrm{BP}}(\alpha) = \lim_{\ell\to\infty} \lim_{n\to\infty} \mathsf{g}_{\mathcal{C}}^{\mathrm{BP},\ell,n}(\alpha)$. For fixed ℓ, in the limit of $n \to \infty$, the computation graph of each function node becomes tree-like with high probability. This implies that the computation graphs of the two variable nodes

(which themselves become tree-like) connected to the function node do not overlap with high probability. The extrinsic estimate of X_i can then be computed via the extrinsic estimates of $X_i^{[1]}$ and $X_i^{[2]}$. The asymptotic BP-GEXIT function can be computed through the fixed points of density evolution $(\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})}, \mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})}, \mathsf{a}, \mathsf{b})$ which satisfy (2.8) and is discussed in the following lemma.

Lemma III.2. *Consider a monotonic curve $\mathcal{C}$ and transmission over the channel pair $(\alpha^{[1]}(\alpha), \alpha^{[2]}(\alpha))$ and let $(\mathsf{a}_{BMSC(\alpha^{[1]})}, \mathsf{a}_{BMSC(\alpha^{[2]})}, \mathsf{a}, \mathsf{b})$ be a fixed point of DE. Define the BP-GEXIT value of the fixed point by*

$$\mathsf{G}_{\mathcal{C}}^{BP}(\mathsf{a}_{BMSC(\alpha^{[1]})}, \mathsf{a}_{BMSC(\alpha^{[2]})}, \mathsf{a}, \mathsf{b}) \triangleq \frac{\partial \alpha^{[1]}}{\partial \alpha} \int_u \mathsf{a}(u)\kappa(\mathsf{a}_{BMSC(\alpha^{[1]})}, u)du + \frac{\partial \alpha^{[2]}}{\partial \alpha} \int_u \mathsf{b}(u)\kappa(\mathsf{a}_{BMSC(\alpha^{[2]})}, u)du.$$

The GEXIT kernel $\kappa(\cdot,\cdot)$ is defined as in (3.2). The BP-GEXIT curve $\mathsf{g}_{\mathcal{C}}^{BP}(\alpha)$ is given by $(\alpha, \mathsf{G}_{\mathcal{C}}^{BP}(\mathsf{a}_{BMSC(\alpha^{[1]})}, \mathsf{a}_{BMSC(\alpha^{[2]})}, \mathsf{a}, \mathsf{b})), \alpha \in [0,1]$.

Proof. The proof follows immediately from the definition of the BP-GEXIT curve. ■

2. MAP Upper Bound

The GEXIT kernel preserves degradation (see [2, Chapter 4]) and hence the BP-GEXIT curve always lies above the GEXIT curve, allowing one to bound the MAP threshold. Consider transmission using codes from the ensemble LDPC(n, λ, ρ). For a fixed curve $\mathcal{C}$, we define the MAP threshold as

$$\alpha_{\mathcal{C}}^{\mathrm{MAP}} = \inf\left\{\alpha : \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha, \mathcal{C}))] > 0\right\},$$

where the expectation is taken over all codes in the ensemble. The set of parameters $\bigcup_{\mathcal{C}}(\alpha^{[1]}(\alpha_{\mathcal{C}}^{\mathrm{MAP}}), \alpha^{[2]}(\alpha_{\mathcal{C}}^{\mathrm{MAP}}))$ is called the MAP boundary and the set of all channel

parameters which are degraded with respect to the boundary is called the MAP-ACPR. By definition of the GEXIT function, this gives

$$\begin{aligned}\int_{\alpha_{\mathcal{C}}^{\mathrm{MAP}}}^{1} \mathsf{g}_{\mathcal{C}}(\alpha)\mathrm{d}\alpha &= \frac{1}{n}\int_{\alpha_{\mathcal{C}}^{\mathrm{MAP}}}^{1} \frac{\mathrm{d}H(\boldsymbol{X}|\boldsymbol{Y}(\alpha,\mathcal{C}))}{\mathrm{d}\alpha}\mathrm{d}\alpha \\ &= \frac{1}{n}H(\boldsymbol{X}|\boldsymbol{Y}(1)) \\ &= \frac{\gamma H(U_1,U_2)}{2(1-\gamma)}.\end{aligned}$$

The above equation gives us a procedure to compute an upper bound on the MAP threshold, using GEXIT curves. For a fixed curve $\mathcal{C}$, let $\bar{\alpha}_{\mathcal{C}}$ denote the largest positive number such that

$$\int_{\bar{\alpha}_{\mathcal{C}}}^{1} \mathsf{g}_{\mathcal{C}}^{\mathrm{BP}}(\alpha)\mathrm{d}\alpha = \frac{\gamma H(U_1,U_2)}{2(1-\gamma)}.$$

Then the MAP threshold $\alpha_{\mathcal{C}}^{\mathrm{MAP}} \leq \bar{\alpha}_{\mathcal{C}}$ and the MAP boundary is degraded with respect to the set $\bigcup_{\mathcal{C}}(\alpha^{[1]}(\bar{\alpha}_{\mathcal{C}}), \alpha^{[2]}(\bar{\alpha}_{\mathcal{C}}))$. This set is indeed equal to the MAP boundary for some cases (as shown in the next section) and we conjecture that this is true in general for the noisy SW problem. Henceforth, we shall use the term MAP boundary loosely to denote this outer bound.[1] The BP-GEXIT curve and the MAP threshold for the punctured LDPC$(4,6)$ ensemble are shown in Fig. 18 for the erasure case and the MAP threshold for symmetric channel conditions is $\alpha^{\mathrm{MAP}} \approx 0.6245$. The BP-GEXIT curve and the MAP threshold for the punctured LDPC$(4,6)$ ensemble are shown in Fig. 19 for the BSC case and the MAP threshold for symmetric channel conditions is $\alpha^{\mathrm{MAP}} \approx 0.6324$.

[1]For computation of the upper bound, it is easiest to consider straight lines passing through $(1,1)$ with different slopes.

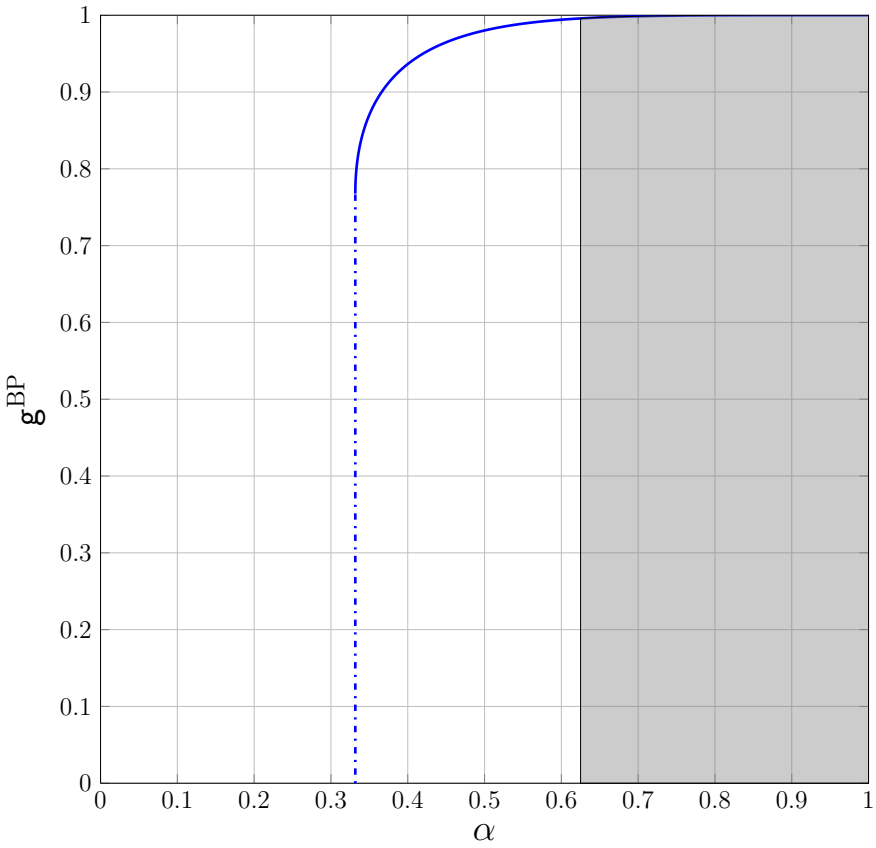

Fig. 18. The BP-GEXIT function for the SW problem with erasures (SWE) along the curve $\alpha^{[1]} = \alpha^{[2]}$. The upper bound on the MAP threshold is given by $\bar{\alpha} \approx 0.6425$. The correlation parameter is $p = 0.5$.

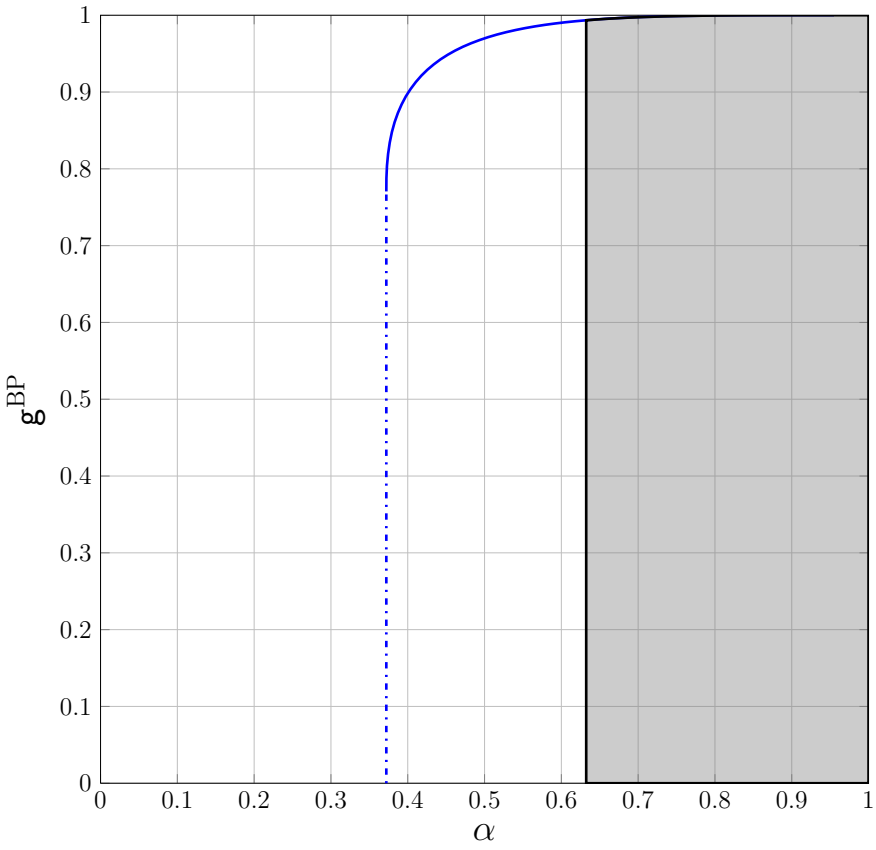

Fig. 19. The BP-GEXIT function for the SW problem with BSC correlation along the curve $\alpha^{[1]} = \alpha^{[2]}$. The upper bound on the MAP threshold is given by $\bar{\alpha} \approx 0.6324$. The correlation parameter is $p = 0.9$.

3. Tightness of the Upper Bound

In this section, we rigorously analyze the upper bound on the MAP threshold for the erasure SW case and show that the bound is tight. The discussion closely follows the proof in [2]. Throughout this section, fixed points refer to both stable and unstable fixed points of DE. First consider a monotonic curve through parameter space $\mathcal{C}$ as defined above, which is differentiable and degraded with respect to α. For simplicity of notation, we omit the dependence of the various quantities on $\mathcal{C}$ throughout this section. Let $(\alpha, a, b) \triangleq (\alpha^{[1]}(\alpha), \alpha^{[2]}(\alpha), a, b)$ be a fixed point of density evolution and define $g(x) = 1 - \rho(1-x)$. For the case when there is a unique α such that (α, a, b) is a fixed point of DE, we write $\alpha(a, b)$ to denote the channel parameter associated with the fixed point (α, a, b).

The extended belief propagation (EBP) GEXIT function can be computed by also considering the unstable fixed points of density evolution. These unstable fixed points can be computed numerically by running density evolution at fixed entropy as discussed in [39, Section VIII]. The the EBP-GEXIT curve, which is an analytical extension of the BP-GEXIT curve, $\mathsf{g}^{\mathrm{EBP}}(a, b)$, is given in parametric form by $(\alpha(a, b), \mathsf{G}^{\mathrm{EBP}}(a, b))$, where $\mathsf{G}^{\mathrm{EBP}}$ is an analytic extension of G^{BP}.

Lemma III.3. *The trial entropy for the SWE problem, which is obtained by integrating the EBP-GEXIT function, is given by*

$$\begin{aligned} P(a,b) = &-\frac{p\gamma}{1-\gamma} L(g(a))L(g(b)) \\ &+ \frac{1}{1-\gamma}\left(L(g(a))\frac{a}{\lambda(g(a))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-a) - aR'(1-a)\right]\right) \\ &+ \frac{1}{1-\gamma}\left(L(g(b))\frac{b}{\lambda(g(b))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-b) - bR'(1-b)\right]\right). \end{aligned}$$

Proof. The proof is given in Appendix A. ∎

The following Lemma enables us to compute the expected residual degree distribution for the erasure SW problem. This can then be used to compute the MAP threshold along any curve $\mathcal{C}$ as described above.

Lemma III.4. *Assume that we run the iterative decoder until it reaches a fixed point. At the fixed point $(\alpha^{[1]}, \alpha^{[2]}, a, b)$, the expected degree distribution of the residual graph has the form*

$$\begin{aligned}\tilde{R}(z) &\propto R(1-a+za) - R(1-a) - zaR'(1-a) \\ &\quad + R(1-b+zb) - R(1-b) - zbR'(1-b) \\ \tilde{L}(z) &\propto (1-\gamma)\left[\alpha^{[1]} L(g(a)z) + \alpha^{[2]} L(g(b)z)\right] \\ &\quad + \gamma(1-p)\left[L(g(a)z) + L(g(b)z)\right] + \frac{\gamma p}{2} L(g(a)z) L(g(b)z).\end{aligned}$$

Proof. The proof is given in Appendix A. ■

Theorem III.5. *Consider the parametrization of the fixed point (a, b) by $a(x)$ and $b(x)$ along the curve $\mathcal{C}$ as defined in Appendix A. Let x^{MAP} be the unique non-zero solution corresponding to $P(x) = P(a(x), b(x)) = 0$. Then, $\alpha^{\mathrm{MAP}} = \alpha(x^{\mathrm{MAP}})$ is the MAP threshold along the curve $\mathcal{C}$.*

Proof. We first show that, at $\alpha = \alpha^{\mathrm{MAP}}$, the design rate of the residual ensemble is zero. The design rate of residual graph

$$\begin{aligned}&\mathtt{R}(\tilde{L}, \tilde{R}) \propto \tilde{L}(1) - \frac{L'(1)}{R'(1)} \cdot \tilde{R}(1) \\ &= \frac{1}{1-\gamma}\left(\left[(1-\gamma)\alpha^{[1]}(x^{\mathrm{MAP}}) + \gamma(1-p)\right] L(g(a(x^{\mathrm{MAP}}))) + \gamma p \frac{L^2(g(a(x^{\mathrm{MAP}})))}{2}\right. \\ &\quad \left. - \frac{L'(1)}{R'(1)}\left[1 - R(1-a(x^{\mathrm{MAP}})) - a(x^{\mathrm{MAP}}) R'(1-a(x^{\mathrm{MAP}}))\right]\right) \\ &\quad + \frac{1}{1-\gamma}\left(\left[(1-\gamma)\alpha^{[2]}(x^{\mathrm{MAP}}) + \gamma(1-p)\right] L(g(b(x^{\mathrm{MAP}}))) + \gamma p \frac{L^2(g(b(x^{\mathrm{MAP}})))}{2}\right. \\ &\quad \left. - \frac{L'(1)}{R'(1)}\left[1 - R(1-b(x^{\mathrm{MAP}})) - b(x^{\mathrm{MAP}}) R'(1-b(x^{\mathrm{MAP}}))\right]\right)\end{aligned}$$

$$
\begin{aligned}
&- \frac{p\gamma}{1-\gamma} \frac{\left(L(g(a(x^{\mathrm{MAP}}))) - L(g(b(x^{\mathrm{MAP}})))\right)^2}{2} \\
&= P(x^{\mathrm{MAP}}) \\
&= 0,
\end{aligned}
$$

by assumption. Thus, the design rate of the residual graph at α^{MAP} is zero. It remains to show that the actual rate is zero. To see this, one observes that the residual graph is a two-edge LDPC ensemble and the numerical technique in [42, Theorem IV.9] (which is a generalization of [2, Lemma 3.22]) can be applied. From this, it follows that α^{MAP} is indeed the MAP threshold. ■

Theorem III.6. *Consider transmission using regular* $(\mathtt{l}, \mathtt{r})$ *LDPC codes of rate* $\mathtt{r}$ *for the erasure SW problem. Then, for any curve* $\mathcal{C}$ *through the parameter space as described above*

$$\lim_{\substack{\mathtt{l},\mathtt{r}\to\infty \\ 1-\frac{\mathtt{l}}{\mathtt{r}}=\mathtt{R}}} \alpha_{\mathcal{C}}^{\mathrm{MAP}} = \alpha_{\mathcal{C}}^{SW}$$

where $\alpha_{\mathcal{C}}^{SW}$ *corresponds to the SW conditions* (2.5) *along the curve* $\mathcal{C}$.

Proof. First, we observe that $x = 0$ implies $g(x) = 0, L(g(x)) = 0, R(1-x) = 1$, and $R'(1-x) = 1$. Also, in the limit $\mathtt{l}, \mathtt{r} \to \infty$ with $1 - \frac{\mathtt{l}}{\mathtt{r}} = \mathtt{R}$ constant, $g(x) = 1, L(g(x)) = 1, R(1-x) = 0, R'(1-x) = 0$ if $x \neq 0$. Based on this observation and the fact that $P(x) = 0$ at the MAP threshold, the result follows by considering three different cases namely $a = 0, b \neq 0$, $a \neq 0, b = 0$ and $a \neq 0, b \neq 0$ for fixed points (a, b) of density evolution. The exact case to be considered depends on the curve $\mathcal{C}$ and the three cases give the three boundaries of the SW region. ■

Remark III.1. From this, we see that the MAP boundary of regular LDPC codes with large degrees approaches the SW boundary. Hence regular LDPC codes with large

degrees are universal under MAP decoding for the SWE problem. We conjecture that this observation holds for the SW problem with BSC correlation.

B. The Gaussian Multiple-Access Channel

Consider the Gaussian multiple-access channel with channel parameters $(\alpha^{[1]}, \alpha^{[2]})$. Suppose each user uses a code chosen independently from the ensemble $\mathrm{LDPC}(n, \lambda, \rho)$. The following discussion is easily extended to the case where the users use different LDPC ensembles. Let $X_i^{[1]}$ and $X_i^{[2]}$ denote the ith bit of user 1 and 2 respectively. Let $X_i = \left(X_i^{[1]}, X_i^{[2]}\right)$, $\boldsymbol{X} = \left(\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]}\right)$, $Y_i = \left(Y_i^{[1]}, Y_i^{[2]}\right)$ and $\boldsymbol{Y} = \left(\boldsymbol{Y}^{[1]}, \boldsymbol{Y}^{[2]}\right)$. Also, we denote $\boldsymbol{Y}(\alpha^{[1]}, \alpha^{[2]}) = \left(\boldsymbol{Y}^{[1]}(\alpha^{[1]}), \boldsymbol{Y}^{[2]}(\alpha^{[2]})\right)$ to emphasize the dependence on the channel parameter. We define the GEXIT function

$$\mathsf{g}(\alpha^{[1]}, \alpha^{[2]}) \triangleq \frac{1}{2n} \nabla \mathbb{E}\left[H(\boldsymbol{X}|\boldsymbol{Y}(\alpha^{[1]}, \alpha^{[2]}))\right].$$

By definition, the line integral of the GEXIT function is path independent. As seen in the next section, it is instructive to consider line integrals along monotonic curves. The projection of the GEXIT function along such curves satisfies a natural area theorem, enabling us to obtain a bound on the MAP performance, along the lines of Section 3.

1. GEXIT Curves

We consider projections of the GEXIT function along monotonic curves, where monotonicity is defined with respect to the partial order implied by channel degradation, in parameter space. For the binary-input Gaussian MAC defined by (2.9), rays through the origin characterized by a parameter $\alpha \in [0, \infty)$, with $\alpha^{[1]} = \alpha$ and $\alpha^{[2]} = \theta\alpha$ are monotonic curves for some fixed $\theta \in [0, \infty)$. The following approach can be ap-

plied to any binary-input MAC characterized by a single parameter, whose density is differentiable and degraded with respect to that parameter.

Now, suppose that the ith bit is transmitted through a channel with parameter α_i and that each α_i is a differentiable function of α. The GEXIT curve is defined by (3.1). A more convenient expression for the GEXIT curve can be derived following the procedure given in [43] for non-binary codes. Let $\mathbf{y}_{\sim i} = \mathbf{y} \backslash y_i$, $\phi_i(\mathbf{y}_{\sim i}) = \{p_{X_i|\boldsymbol{Y}_{\sim i}}(x|\mathbf{y}_{\sim i}), x \in \mathcal{X}\}$ and $\Phi_i \triangleq \phi_i(\boldsymbol{Y}_{\sim i})$ be the corresponding random variable. Note that $\phi_i(\cdot)$ is the extrinsic MAP estimator of X_i.[2]

Lemma III.7. *Suppose that all bits are transmitted through channel with parameter α. Then, the ith GEXIT function is given by*

$$\mathsf{g}_i(\alpha) = \sum_{x \in \mathcal{X}} p(x) \int_{\mathbf{u}} \mathsf{a}_{x,i}(\mathbf{u}) \kappa_x(\mathbf{u}) d\mathbf{u},$$

where $\mathsf{a}_{x,i}(\mathbf{u})$ is the distribution of Φ_i given $X_i = x$ and the GEXIT kernel is given by

$$\kappa_x(\mathbf{u}) = \int \frac{\partial}{\partial \alpha} p(y|x) \log_2 \frac{\sum_{x'} u[x'] p(y|x')}{u[x] p(y|x)} dy, \tag{3.3}$$

where $u[j]$ denotes the jth component of $\mathbf{u}$.

Proof. This proof is given in Appendix A. ■

As discussed in Section 1, the asymptotic BP-GEXIT function can be computed through the fixed points of density evolution $(\mathsf{a}_{\mathrm{BAWGNMA}(\alpha)}, \mathsf{a}, \mathsf{b})$ that satisfy (2.13) and is discussed in the following Lemma.

---

[2]To see this, write

$$p_{\boldsymbol{Y}_{\sim i}|X_i}(\mathbf{y}_{\sim i}|x_i) = \frac{p_{X_i|\boldsymbol{Y}_{\sim i}}(x_i|\mathbf{y}_{\sim i})}{p_{X_i}(x_i)} p_{\boldsymbol{Y}_{\sim i}}(\mathbf{y}_{\sim i}) = \frac{\phi_i \cdot e[x_i]}{p_{X_i}(x_i)} p_{\boldsymbol{Y}_{\sim i}}(\mathbf{y}_{\sim i}),$$

where $e[x_i]$ is the standard basis vector with a 1 in the x_i-th coordinate and use the result in [2, p. 29] regarding sufficient statistics.

Lemma III.8. *Consider transmission over the multiple-access channel* $\mathsf{a}_{BAWGNMA}(\alpha)$ *and let* $(\mathsf{a}_{BAWGNMA(\alpha)}, \mathsf{a}, \mathsf{b})$ *be a fixed point of DE. Define the BP-GEXIT value of the fixed point by*

$$\mathsf{G}^{BP}(\mathsf{a}_{BAWGNMA}, \mathsf{a}, \mathsf{b}) \triangleq \sum_{x \in \mathcal{X}} p(x) \int \mathsf{F}_x[\mathsf{a}, \mathsf{b}](u, v) \kappa_x(u, v) du dv. \tag{3.4}$$

The GEXIT kernel $\kappa_x(\cdot, \cdot)$ *is defined as in* (3.3) *and the operator* $\mathsf{F}_x[\cdot, \cdot]$ *(defined in* (A.3)*) computes the density of the extrinsic BP estimate* Φ^{BP} *given* $X = x$*. The BP-GEXIT curve* $\mathsf{g}^{BP}(\alpha)$ *is given by* $(\alpha, \mathsf{G}(\mathsf{a}_{BAWGNMA(\alpha)}, \mathsf{a}, \mathsf{b})), \alpha \in [0, \infty)$*.*

Proof. The proof is given in Appendix A. ■

2. MAP Upper Bound

It can be shown that the BP-GEXIT function is a lower bound on the GEXIT function (see the discussion in [2, p. 206]). Consider transmission using codes from the ensemble LDPC(n, λ, ρ). For a fixed θ, we define the MAP threshold as

$$\alpha^{\mathrm{MAP}}(\theta) = \sup \left\{ \alpha : \liminf_{n \to \infty} \frac{1}{n} \mathbb{E}[H(\boldsymbol{X} | \boldsymbol{Y}(\alpha, \theta))] > 0 \right\},$$

where the expectation is taken over all codes in the ensemble. By definition of the GEXIT function, this gives

$$\begin{aligned} \int_{\alpha^{\mathrm{MAP}}(\theta)}^{0} \mathsf{g}(\alpha) \mathrm{d}\alpha &= \frac{1}{n} \int_{\alpha^{\mathrm{MAP}}(\theta)}^{0} \frac{\mathrm{d}H(\boldsymbol{X} | \boldsymbol{Y}(\alpha))}{\mathrm{d}\alpha} \mathrm{d}\alpha \\ &= \frac{1}{n} H(\boldsymbol{X} | \boldsymbol{Y}(0)) \\ &= 2\mathtt{R}(\lambda, \rho). \end{aligned}$$

The above equation gives us a procedure to compute the MAP threshold, using the GEXIT curve. Let $\bar{\alpha}$ denote the smallest positive number such that

$$\int_{\bar{\alpha}}^{0} \mathsf{g}^{\mathrm{BP}}(\alpha)\mathrm{d}\alpha = 2\mathtt{R}(\lambda,\rho),$$

where $\mathtt{R}(\lambda,\rho)$ is the design rate of the ensemble $\mathrm{LDPC}(\lambda,\rho)$. Then the MAP threshold $\alpha^{\mathrm{MAP}} \leq \bar{\alpha}$. The set of all points $(\alpha', \theta\alpha')$ such that $\alpha' \geq \alpha^{\mathrm{MAP}}(\theta)$ form the MAP-ACPR and its boundary is called the MAP boundary. The BP-GEXIT curve and the upper bound on the MAP threshold for the $\mathrm{LDPC}(3,6)$ ensemble is shown in Fig. 20, for $\theta = 1$. Using this procedure, we can compute an outer bound to the MAP boundary by considering different values of θ.

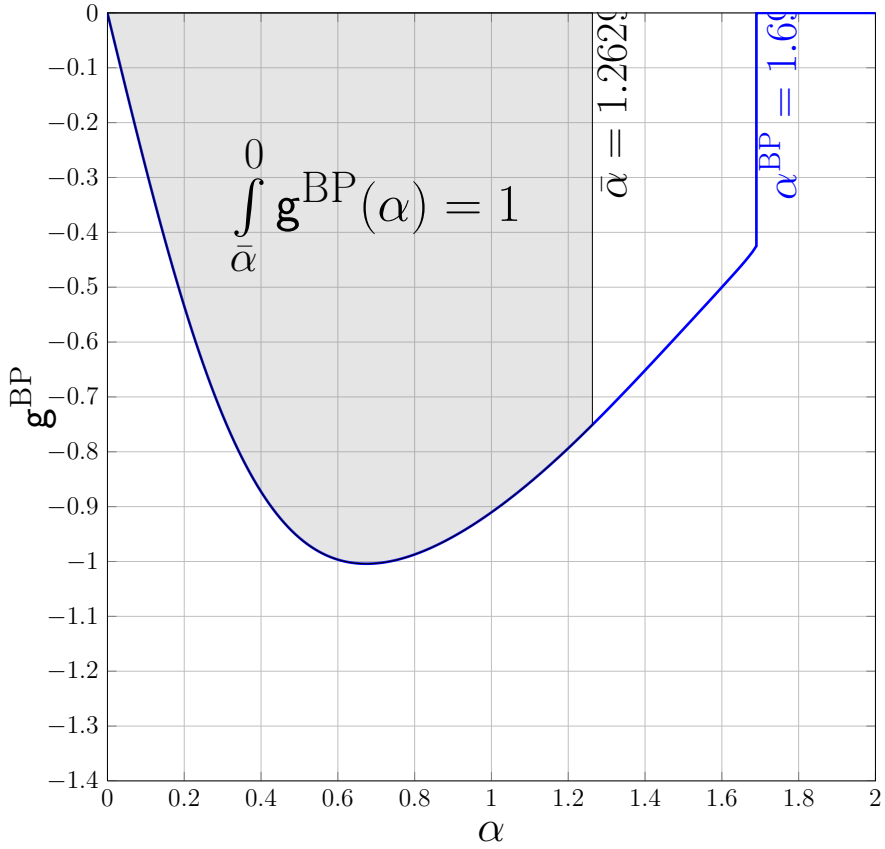

Fig. 20. The BP-GEXIT curve for the regular LDPC(3, 6) ensemble and the upper bound on the MAP threshold is shown above for $\theta = 1$. GEXIT curves in literature are typically parametrized by the channel entropy and the channels get *worse* as the entropy increases. However, the channel gains are a natural parameterization for this problem and the channel gets *better* by increasing the channel gains. So the GEXIT values are negative for this parametrization.

CHAPTER IV

THRESHOLD SATURATION AND SPATIAL COUPLING

The phenomenon of threshold saturation was introduced by Kudekar et al. [44] to explain the impressive performance of convolutional LDPC ensembles [45, 46]. They observed that the belief-propagation (BP) threshold of a spatially-coupled ensemble is very close to the maximum-a-posteriori (MAP) threshold of its underlying ensemble; a similar statement was formulated independently, as a conjecture in [47]. This phenomenon has been termed "threshold saturation via spatial coupling". Kudekar et al. prove in [44] that threshold saturation occurs for the binary erasure channel (BEC) and a particular convolutional LDPC ensemble. For general binary-input memoryless symmetric (BMS) channels, threshold saturation was empirically observed first [48,49] and then shown analytically [50]. It is known that the MAP threshold of regular LDPC codes approaches the Shannon limit for binary memoryless symmetric (BMS) channels with increasing left degree, while keeping the rate fixed (though such codes have a vanishing BP threshold) [44]. So, spatial coupling appears to provide us with a new paradigm to construct capacity approaching codes for BMS channels.

From the observation in Section 3, spatially-coupled codes are potential candidates for universal codes, for multi-terminal problems.

A. Spatially Coupled Codes

This section describes spatially-coupled codes and is included here for completeness. The material closely follows the description in [44].

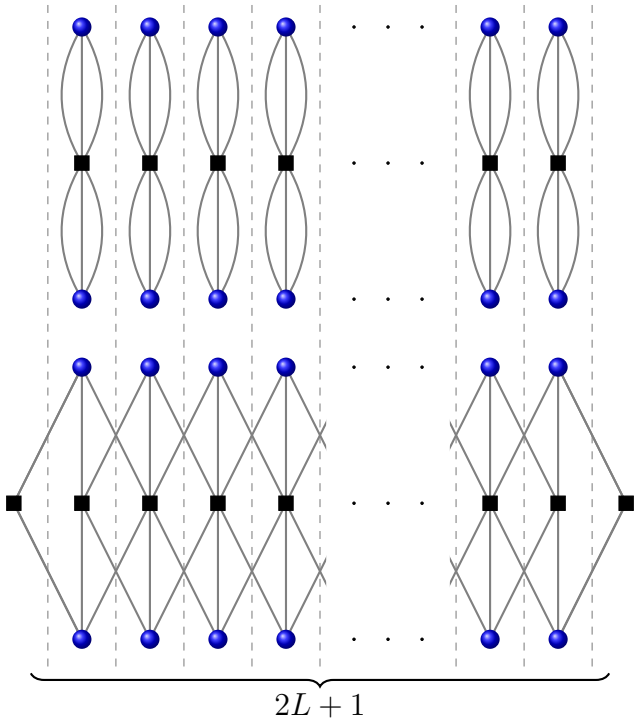

Fig. 21. The protograph of the $(3, 6, L)$ ensemble is shown above.

1. The $(\mathtt{l}, \mathtt{r}, L)$ Ensemble

Consider ensembles with parameters $(\mathtt{l}, \mathtt{r})$, where $\mathtt{l} \leq \mathtt{r}$ is odd. Choose M such that $\frac{\mathtt{l}}{\mathtt{r}}M$ is an integer. Place variable nodes at positions $[-L, L] \triangleq \{-L, -L+1, \cdots, L\}$, such that there are M variable nodes at each position. Define $\hat{\mathtt{l}} = \frac{\mathtt{l}-1}{2}$. Place check nodes at positions $[-L-\hat{\mathtt{l}}, L+\hat{\mathtt{l}}]$, with $\frac{\mathtt{l}}{\mathtt{r}}M$ check nodes at each position. Each of the $\mathtt{l}$ edges of a variable node at position i is connected to exactly one check node at position $i-\hat{\mathtt{l}}, \cdots, i+\hat{\mathtt{l}}$. For each check node position in $[-L-\hat{\mathtt{l}}, L+\hat{\mathtt{l}}]$, there are $\frac{\mathtt{l}}{\mathtt{r}}M\mathtt{r} = M\mathtt{l}$ sockets. The probability distribution on the ensemble is defined by choosing a random permutation on the edges at each check node position. The protograph is shown in Fig. 21. The BP-GEXIT curves are shown for the $(4, 6, L)$ ensemble for transmission over erasure channels in Fig. 22. As seen in Fig. 22 the

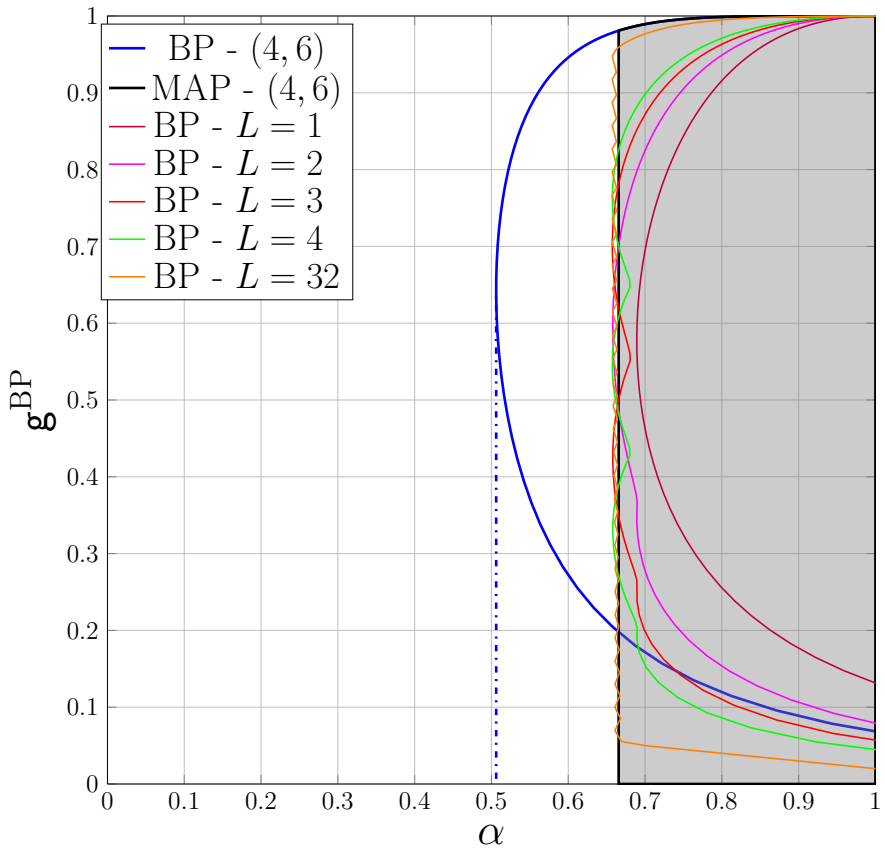

Fig. 22. The BP-GEXIT curves for the spatially coupled $(4, 6, L)$ are shown above for transmission over the BEC. We observe that the BP threshold of the spatially-coupled codes saturates towards the MAP threshold of the $(4, 6)$ ensemble.

GEXIT curves exhibit wiggles which prevent it from saturating to the MAP threshold of the underlying ensemble. It turns out that the size of these wiggles does not decay with L [44].

2. The $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble

This ensemble is defined to simplify the analysis of the $(\mathtt{l}, \mathtt{r}, L)$ ensemble and overcome the gap due to the wiggles. Although this ensemble loses the protograph structure, the analysis is much simpler and its BP threshold does saturate towards the MAP

threshold of the underlying ensemble. As before, the variable nodes are placed at positions $[-L, L]$, with M nodes at each position. The check nodes are placed at positions $[-L, L + w - 1]$ (w can be thought of as a "smoothing" parameter).

Each of the $\mathtt{l}$ connections, of a variable node at position i, are uniformly and independently chosen from $[i, i + w - 1]$. Define the *type* of a variable node by a w-tuple $t = (t_0, t_1, \cdots, t_{w-1})$ of non-negative integers such that $\sum t_j = l$. This means that the variable node has t_j edges that connect to a check node at position $t + j$. Note that these edges are not ordered. Assume that we fix an arbitrary order for the edges of each variable node. We can then define the constellation of a variable node by an $\mathtt{l}$-tuple $c = (c_1, \cdots, c_{\mathtt{l}})$, with elements in $[0, w-1]$, which means that the k-th edge is connected to a check node at position $i + c_k$. Note that there are many constellations for a given type (permute the elements of c). Let $\tau(c)$ denote the type of a constellation. A uniform distribution is imposed on the set of all constellations (due to the requirement that each edge is chosen independently). This induces a probability distribution on the types, given by

$$p(t) = \frac{|\{c|\tau(c) = t\}|}{w^{\mathtt{l}}}.$$

Choose M so that $Mp(t)$ is an integer for all t. For each position $i \in [-L, L]$, pick $Mp(t)$ variable nodes of type t. A random permutation is chosen to map the type to a constellation.

For each check position i, away from the boundary, the number of edges that come from variable nodes at position $i - j, j \in [0, w-1]$ is $M\frac{\mathtt{l}}{w}$ i.e., it is a fraction $\frac{1}{w}$ of the Ml sockets at position i. These edges are mapped to the sockets by choosing a uniform random permutation of size $M\mathtt{l}$.

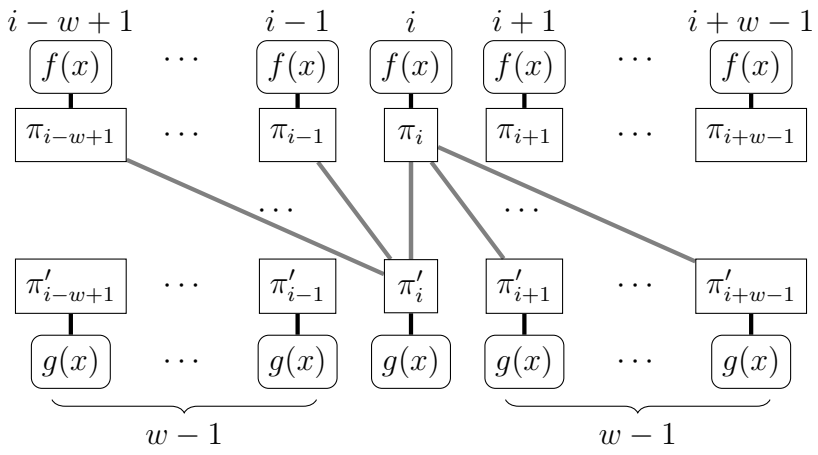

Fig. 23. A portion of a generic SC system. The f-node at position i is coupled with the g-nodes at positions $i-w+1,\ldots,i$ and, by reciprocity, g-node at position i is coupled with the f-nodes at positions $i,\ldots,i+w-1$. Here, π_i and π'_i are random permutations.

The design rate of this ensemble is given by [44]

$$R(\mathtt{l},\mathtt{r},L,w)=\left(1-\frac{\mathtt{l}}{\mathtt{r}}\right)+\frac{\mathtt{l}}{\mathtt{r}}\frac{\left[(w+1)-2\sum_{i=0}^{w}\left(1-\left(\frac{w-i-1}{w}\right)^{\mathtt{r}}\right)\right]}{2L+1}.$$

Note that this is a lower bound on the true rate of the code. This construction can be extended to irregular LDPC(Λ,P) codes as shown in Fig. 23. The f-nodes at each position are replaced by M copies of the node degree profile $\Lambda(x)=\sum_i \Lambda_i x^i$, where Λ_i is the number of bit nodes of degree i. The g-nodes at each position are replaced by M copies of the node degree profile $P(x)=\sum_i P_i x^i$, where P_i is the number of check nodes of degree i. For sufficiently large M, these nodes can be coupled uniformly using an averaging window of length w in a manner similar to the $(\mathtt{l},\mathtt{r},L,w)$ ensemble defined above.

3. Density Evolution of the $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble

Let $\mathsf{a}_i^{(\ell)}$ denote the density of the messages emitted by a variable node at position i at iteration ℓ. Set $\mathsf{a}_i = \Delta_{+\infty}$, if $i \notin [-L, L]$. For $i \in [-L, L]$, we can write down the density evolution equations

$$\mathsf{a}_i^{(\ell+1)} = \mathsf{a}_{\mathrm{BMSC}} \circledast \lambda\left(\frac{1}{w}\sum_{j=0}^{w-1} \rho\left(\frac{1}{w}\sum_{k=0}^{w-1} \mathsf{a}_{i+j-k}^{(\ell)}\right)\right). \tag{4.1}$$

It is observed in [44] that the size of the wiggles reduces with w.

B. A Simple Proof of Threshold Saturation

In this section, we provide a simple proof of threshold saturation via spatial-coupling for a broad class of vector recursions over erasure-type channels. The main tool is a potential theory for vector recursions that extends naturally to coupled systems of vector recursions.

1. Notation

The following notation is used throughout this section. We let $d \in \mathbb{N}$ be the dimension of the vector recursion, $\mathcal{X} \triangleq [0,1]^d$ be the space on which the recursion is defined, and $\mathcal{E} \triangleq [0,1]$ be the parameter space of the recursive system. We also use $\mathcal{X}_\circ$ and $\mathcal{E}_\circ$ to denote $\mathcal{X} \setminus \{\mathbf{0}\}$ and $\mathcal{E} \setminus \{0\}$ respectively. Vectors are considered to be row vectors and are denoted in boldface (e.g. $\boldsymbol{x}$). For two vectors $\boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{X}$, the partial orders $\boldsymbol{x} \succeq \boldsymbol{y}$ and $\boldsymbol{x} \preceq \boldsymbol{y}$ are defined by $x_i \geq y_i$ for $i = 1, \ldots, d$ and $x_i \leq y_i$ for $i = 1, \ldots, d$, respectively. We use lower case (e.g., $f(\boldsymbol{x})$) to denote scalar functions of a vector argument and lower case bold (e.g., $\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), \cdots, f_d(\boldsymbol{x})]$) to denote a vector function of a vector argument. The gradient of a scalar function is denoted by an apostrophe and is defined by $f'(\boldsymbol{x}) \triangleq [\partial f(\boldsymbol{x})/\partial x_1, \cdots, \partial f(\boldsymbol{x})/\partial x_d]$, and the Jacobian

of a vector function is defined by

$$\boldsymbol{f}'(\boldsymbol{x}) = \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} \triangleq \begin{bmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\boldsymbol{x})}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_d(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_d(\boldsymbol{x})}{\partial x_d} \end{bmatrix}.$$

If $\boldsymbol{X}$ is a matrix, we use the notation $\boldsymbol{x}_i$ or $[\boldsymbol{X}]_i$ to denote the i-th row of $\boldsymbol{X}$ and $x_{i,j}$ to denote the (i,j)-th element of $\boldsymbol{X}$. Abusing notation, we also allow vector functions to take matrix arguments and define $\boldsymbol{f}(\boldsymbol{X})$ via $[\boldsymbol{f}(\boldsymbol{X})]_i = \boldsymbol{f}(\boldsymbol{x}_i)$. We use the notation $\mathrm{vec}(\boldsymbol{X})$ to denote the transpose of the vector obtained by stacking the columns of $\boldsymbol{X}$ together. The Jacobian of a matrix function is defined by

$$\boldsymbol{f}'(\boldsymbol{X}) = \frac{\partial \mathrm{vec}(\boldsymbol{f}(\boldsymbol{X}))}{\partial \mathrm{vec}(\boldsymbol{X})},$$

and the Hessian of a vector function is defined by

$$\boldsymbol{f}''(\boldsymbol{x}) = \frac{\partial \mathrm{vec}(\boldsymbol{f}'(\boldsymbol{x}))}{\partial \boldsymbol{x}}.$$

2. Single System Potential

First, we define potential functions for a class of vector recursions and discuss threshold parameters associated with the potential.

Definition IV.1. An *admissible vector system* $(\boldsymbol{f}, \boldsymbol{g})$ parametrized by $\epsilon \in \mathcal{E}$, is defined by the recursion

$$\boldsymbol{x}^{(\ell+1)} = \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}^{(\ell)}); \epsilon), \tag{4.2}$$

where $\boldsymbol{f} = [f_1, \cdots, f_d]$ and $\boldsymbol{g} = [g_1, \cdots, g_d]$. Here, $f_i : \mathcal{X} \times \mathcal{E} \to [0,1]$ is strictly increasing in all its arguments for $\boldsymbol{x} \in \mathcal{X}_\circ$, $\epsilon \in \mathcal{E}_\circ$, and $g_i : \mathcal{X} \to [0,1]$, $i = 1, \cdots, d$, satisfies $g_i'(\boldsymbol{x}) \succ 0$ for $\boldsymbol{x} \in \mathcal{X}_\circ$. We also assume that $\boldsymbol{f}(\boldsymbol{0}; \epsilon) = \boldsymbol{g}(\boldsymbol{0}) = \boldsymbol{f}(\boldsymbol{x}; \boldsymbol{0}) = \boldsymbol{0}$,

and that $\boldsymbol{f}, \boldsymbol{g}$ have bounded and continuous second differentials w.r.t. all arguments.

Definition IV.2. Suppose there exist functionals $F : \mathcal{X} \times \mathcal{E} \to \mathbb{R}, G : \mathcal{X} \to \mathbb{R}$ such that $F' = \boldsymbol{f}$ and $G' = \boldsymbol{g}$. Then, the single-system *potential function* $U(\boldsymbol{x}; \epsilon)$ of an admissible vector system $(\boldsymbol{f}, \boldsymbol{g})$ is defined by

$$\begin{aligned} U(\boldsymbol{x}; \epsilon) &\triangleq \int_0^{\boldsymbol{x}} [\boldsymbol{g}'(\boldsymbol{z}) (\boldsymbol{z} - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{z}); \epsilon))] \cdot \mathrm{d}\boldsymbol{z} \\ &= \boldsymbol{x} \cdot \boldsymbol{g}(\boldsymbol{x}) - G(\boldsymbol{x}) - F(\boldsymbol{g}(\boldsymbol{x}); \epsilon). \end{aligned} \tag{4.3}$$

Definition IV.3. For $\boldsymbol{x} \in \mathcal{X}, \epsilon \in \mathcal{E}$, we have the following terms.

- For fixed ϵ, $\boldsymbol{x}$ is a *fixed point* (f.p.) iff $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}); \epsilon)$.
- For fixed ϵ, $\boldsymbol{x}$ is a *stationary point* (s.p.) if $U'(\boldsymbol{x}; \epsilon) = \boldsymbol{0}$.

Lemma IV.1. *The potential function of an admissible vector system has the following properties:*

1. *$U(\boldsymbol{x}; \epsilon)$ is strictly decreasing in ϵ, for $\epsilon \in \mathcal{E}_\circ$.*
2. *An $\boldsymbol{x} \in \mathcal{X}_\circ$, such that $x_i > 0, \forall\, i$, is a f.p. iff it is a s.p. of the potential.*

Proof. These properties hold because the potential function is the scaled line integral of the DE update $\boldsymbol{g}'(\boldsymbol{z}) (\boldsymbol{z} - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{z}); \epsilon))$, which is strictly decreasing in ϵ, for $\epsilon \in \mathcal{E}_\circ$, and zero iff $\boldsymbol{z}$ is a fixed point of the recursion. ■

Definition IV.4. Let $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{x}^{(0)} = \boldsymbol{x}$. Denote by $\boldsymbol{x}^\infty(\boldsymbol{x}; \epsilon)$, the limit (if it exists) of $\boldsymbol{x}^{(\ell+1)} = \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}^{(\ell)}); \epsilon)$.

Lemma IV.2. *Consider the recursion (4.2), with $\boldsymbol{x}^{(0)} = \boldsymbol{1}$. Then $\boldsymbol{x}^{(\ell)}$ converges to $\boldsymbol{x}^\infty(\boldsymbol{1}; \epsilon)$.*

Proof. Note that $\boldsymbol{x}^{(1)} = \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}^{(0)}; \epsilon) \preceq \boldsymbol{1} = \boldsymbol{x}^{(0)}$ as $\boldsymbol{1}$ is the greatest element of $\mathcal{X}$. It follows by induction on ℓ that $\boldsymbol{x}^{(0)} \succeq \boldsymbol{x}^{(1)} \succeq \cdots \succeq \boldsymbol{x}^{(\ell)} \succeq \cdots \succeq \boldsymbol{0}$. Hence the sequence has a limit $\boldsymbol{x}^{(\infty)} = \boldsymbol{x}^\infty(\boldsymbol{1}; \epsilon)$. ■

Definition IV.5. The *single-system threshold* is defined to be

$$\epsilon_s^* = \sup\left\{\epsilon \in \mathcal{E} \,|\, \boldsymbol{x}^\infty(\mathbf{1};\epsilon) = \mathbf{0}\right\},$$

and is the ϵ-threshold for convergence of the single system recursion to $\mathbf{0}$.

Remark IV.1. The recursion (4.2) has no f.p.s in $\mathcal{X}_\circ$ iff $\epsilon < \epsilon_s^*$. For DE recursions associated with BP decoding, the threshold ϵ_s^* is called the BP threshold.

Definition IV.6. The *basin of attraction* for $\mathbf{0}$ is defined by

$$\mathcal{U}_x(\epsilon) = \left\{\boldsymbol{x} \in \mathcal{X} \,|\, \boldsymbol{x}^\infty(\boldsymbol{x};\epsilon) = \mathbf{0}\right\}.$$

Notice that this equals $\mathcal{X}$ if $\epsilon < \epsilon_s^*$ but it is a strict subset of $\mathcal{X}$ if $\epsilon \geq \epsilon_s^*$.

Definition IV.7. We define the *energy gap* $\Delta E(\epsilon) = \inf\left\{U(\boldsymbol{x};\epsilon) \,|\, x \in \mathcal{X} \setminus \mathcal{U}_x(\epsilon)\right\}$ and the potential threshold

$$\epsilon^* = \sup\left\{\epsilon \in (\epsilon_s^*, 1] \,|\, \Delta E(\epsilon) > 0\right\}. \tag{4.4}$$

Since $\Delta E(\epsilon)$ is strictly decreasing in ϵ, this is well defined and $\epsilon < \epsilon^*$ implies $\Delta E(\epsilon) > 0$. For DE recursions associated with BP decoding, the potential threshold is analogous to the threshold predicted by the Maxwell conjecture [51, Conj. 1].

Example IV.1. For the standard irregular ensemble of LDPC codes (e.g., see [2]), the DE recursion,

$$x^{(\ell+1)} = \epsilon\lambda(1 - \rho(1 - x^{(\ell)})),$$

is an admissible scalar system with $d = 1$, $f(x;\epsilon) = \epsilon\lambda(x)$ and $g(x) = 1 - \rho(1-x)$. In this case, the single system potential is given by (4.3) and shown in Fig. 24 for the $(3,6)$-regular LDPC code ensemble defined by $(\lambda,\rho) = (x^2, x^5)$.

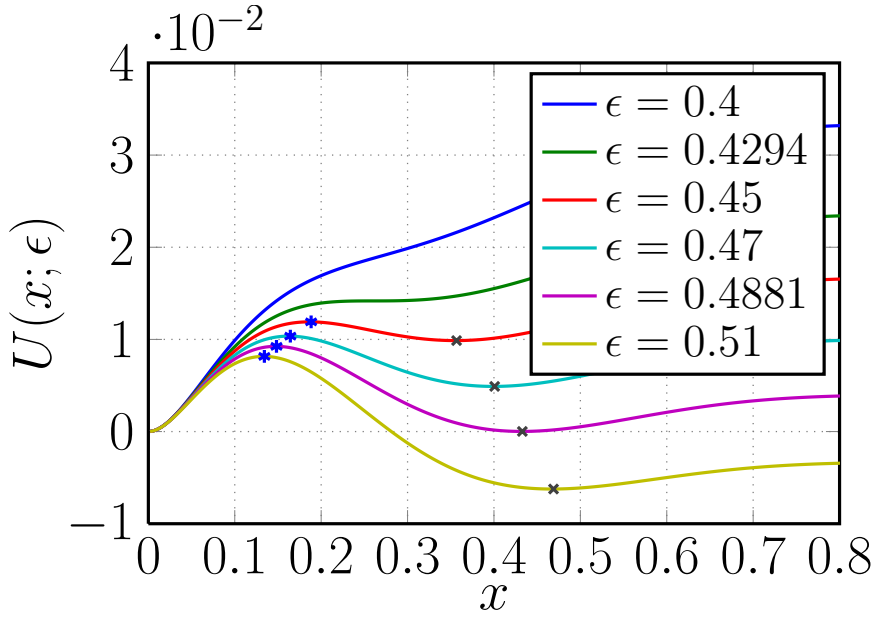

Fig. 24. The potential function of the (3,6)-regular LDPC ensemble is shown for a range of ϵ. Here $\epsilon_s^* \approx 0.4294$, $\epsilon^* \approx 0.4881$, and the stationary points are marked. Notice that, for $\epsilon < \epsilon_s^*$, $U(x;\epsilon)$ has no stationary points.

3. Coupled System Potential

Now, we extend our definition of potential functions to coupled systems of vector recursions. In particular, we consider a "spatial-coupling" of the single system recursion, (4.2), that leads to the recursion (4.5) and a closely related matrix recursion (4.6). For the matrix recursion of the coupled system, we define a potential function and show that, for $\epsilon < \epsilon^*$, the only fixed point of the coupled system is the zero matrix.

Definition IV.8 (cf. [44])**.** The basic *spatially-coupled vector system* is defined by placing $2L+1$ single systems at positions in the set $\mathcal{L} = \{-L, -L+1, \ldots, L\}$ and coupling them with w systems on one side as shown in Fig. 23. Let $\boldsymbol{x}_i^{(\ell)}$ be the input to the g-function in the i-th position after $\ell+1$ iterations and define $\boldsymbol{x}_i^{(\ell)} = 0$ for $i \notin \mathcal{L} \triangleq \{-L, -L+1, \ldots, L+w-1\}$ and all ℓ. For the coupled system, this leads to

the recursion

$$\boldsymbol{x}_i^{(\ell+1)} = \frac{1}{w}\sum_{k=0}^{w-1} \boldsymbol{f}\left(\frac{1}{w}\sum_{j=0}^{w-1} \boldsymbol{g}(\boldsymbol{x}_{i+j-k}^{(\ell)}); \epsilon_{i-k}\right), \tag{4.5}$$

where $\boldsymbol{x}_i^{(0)} = \mathbf{1}$ for $i \in \mathcal{L}$ and $\boldsymbol{x}_i^{(\ell)} = \mathbf{0}$ for $i \notin \mathcal{L}$ and all ℓ. Also $\epsilon_i = \epsilon$ for $i \in \mathcal{L}_0$ and $\epsilon_i = 0$ for $i \notin \mathcal{L}_0$.

Definition IV.9 (cf. [44])**.** Let $i_0 \triangleq \left\lfloor \frac{w-1}{2} \right\rfloor$. The *one-sided spatially-coupled vector system* is a modification of (4.5) defined by fixing the values of positions outside $\mathcal{L}' = \{-L, L+1, \ldots, i_0\}$. It fixes the left boundary to zero by defining $\boldsymbol{x}_i^{(\ell)} = \mathbf{0}$ for $i < -L$ and all ℓ. It forces the right boundary to a floating constant by setting $\boldsymbol{x}_i^{(\ell)} = \boldsymbol{x}_0^{(\ell)}$ for $i \geq 1$ and all ℓ.

Definition IV.10 (cf. [52])**.** Let the *matrix one-sided SC vector recursion* be

$$\boldsymbol{X}^{(\ell+1)} = \boldsymbol{A}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{A}\boldsymbol{g}(\boldsymbol{X}^{(\ell)}); \epsilon), \tag{4.6}$$

where $\boldsymbol{X} = [\boldsymbol{x}_{-L-w}^{\mathsf{T}}, \cdots, \boldsymbol{x}_{2w+i_0}^{\mathsf{T}}]^{\mathsf{T}}$ and $\boldsymbol{A}$ is the $(L+3w+i_0+1) \times (L+3w+i_0+1)$ matrix given by

$$\boldsymbol{A} = \frac{1}{w}\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 & 1 & \ddots & 1 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Remark IV.2. The right hand side of (4.6) accurately represents a single iteration of the one-sided SC system update for $i \in \mathcal{L}'$, but cannot be used recursively because

the boundary condition $\boldsymbol{x}_i^{(\ell)} = \boldsymbol{x}_0^{(\ell)}$ for $i \geq 1$ is not preserved after the first step.

Lemma IV.3 (cf. [44, Lem. 14])**.** *For both the basic and one-sided SC systems, the recursions are component-wise decreasing with iteration and converge to well-defined fixed points. The one-sided recursion defined in Def. IV.9 is also a component-wise upper bound on the basic SC recursion for $i \in \mathcal{L}$ and it converges to a non-decreasing fixed-point vector.*

Sketch of Proof. The proof follows from the monotonicity of $\boldsymbol{f}, \boldsymbol{g}$. For the one-sided SC system, the right boundary condition is also needed to show this result. ■

Definition IV.11. The *coupled system potential* for general matrix recursions, in the form of (4.6), is given by

$$U(\boldsymbol{X}; \epsilon) = \mathrm{Tr}(\boldsymbol{X}^\intercal \boldsymbol{g}(\boldsymbol{X})) - G(\boldsymbol{X}) - F(\boldsymbol{A}\boldsymbol{g}(\boldsymbol{X}); \epsilon),$$

where $G(\boldsymbol{X}) = \sum_i G(\boldsymbol{x}_i)$ and $F(\boldsymbol{X}; \epsilon) = \sum_i F(\boldsymbol{x}_i; \epsilon)$.

Remark IV.3. A key observation of this work is that a potential function for coupled vector systems can be written in the simple form given in Def. IV.11. Remarkably, this holds for general coupling coefficients because of the $\boldsymbol{A}, \boldsymbol{A}^\intercal$ reciprocity that appears naturally in SC.

Lemma IV.4. *Let $\boldsymbol{X} \in \mathcal{X}^n$ be a matrix with non-decreasing columns generated by averaging the rows of $\boldsymbol{Z} \in \mathcal{X}^n$ over a sliding window of size w. Let the down-shift operator $\boldsymbol{S}_n : \mathcal{X}^n \to \mathcal{X}^n$ be defined by $[\boldsymbol{S}_n \boldsymbol{X}]_1 = \boldsymbol{0}$ and $[\boldsymbol{S}_n \boldsymbol{X}]_i = \boldsymbol{x}_{i-1}$ for $i = 2, \ldots, n$. Then, we have the bounds $\|vec(\boldsymbol{S}_n \boldsymbol{X} - \boldsymbol{X})\|_\infty \leq \frac{1}{w}$ and $\|vec(\boldsymbol{S}_n \boldsymbol{X} - \boldsymbol{X})\|_1 = \|\boldsymbol{x}_n\|_1 = \|\boldsymbol{X}\|_\infty$.*

Proof. The bound $\|\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})\|_\infty \leq \frac{1}{w}$ follows from

$$|x_{i,j} - x_{i-1,j}| = \left| \tfrac{1}{w} \sum\nolimits_{k=0}^{w-1} z_{i+k,j} - \tfrac{1}{w} \sum\nolimits_{k=0}^{w-1} z_{i-1+k,j} \right| \leq \tfrac{1}{w}.$$

Since the columns of $\boldsymbol{X}$ are non-decreasing, the 1-norm sum telescopes and we get $\|\mathrm{vec}(\boldsymbol{S}\boldsymbol{X}-\boldsymbol{X})\|_1 = \|\boldsymbol{x}_n\|_1 = \|\boldsymbol{X}\|_\infty$. ■

Lemma IV.5. *For the vector one-sided SC system, a shift changes the potential by* $U(\boldsymbol{S}\boldsymbol{X};\epsilon) - U(\boldsymbol{X};\epsilon) = -U(\boldsymbol{x}_{i_0};\epsilon)$.

Proof. First, we rewrite the potential as the summation

$$U(\boldsymbol{X};\epsilon) = \sum\nolimits_{i=-L-w}^{2w} \left[\boldsymbol{g}(\boldsymbol{x}_i)\cdot\boldsymbol{x}_i - G(\boldsymbol{x}_i) - F\left([\boldsymbol{A}\boldsymbol{g}(\boldsymbol{X})]_i;\epsilon\right)\right].$$

Since the first w rows of $\boldsymbol{X}$ are $\mathbf{0}$ and the last $2w+1$ rows of $\boldsymbol{X}$ equal $\boldsymbol{x}_0$, it can be shown that $\sum_{i=-L-w}^{2w+i_0} F\left([\boldsymbol{A}\boldsymbol{g}(\boldsymbol{S}\boldsymbol{X})]_i;\epsilon\right) - F\left([\boldsymbol{A}\boldsymbol{g}(\boldsymbol{X})]_i;\epsilon\right) = F(\boldsymbol{g}(\mathbf{0});\epsilon) - F\left(\boldsymbol{g}(\boldsymbol{x}_{i_0});\epsilon\right)$. Thus, we have

$$U(\boldsymbol{S}\boldsymbol{X};\epsilon) - U(\boldsymbol{X};\epsilon) = U(\mathbf{0};\epsilon) - U(\boldsymbol{x}_{2w};\epsilon) = -U(\boldsymbol{x}_0;\epsilon).$$

■

Lemma IV.6. *The norm of the Hessian* $U''(\boldsymbol{X};\epsilon)$ *of the SC potential is bounded by a constant independent of* L *and* w *and satisfies*

$$\|U''(\boldsymbol{X};\epsilon)\|_\infty \triangleq \|\boldsymbol{g}'\|_\infty + \|\boldsymbol{g}'\|_\infty^2\|\boldsymbol{f}'\|_\infty + \|\boldsymbol{g}''\|_\infty,$$

where $\|\boldsymbol{h}\|_\infty = \sup_{x\in\mathcal{X}} \max_i |h_i(\boldsymbol{x})|$ *for functions* $\boldsymbol{h}:\mathcal{X}\to\mathbb{R}$. *We also define* $K_{\boldsymbol{f},\boldsymbol{g}} = \|\boldsymbol{g}'\|_\infty + \|\boldsymbol{g}'\|_\infty^2\|\boldsymbol{f}'\|_\infty + \|\boldsymbol{g}''\|_\infty$.

Proof. First note that

$$U''(\boldsymbol{X};\epsilon) = \frac{\partial \mathrm{vec}(U'(\boldsymbol{X};\epsilon))}{\partial \mathrm{vec}(\boldsymbol{X})}.$$

By direct computation, we obtain

$$\|U''(\boldsymbol{X};\epsilon)\|_\infty \le \|g'\|_\infty + \|g'\|_\infty^2\|f'\|_\infty + \|g''\|_\infty.$$

■

We now state the main result of this chapter. Roughly speaking, it says that, if $\epsilon < \epsilon^*$ and w is sufficiently large, then one can always decrease the coupled potential of a non-zero matrix by down shifting. Since this implies that the next step of the recursion must reduce some value, the only valid fixed point is the zero matrix.

Theorem IV.7. *Consider an admissible vector system* $(\boldsymbol{f}, \boldsymbol{g})$. *If* $\epsilon < \epsilon^*$ *and* $w > dK_{\boldsymbol{f},\boldsymbol{g}}/\Delta E(\epsilon)$, *then the only fixed point of the spatially-coupled system, defined by* (4.5), *is* $\boldsymbol{X} = \boldsymbol{0}$.

Proof. Using Lem. IV.3, let $\boldsymbol{X}$ be the unique fixed point of the one-sided recursion defined in Def. IV.9. This fixed point upper bounds the fixed point of the basic SC system defined in Def. IV.8. If $\boldsymbol{X} \neq \boldsymbol{0}$, then $\boldsymbol{x}_0 \notin \mathcal{U}_x(\epsilon)$ because the system has no fixed points with $\boldsymbol{x}_i \in \mathcal{U}_x(\epsilon)$ for all i. From Lemma IV.5, we have $\Delta U \triangleq U(\boldsymbol{S}\boldsymbol{X};\epsilon) - U(\boldsymbol{X};\epsilon) < -U(\boldsymbol{x}_{i_0};\epsilon)$. Expanding $U(\boldsymbol{S}\boldsymbol{X};\epsilon)$ in a Taylor series (with remainder) around $\boldsymbol{X}$, we get

$$
\begin{aligned}
&\mathrm{vec}(U'(\boldsymbol{X};\epsilon)) \cdot \mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X}) = U(\boldsymbol{S}\boldsymbol{X};\epsilon) - U(\boldsymbol{X};\epsilon) \\
&\quad - \int_0^1 (1-t)\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})^\intercal U''(\boldsymbol{X}(t);\epsilon)\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})\,\mathrm{d}t \\
&\leq \Delta U + \left| \int_0^1 (1-t)\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})^\intercal U''(\boldsymbol{X}(t);\epsilon)\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})\,\mathrm{d}t \right| \\
&\leq \Delta U + \|\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})\|_1 \max_{t\in[0,1]} \|U''(\boldsymbol{X}(t);\epsilon)\|_\infty \|\mathrm{vec}(\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X})\|_\infty \\
&\leq -U(\boldsymbol{x}_{i_0};\epsilon) + \frac{\|\boldsymbol{x}_{i_0}\|_1}{w} \max_{t\in[0,1]} \|U''(\boldsymbol{X}(t);\epsilon)\|_\infty \\
&\leq -U(\boldsymbol{x}_{i_0};\epsilon) + \frac{dK_{\boldsymbol{f},\boldsymbol{g}}}{w} \\
&< -U(\boldsymbol{x}_{i_0};\epsilon) + \Delta E(\epsilon) \leq 0,
\end{aligned}
$$

where the last steps hold because $w > dK_{\boldsymbol{f},\boldsymbol{g}}/\Delta E(\epsilon)$ and $\|\boldsymbol{x}_{i_0}\|_1 \leq d$ and $U(\boldsymbol{x};\epsilon) \geq \Delta E(\epsilon)$ for $\boldsymbol{x} \in \mathcal{X} \setminus \mathcal{U}_x(\epsilon)$.

Now, we observe that $\boldsymbol{SX} - \boldsymbol{X} \preceq \mathbf{0}$ (i.e., the fixed point is non-decreasing) and $[\boldsymbol{SX} - \boldsymbol{X}]_i$ is zero for $i \notin \mathcal{L}'$. So, $U'(\boldsymbol{X}; \epsilon)$ is positive in at least one entry of one row (i.e., there exists $i \in \mathcal{L}'$ such that $[U'(\boldsymbol{X}; \epsilon)]_i > \mathbf{0}$). Since $[U'(\boldsymbol{X}; \epsilon)]_i = (\boldsymbol{x}_i - [\boldsymbol{A}^\intercal \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{AX}); \epsilon)]_i)\boldsymbol{g}'(\boldsymbol{x}_i)$, it follows that $[\boldsymbol{g}'(\boldsymbol{x}_i)]_j > 0$ and $[\boldsymbol{A}^\intercal \boldsymbol{f}(\boldsymbol{Ag}(\boldsymbol{X}); \epsilon)]_{i,j} < \boldsymbol{X}_{i,j}$. Therefore, one more decoding iteration must reduce the value of the i-th component for some $i \in \mathcal{L}'$. This contradicts the fact that $\boldsymbol{X}$ is a fixed point and shows that the only fixed point of the one-sided SC system is $\boldsymbol{X} = \mathbf{0}$. Since the fixed point of the basic SC system is upper bounded by this, we conclude that it must also be zero. ■

CHAPTER V

APPLICATIONS OF SPATIAL COUPLING*

A. The Noisy Slepian-Wolf Problem

Spatial coupling is most easily described by the $(\mathtt{l}, \mathtt{r}, L)$ ensemble interms of protographs [44, 53]. We briefly review the protograph structure at the joint decoder here. Consider the protograph of a standard LDPC$(4, 6)$ ensemble. There are two check nodes and three variable nodes. For each user, take a collection of $(2L + 1)$ protographs at positions $\mathcal{L} \triangleq \{-L, \cdots, L\}$ and couple them as described in [44]. One variable node at each position $i \in \mathcal{L}$ from the first user is punctured and connected to a punctured variable node at the same position of the second user. The resulting protograph, shown in Fig. 25, is then expanded M times to form the parity-check matrix of the joint system. This structure is fundamental to the phenomenon of threshold saturation observed at the joint decoder. It is simply not sufficient to use spatially-coupled codes with random connections between the information nodes. Such a coupling will only result in pushing the threshold of the component codes to the MAP threshold, but may have little effect on the BP threshold of the joint system.

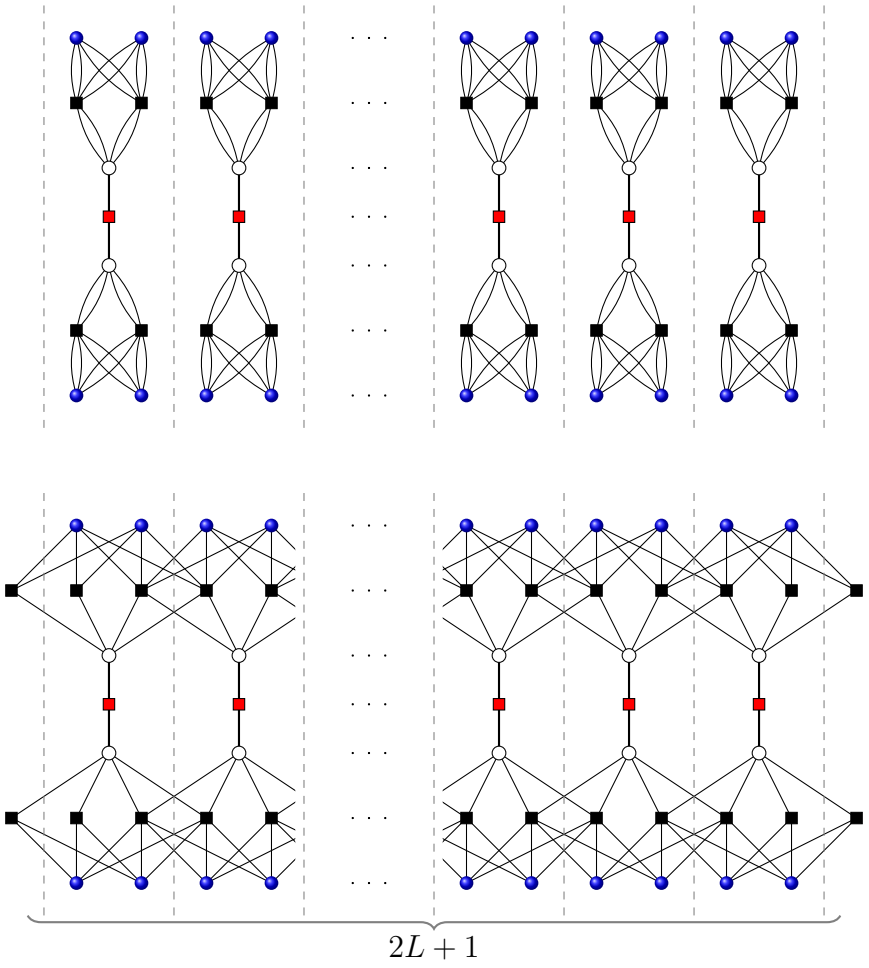

Fig. 25. The protograph of the joint decoder for the $(4, 6, L)$ ensemble is shown above for the noisy SW problem.

1. The $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble

The $(\mathtt{l}, \mathtt{r}, L, w)$ spatially-coupled ensemble can be described as follows: Place M variable nodes at each position in $[-L, L]$. The check nodes are placed at positions $[-L, L + w - 1]$, with $\frac{\mathtt{l}}{\mathtt{r}}M$ check nodes at each position. The connections are made as described in [44]. This procedure generates a Tanner graph for the $(\mathtt{l}, \mathtt{r}, L, w)$ ensemble.

For this work we consider codes of rate 1/3, punctured to a rate 1/2. Two such

graphs (generated by the above procedure) are taken and $2M/3$ variable nodes ($M/3$ from each graph) at each position are connected by a random (uniform) permutation of size $M/3$ via correlation nodes. This procedure ensures that all the variable node positions are symmetric (as opposed to Fig. 25) with respect to puncturing and correlation, enabling us to write down the density evolution (DE) equations as described in the following section.

2. Density Evolution of the $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble and GEXIT Curves

Let $\mathsf{a}_i^{(\ell)}$ and $\mathsf{b}_i^{(\ell)}$ denote the average density emitted, after ℓ iterations of decoding, by variable nodes at position i, at iteration ℓ, for codes 1 and 2 respectively. Let $\Delta_{+\infty}$ denote the delta function at $+\infty$ and set $\mathsf{a}_i^{(\ell)} = \mathsf{b}_i^{(\ell)} = \Delta_{+\infty}$ for $i \notin \mathcal{L}$. The channel densities for codes 1 and 2 are denoted by $\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})}$ and $\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})}$ respectively. All the above densities are L-densities conditioned on the transmission of the all-zero codeword (see Section A). We consider the parallel schedule for each user (as described in [44]) and update the correlation nodes before proceeding to the next iteration. Let us define

$$g(\boldsymbol{x}_{i-w+1}, \cdots, \boldsymbol{x}_{i+w-1}) \triangleq \left(\frac{1}{w} \sum_{j=0}^{w-1} \left(\frac{1}{w} \sum_{k=0}^{w-1} \boldsymbol{x}_{i+j-k} \right)^{\circledast(r-1)} \right)^{\circledast(l-1)},$$

$$\Gamma(\boldsymbol{x}_{i-w+1}, \cdots, \boldsymbol{x}_{i+w-1}) \triangleq \left(\frac{1}{w} \sum_{j=0}^{w-1} \left(\frac{1}{w} \sum_{k=0}^{w-1} \boldsymbol{x}_{i+j-k} \right)^{\circledast(r-1)} \right)^{\circledast l}.$$

The DE equations for the joint spatially-coupled system can be written as

$$\begin{aligned}
\mathsf{a}_i^{(\ell+1)} &= \left[\gamma\zeta\left(\Gamma(\mathsf{b}_{i-w+1}^{(\ell)},\cdots,\mathsf{b}_{i+w-1}^{(\ell)})\right)+(1-\gamma)\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})}\right]\circledast \\
&\qquad g(\mathsf{a}_{i-w+1}^{(\ell)},\cdots,\mathsf{a}_{i+w-1}^{(\ell)}), \\
\mathsf{b}_i^{(\ell+1)} &= \left[\gamma\zeta\left(\Gamma(\mathsf{a}_{i-w+1}^{(\ell)},\cdots,\mathsf{a}_{i+w-1}^{(\ell)})\right)+(1-\gamma)\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})}\right]\circledast \\
&\qquad g(\mathsf{b}_{i-w+1}^{(\ell)},\cdots,\mathsf{b}_{i+w-1}^{(\ell)}),
\end{aligned}$$

for $i \in \mathcal{L}$. For a further discussion of the DE equations for the $(\mathtt{l},\mathtt{r},L,w)$ spatially-coupled ensembles on BMS channels, see [49]. Let $\underline{\mathsf{a}} = (\mathsf{a}_{-L},\cdots,\mathsf{a}_L)$ and $\underline{\mathsf{b}} = (\mathsf{b}_{-L},\cdots,\mathsf{b}_L)$. The fixed points of SC DE are given by $(\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})},\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})},\underline{\mathsf{a}},\underline{\mathsf{b}})$. Define

$$\mathsf{G}(\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})},\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})},\underline{\mathsf{a}},\underline{\mathsf{b}}) = \frac{1}{2L+1}\sum_{i=-L}^{L}\mathsf{G}(\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})},\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]}},\mathsf{a}_i,\mathsf{b}_i).$$

The BP-GEXIT curve is the set of points $(\alpha,\mathsf{G}(\mathsf{a}_{\mathrm{BMSC}(\alpha^{[1]})},\mathsf{a}_{\mathrm{BMSC}(\alpha^{[2]})},\underline{\mathsf{a}},\underline{\mathsf{b}}))$. The resulting curves for the erasure channel with erasure correlated sources are shown in Fig. 26 and those for the AWGN channel with BSC correlated sources are shown in Fig. 27. These curves are very similar to the single user case and demonstrate the phenomenon of threshold saturation at the joint decoder, for symmetric channel conditions. For channel parameters not on the symmetric line, this implies threshold saturation towards the MAP boundary.

Consider the SWE problem and a monotonic curve $\mathcal{C}$ parametrized by ϵ i.e., let $\epsilon_1 = \epsilon_1(\epsilon)$ and $\epsilon_2 = \epsilon_2(\epsilon)$. In this case, the DE recursion can be written as (4.2),

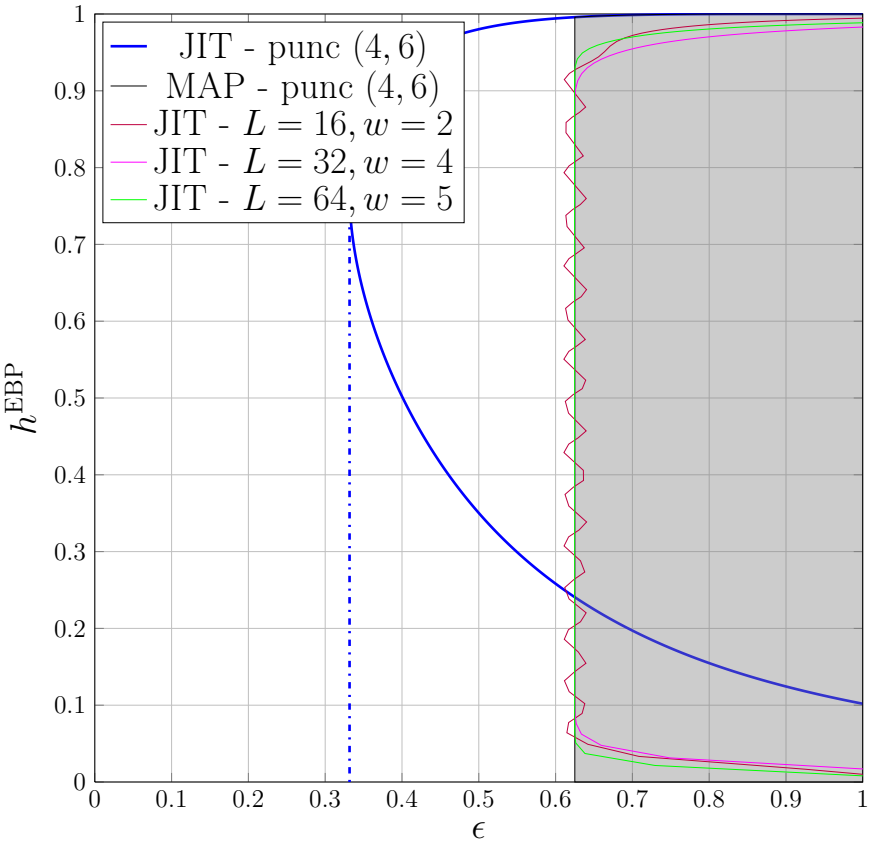

Fig. 26. EBP-EXIT curves of the $(4, 6, L, w)$ and $(4, 6)$ ensembles for transmission over erasure channels with erasure correlated sources.

where

$$\boldsymbol{f}(\boldsymbol{x};\epsilon) \triangleq [\psi(L(x_2);\epsilon_1(\epsilon))\lambda(x_1), \psi(L(x_1);\epsilon_2(\epsilon))\lambda(x_2)],$$

$$\psi(x;\epsilon) = (1-\gamma)\epsilon + \gamma(1-p+px),$$

$$\boldsymbol{g}(\boldsymbol{x}) \triangleq [1-\rho(1-x_1), 1-\rho(1-x_2)].$$

From this, we can compute

$$F(\boldsymbol{x};\epsilon) = \frac{\psi(L(x_1);\epsilon_2(\epsilon))\psi(L(x_2);\epsilon_1(\epsilon))}{\gamma p L'(1)}$$

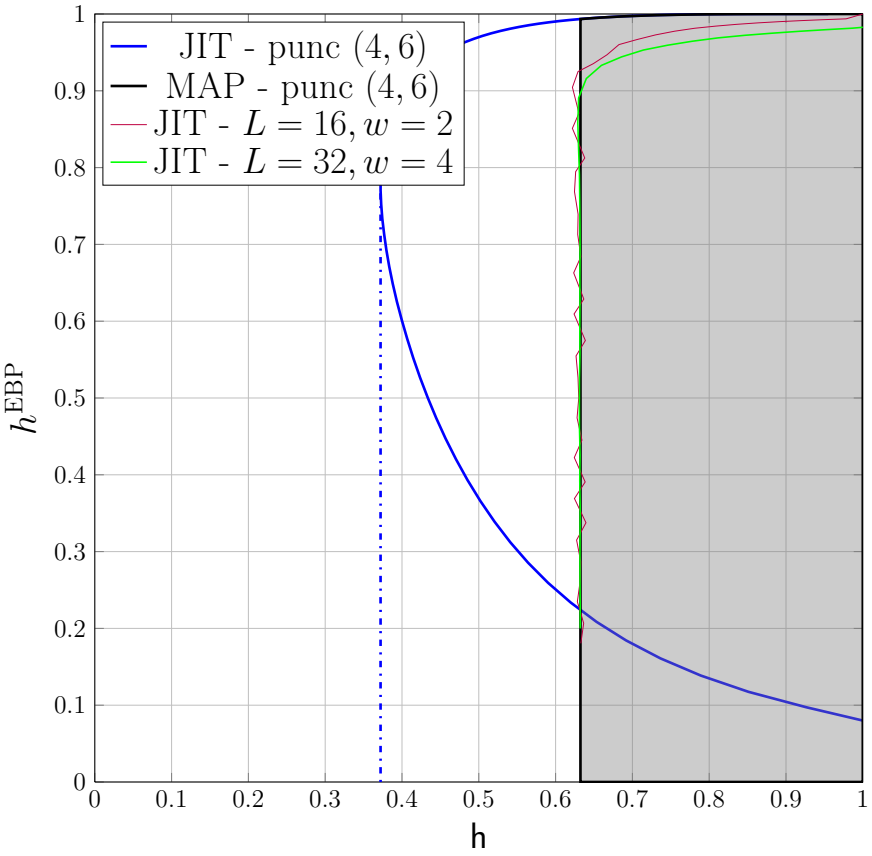

Fig. 27. EBP-GEXIT curves of the $(4,6,L,w)$ and $(4,6)$ ensembles for transmission over AWGN channels which BSC correlation between the sources.

and

$$G(\boldsymbol{x}) = \int_{\mathbf{0}}^{\boldsymbol{x}} \boldsymbol{g}(\boldsymbol{y}) \cdot \mathrm{d}\boldsymbol{y} = (\boldsymbol{x} - R(\mathbf{1}-\boldsymbol{x})/R'(1)) \cdot \mathbf{1}.$$

Let $P(\boldsymbol{x})$ be the trial entropy defined in Section 3 and let $\epsilon(\boldsymbol{x})$ (see Appendix A) be differentiable. Also define $\boldsymbol{\epsilon}(\boldsymbol{x}) = [\epsilon_1(\epsilon(\boldsymbol{x})), \epsilon_2(\epsilon(\boldsymbol{x}))]$. Then, we have

$$U(\boldsymbol{x};\epsilon) = \frac{1-\gamma}{L'(1)} \left((\boldsymbol{\epsilon}(\boldsymbol{x}) - [\epsilon_1(\epsilon(\boldsymbol{x})), \epsilon_2(\epsilon(\boldsymbol{x}))]) \cdot L(\boldsymbol{g}(\boldsymbol{x})) - P(\boldsymbol{x}) \right)$$

Lemma V.1. *Consider the potential threshold ϵ^* defined by (4.4). Let ϵ^{Max} be the*

Maxwell threshold defined by

$$\epsilon^{Max} = \min\left\{\epsilon(\boldsymbol{x}) \,|\, P(\boldsymbol{x}) = 0, \boldsymbol{x} \in \mathcal{X}\right\}. \tag{5.1}$$

Then, $\epsilon^* = \epsilon^{Max}$ for this problem.

Proof. Let $\boldsymbol{x}^{\text{Max}}$ be the $\boldsymbol{x}$-value that achieves the minimum. Then, $U(\boldsymbol{x}^{\text{Max}}; \epsilon^{\text{Max}}) = U(\boldsymbol{x}^{\text{Max}}; \epsilon(\boldsymbol{x}^{\text{Max}})) = -P(\boldsymbol{x}^{\text{Max}})(1-\gamma)/L'(1) = 0$. From Def. IV.7, we know $\epsilon^* \leq \epsilon(\boldsymbol{x}^{\text{Max}})$. Also, it can be shown that $P(\boldsymbol{x}^*) = 0$, and thus, $\epsilon^{\text{Max}} \leq \epsilon^*$. Therefore, we have equality. ■

Corollary V.2. *Applying Theorem IV.7 shows that, if $\epsilon < \epsilon^{Max}$ and $w > K_{\boldsymbol{f},\boldsymbol{g}}/\Delta E(\epsilon)$, then the SC DE recursion must converge to the zero matrix. This shows the universality of spatially-coupled codes for the SWE problem, along with the results in Section III.3.*

B. The Gaussian Multiple-Access Channel

We first describe the $(\mathtt{l}, \mathtt{r}, L)$ ensemble through a protograph. The protograph structure at the joint decoder is shown in Fig. 28 for a LDPC$(3, 6)$ base code. The protograph is generated as follows: Consider the protograph of a $(3, 6)$ regular LDPC code. It has two variable nodes of degree 3 and one check node of degree 6. Connect both the variable nodes to the variable nodes of another protograph via function nodes. The resulting protograph represents the joint decoder when both users are using $(3, 6)$ regular LDPC codes for transmission over the 2-user binary-input Gaussian MAC. Place $2L + 1$ protographs at positions $-L, \cdots, L$. Each of the 3 edges of a variable node at position i is connected to exactly one check node at position $i-1, i, i+1$, for each user.

As noted in Section A, we use the $(\mathtt{l}, \mathtt{r}, L, w)$ ensemble for the remainder of this

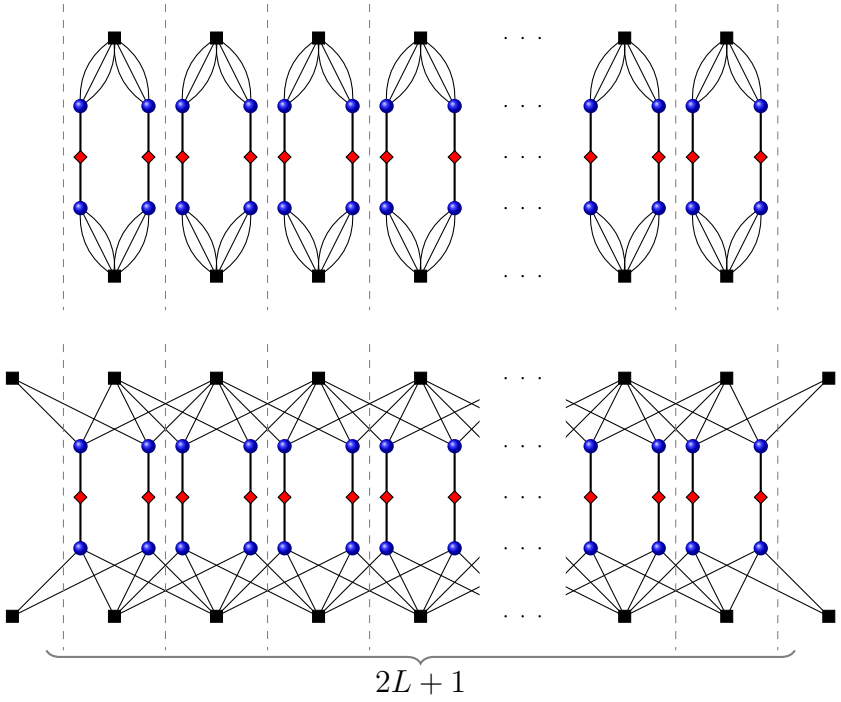

Fig. 28. Protograph of the joint decoder. Shown above are $2L+1$ copies of the protograph of the joint decoder for a $(3, 6)$ regular LDPC code. The bottom graph shows the protograph of the joint decoder for the corresponding spatially coupled code.

work.

1. The $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble

Two single-user graphs, which are generated by the procedure described in Section 2, are taken and the variable nodes (of each graph) at each position are connected by a uniform random permutation of size M via channel nodes. This procedure ensures that all the variable node positions are symmetric and enables us to write down the density evolution (DE) equations in a simple manner, as described in the following section.

2. Density Evolution of the $(\mathtt{l}, \mathtt{r}, L, w)$ Ensemble and GEXIT Curves

Let $\mathsf{a}_i^{(\ell)}$ and $\mathsf{b}_i^{(\ell)}$ denote the average density emitted by the variable node at position i, at iteration ℓ, for codes 1 and 2 respectively. Set $\mathsf{a}_i^{(\ell)} = \mathsf{b}_i^{(\ell)} = \Delta_{+\infty}$ for $i \notin \mathcal{L}$. The channel density is denoted by $\mathsf{a}_{\text{BAWGNMA}}$. All the above densities are L-densities conditioned on the transmission of the all-zero codeword (see Section 1). We consider the parallel schedule for each user (as described in [44]) and update the correlation nodes before proceeding to the next iteration. Let us define

$$g(\boldsymbol{x}_{i-w+1}, \cdots, \boldsymbol{x}_{i+w-1}) \triangleq \left(\frac{1}{w} \sum_{j=0}^{w-1} \left(\frac{1}{w} \sum_{k=0}^{w-1} \boldsymbol{x}_{i+j-k} \right)^{\circledast(r-1)} \right)^{\circledast(l-1)},$$

$$\Gamma(\boldsymbol{x}_{i-w+1}, \cdots, \boldsymbol{x}_{i+w-1}) \triangleq \left(\frac{1}{w} \sum_{j=0}^{w-1} \left(\frac{1}{w} \sum_{k=0}^{w-1} \boldsymbol{x}_{i+j-k} \right)^{\circledast(r-1)} \right)^{\circledast l}.$$

The DE equations for the joint spatially-coupled system can be written as

$$\begin{aligned}
\mathsf{a}_i^{(\ell+1)} &= \zeta_{2\to 1} \left(\Gamma(\mathsf{b}_{i-w+1}^{(\ell)}, \cdots, \mathsf{b}_{i+w-1}^{(\ell)}), \mathsf{a}_{\text{BAWGNMA}} \right) \circledast \\
&\qquad g(\mathsf{a}_{i-w+1}^{(\ell)}, \cdots, \mathsf{a}_{i+w-1}^{(\ell)}), \\
\mathsf{b}_i^{(\ell+1)} &= \zeta_{1\to 2} \left(\Gamma(\mathsf{a}_{i-w+1}^{(\ell)}, \cdots, \mathsf{a}_{i+w-1}^{(\ell)}), \mathsf{a}_{\text{BAWGNMA}} \right) \circledast \\
&\qquad g(\mathsf{b}_{i-w+1}^{(\ell)}, \cdots, \mathsf{b}_{i+w-1}^{(\ell)}),
\end{aligned}$$

for $i \in \mathcal{L}$. For a further discussion of the DE equations for the (l, r, L, w) spatially-coupled ensembles on BMS channels, see [49]. Using the notation $\underline{\mathsf{a}} \triangleq (\mathsf{a}_{-L}, \cdots, \mathsf{a}_L)$, the fixed points of DE are given by $(\mathsf{a}_{\text{BAWGNMA}}, \underline{\mathsf{a}}, \underline{\mathsf{b}})$. Define the GEXIT value of a fixed point $(\mathsf{a}_{\text{BAWGNMA}}, \underline{\mathsf{a}}, \underline{\mathsf{b}})$ by

$$\mathsf{G}(\mathsf{a}_{\text{BAWGNMA}}, \underline{\mathsf{a}}, \underline{\mathsf{b}}) \triangleq \frac{1}{2L+1} \sum_{i=-L}^{L} \mathsf{G}(\mathsf{a}_{\text{BAWGNMA}}, \mathsf{a}_i, \mathsf{b}_i),$$

where G is defined in (3.4). For a fixed θ, the BP-GEXIT curve $\mathsf{g}(\alpha)$ is given by the set of points $(\alpha, \mathsf{G}(\mathsf{a}_{\mathrm{BAWGNMA}(\alpha)}, \underline{\mathsf{a}}, \underline{\mathsf{b}}))$. The resulting curves for the spatially-coupled $(3, 6, 16, 2)$ and $(3, 6, 32, 4)$ ensembles are shown in Fig. 29 for symmetric channel conditions. These curves are very similar to the single user case and demonstrate the phenomenon of threshold saturation at the joint decoder, for symmetric channel conditions. For channel parameters not on the symmetric line, this implies threshold saturation towards the MAP boundary.

C. Summary

The density evolution ACPRs for the two scenarios considered in this work are shown in Figs. 30, 31 and 32. These figures show that spatially coupled ensembles are near universal for these problems.

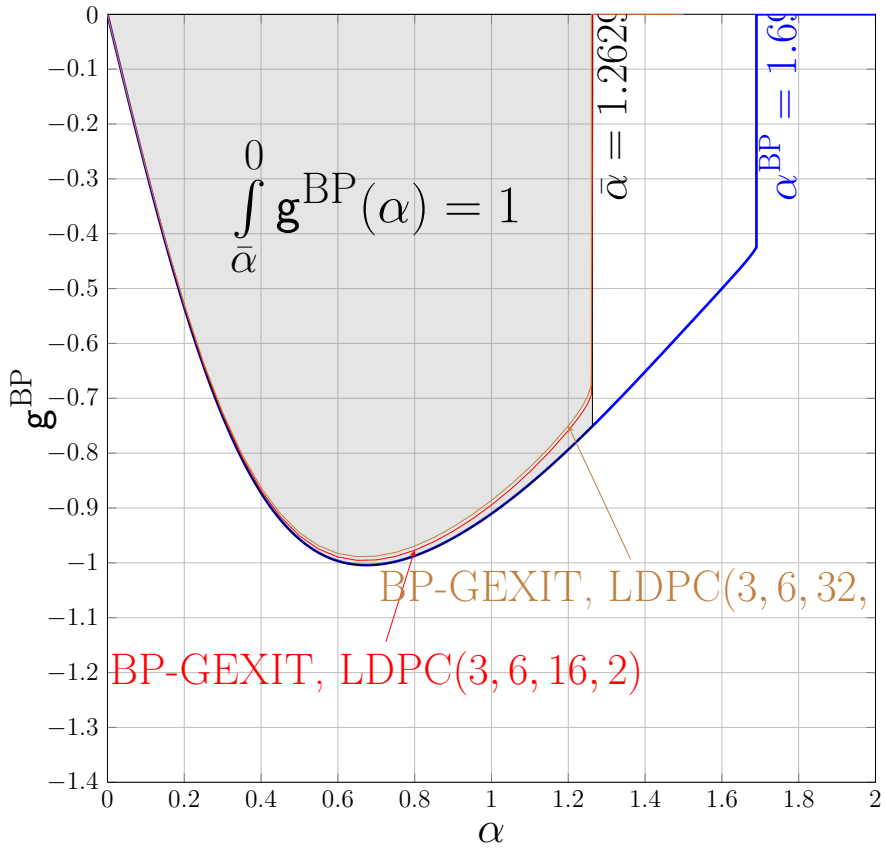

Fig. 29. BP-GEXIT curve and an upper bound on the MAP threshold (computed using the area theorem) for transmission over a 2-user binary-input Gaussian MAC, for $A = 1$, of the $(3, 6)$ regular LDPC ensemble. GEXIT curves in literature are typically parametrized by the channel entropy and the channels get *worse* as the entropy increases. However, the channel gains are a natural parameterization for this problem and the channel gets *better* by increasing the channel gains. So the GEXIT values are negative for this parametrization. Also shown are the BP-GEXIT curves of the $(3, 6, L, w)$ spatially-coupled LDPC ensembles.

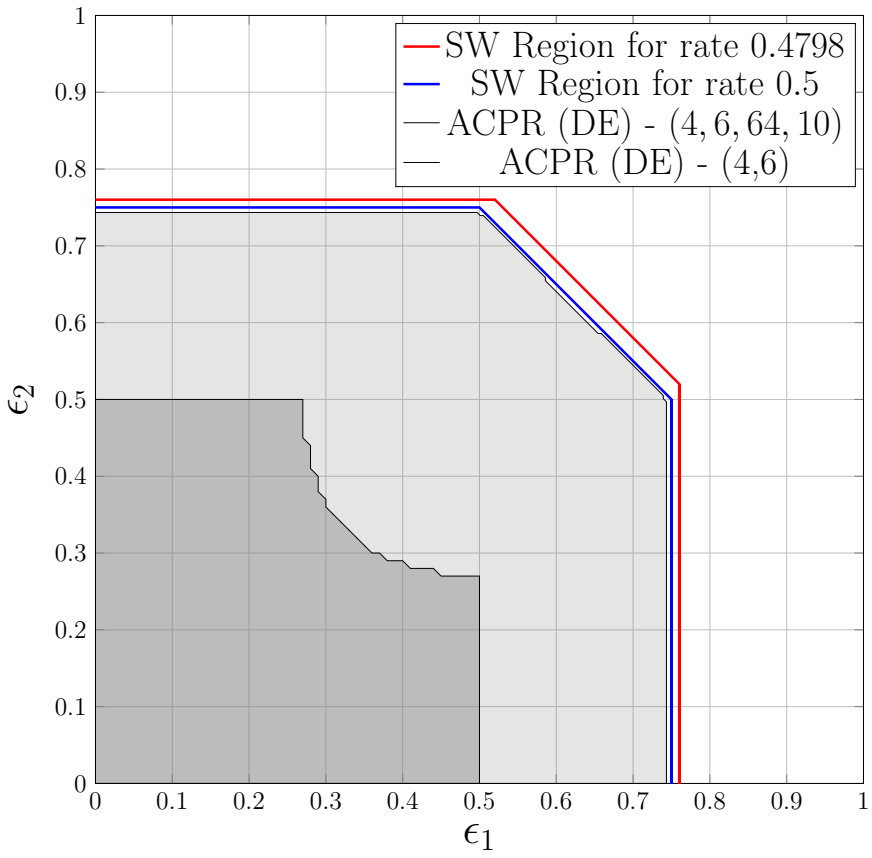

Fig. 30. DE ACPR of the spatially coupled punctured $(4, 6, 64, 10)$ LDPC and the regular punctured LDPC$(4, 6)$ ensembles for transmission over erasure channels with erasure correlated sources.

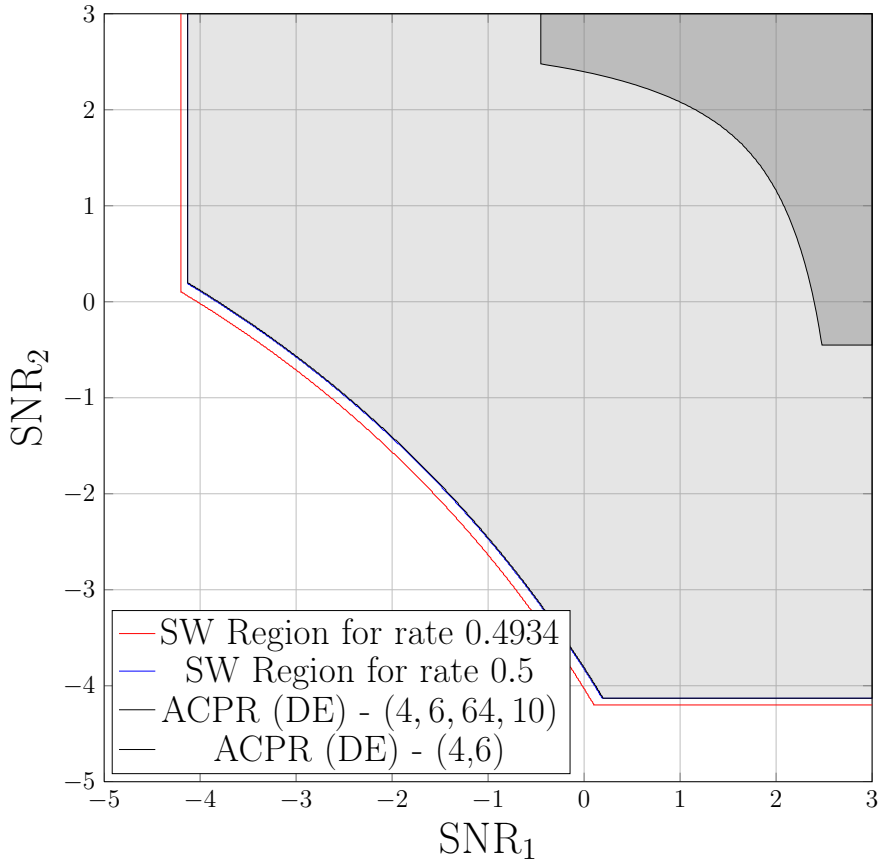

Fig. 31. DE ACPR of the spatially coupled punctured $(4, 6, 64, 10)$ LDPC and the regular punctured LDPC$(4, 6)$ ensembles for transmission over AWGN channels with BSC correlated sources.

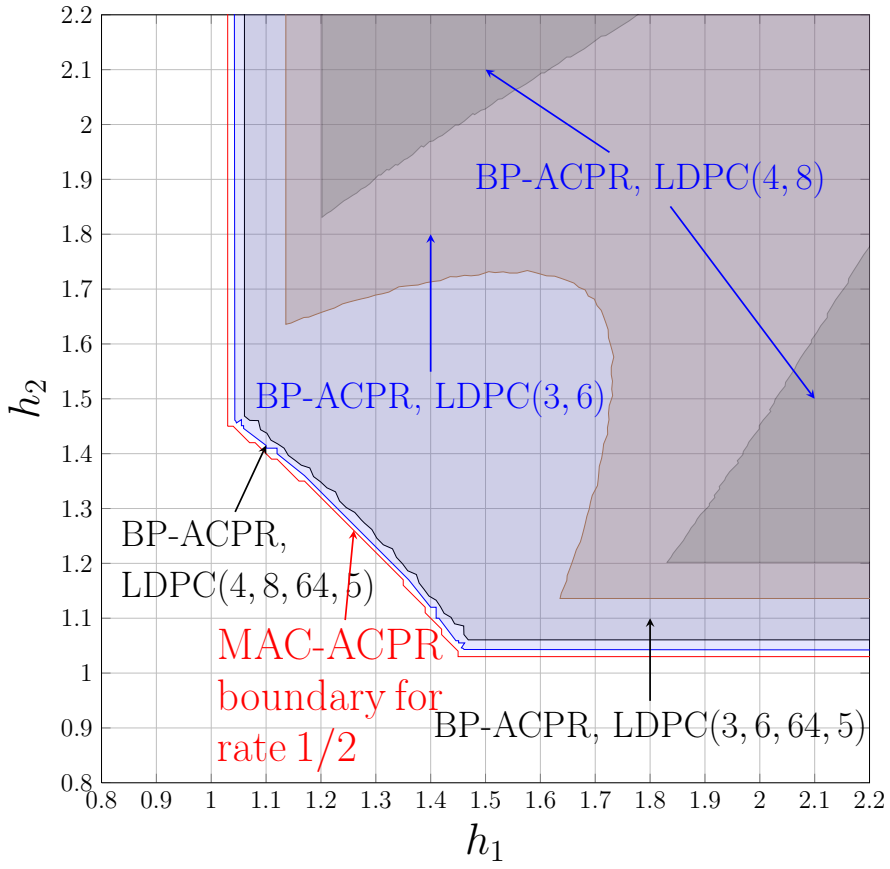

Fig. 32. BP-ACPR of the $(3, 6, 64, 5)$ and $(4, 8, 64, 5)$ spatially-coupled LDPC ensembles for the 2-user binary-input Gaussian MAC. Also shown are the BP-ACPRs for the $(3, 6)$ and $(4, 8)$ regular LDPC ensembles. The BP-ACPR of the $(4, 8, 64, 5)$ spatially-coupled LDPC ensemble is very close to the MAC-ACPR, demonstrating the near-universal performance of spatially-coupled codes.

CHAPTER VI

CONCLUDING REMARKS AND FUTURE WORK

A. Results

The noisy Slepian-Wolf problem and the 2-user Gaussian MAC were considered in this work. The GEXIT functions for these problems are computed and their natural area theorems are derived. By projecting the GEXIT functions along monotone curves, we are able to obtain an upper bound on the MAP decoding threshold. This bound is shown to be tight for the SWE problem for some cases. Based on the observation that regular LDPC codes with large left degrees behave like random codes and the fact that random codes are universal under MAP decoding, we are also able to show that increasing the left degree (keeping the rate constant) will push the MAP boundary towards the boundary of the SW-ACPR for the SWE problem. We also conjecture that increasing the left degree (keeping the rate constant) will push the MAP boundary towards the boundary of the SW-ACPR/MAC-ACPR for more general noise distributions.

We considered spatially-coupled codes for the noisy Slepian-Wolf problem and the Gaussian MAC and observed that spatial coupling boosts the BP threshold of the joint decoder to the MAP threshold of the underlying ensemble. The density evolution ACPRs for the two scenarios considered in this work are shown in Figs. 30, 31 and 32. These figures show that spatially-coupled ensembles are near universal for these problems. We are able to show an analytic proof of this result for the SWE problem. The analytic proof of this result remains an open problem for other general models. Such a proof would essentially show that it is possible to achieve universality for the noisy Slepian-Wolf problem and the MAC channel under iterative decoding.

B. Future Work

This work can be extended in a variety of ways. For example, it is straightforward to dispense with AWGN and compute the ACPRs of any suitably parameterized 2-user binary-input MAC. One can also generalize these results to m-user MACs, larger input alphabets, and multiple-input multiple-output (MIMO) systems. In these cases, the increase in computational complexity makes discretized DE infeasible and Monte Carlo methods must be used to evaluate the DE and GEXIT functions. We conjecture that threshold saturation will continue to occur for all these extensions and that spatially-coupled codes will achieve near-universal performance.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell Syst. Techn. J.*, vol. 27, pp. 379–423, 623–656, July / Oct. 1948.

[2] T. J. Richardson and R. L. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[4] C. Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[5] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 155–170, 2006.

[6] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[7] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. 20, no. 1, pp. 2–10, 1974.

[8] J. Chen, D. He, and A. Jagmohan, "Slepian-Wolf code design via source-channel correspondence," in *Proc. IEEE Int. Symp. Inform. Theory*, 2006, pp. 2433–2437.

[9] P. Delsarte and P. Piret, "Algebraic constructions of shannon codes for regular channels," *IEEE Trans. Inform. Theory*, vol. 28, no. 4, pp. 593–599, July 1982.

[10] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (discus): Design and construction," *Information Theory, IEEE Transactions on*, vol. 49, no. 3, pp. 626–643, 2003.

[11] J. Garcia-Frias, "Joint source-channel decoding of correlated sources over noisy channels," in *Data Comp. Conf.*, Snowbird, UT, 2001, pp. 283–292.

[12] W. Zhong and J. Garcia-Frias, "LDGM codes for channel coding and joint source-channel coding of correlated sources," *EURASIP J. on Appl. Signal Process.*, pp. 942–953, 2005.

[13] R. Hu, R. Viswanathan, and J. Li, "A new coding scheme for the noisy-channel slepian-wolf problem: Separate design and joint decoding," in *Proc. IEEE Global Telecom. Conf.* IEEE, 2004, vol. 1, pp. 51–55.

[14] P. D. Alexander, A. J. Grant, and M. C. Reed, "Iterative detection in code-division multiple-access with error control coding," *European Transactions on Telecommunications*, vol. 9, no. 5, pp. 419–425, 1998.

[15] L. Brunel and J. Boutros, "Code division multiple access based on independent codes and turbo decoding," *Annals of Telecommunications*, vol. 54, no. 7, pp. 401–410, 1999.

[16] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1772–1793, 2002.

[17] C. Schlegel, Zhenning Shi, and M. Burnashev, "Optimal power/rate allocation and code selection for iterative joint detection of coded random CDMA," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 4286–4294, Sept.

2006.

[18] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: near-single-user performance," *Communications, IEEE Transactions on*, vol. 46, no. 12, pp. 1693–1699, Dec 1998.

[19] M. Moher, "An iterative multiuser decoder for near-capacity communications," *Communications, IEEE Transactions on*, vol. 46, no. 7, pp. 870–880, Jul 1998.

[20] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *Communications, IEEE Transactions on*, vol. 47, no. 7, pp. 1046–1061, Jul 1999.

[21] P. D. Alexander, M. C. Reed, J. A. Asenstorfer, and C. B. Schlegel, "Iterative multiuser interference reduction: turbo CDMA," *Communications, IEEE Transactions on*, vol. 47, no. 7, pp. 1008–1014, Jul 1999.

[22] Z. Shi and C. Schlegel, "Joint iterative decoding of serially concatenated error control coded CDMA," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1646–1653, Aug 2001.

[23] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[24] S. Verdú and S. Shamai, "Spectral efficiency of cdma with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, mar 1999.

[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications. Wiley, 1991.

[26] R. Palanki, A. Khandekar, and R. J. McEliece, "Graph-based codes for synchronous multiple access channels," in *Proc. 39th Annual Allerton Conf. on Commun., Control, and Comp.*, Oct. 2001, vol. 39, pp. 1263–1271.

[27] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 364–375, 1996.

[28] A. Amraoui, S. Dusad, and R. Urbanke, "Achieving general points in the 2-user Gaussian MAC without time-sharing or rate-splitting by means of iterative coding," in *Proc. IEEE Int. Symp. Inform. Theory*, 2002, p. 334.

[29] A. Roumy and D. Declercq, "Characterization and optimization of LDPC codes for the 2-user Gaussian multiple access channel," *EURASIP J. on Wireless Commun. and Networking*, vol. Article ID, no. 74890, 2007.

[30] R. Liu, P. Spasojevic, and E. Soljanin, "Reliable channel regions for good binary codes transmitted over parallel channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1405–1424, 2006.

[31] R. G. Gallager, "Low-density parity-check codes," Ph.D. dissertation, M.I.T., Cambridge, MA, USA, 1960.

[32] R. Michael Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, no. 5, pp. 533–547, Sept. 1981.

[33] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, March 1999.

[34] M. G. Luby, M. Mitzenmacher, and M. Amin Shokrollahi, "Practical loss-resilient codes," in *Proc. 29th Annu. ACM Symp. Theory of Computing*, 1997, pp. 150–159.

[35] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 569–584, Feb. 2001.

[36] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[37] S. ten-Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEÜ Int. J. Electron. Commun.*, vol. 54, no. 6, pp. 389–398, Nov. 2000.

[38] C. Méasson, A. Montanari, T. J. Richardson, and R. L. Urbanke, "Life above threshold: From list decoding to area theorem and MSE," *Arxiv preprint cs.IT/0410028*, 2004.

[39] C. Méasson, A. Montanari, T. J. Richardson, and R. Urbanke, "The generalized area theorem and some of its consequences," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4793–4821, Nov. 2009.

[40] M. Luby and A. Shokrollahi, "Slepian-Wolf type problems on the erasure channel," in *Proc. IEEE Int. Symp. Inform. Theory*, St. Petersburg, Russia, July 2011, pp. 2771–2775.

[41] S-Y. Chung, G. D. Forney, Jr., T. J. Richardson, and R. L. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Letters*, vol. 5, no. 2, pp. 58–60, Feb. 2001.

[42] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund, "Performance analysis and design of two edge type LDPC codes for the BEC wiretap

channel," [Online]. Available: http://arxiv.org/abs/1009.4610, Sept. 2010.

[43] C. Méasson, "Conservation laws for coding," Ph.D. dissertation, Swiss Federal Institute of Technology, Lausanne, 2006.

[44] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 803–834, 2011.

[45] J. Felstrom and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2181–2191, 1999.

[46] M. Lentmaier, A. Sridharan, K. S. Zigangirov, and D. J. Costello, "Terminated LDPC convolutional codes with thresholds close to capacity," in *Proc. IEEE Int. Symp. Inform. Theory*, Adelaide, Australia, 2005, pp. 1372–1376.

[47] M. Lentmaier and G. P. Fettweis, "On the thresholds of generalized LDPC convolutional codes based on protographs," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, TX, 2010, pp. 709–713.

[48] M. Lentmaier, A. Sridharan, D. J. Costello, and K. S. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inform. Theory*, vol. 56, no. 10, pp. 5274–5289, Oct. 2010.

[49] S. Kudekar, C. Méasson, T. Richardson, and R. Urbanke, "Threshold saturation on BMS channels via spatial coupling," in *Proc. Int. Symp. on Turbo Codes & Iterative Inform. Proc.*, Sept. 2010, pp. 309–313.

[50] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," Arxiv preprint

arXiv:1201.2999, 2012.

[51] C. Méasson, A. Montanari, and R. L. Urbanke, "Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding," *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5277–5307, Dec. 2008.

[52] A. Yedla, Y. Y. Jian, P. S. Nguyen, and H. D. Pfister, "A simple proof of threshold saturation for coupled scalar recursions," Arxiv preprint, 2012.

[53] A. Sridharan, M. Lentmaier, D. J. Costello, and K. S. Zigangirov, "Convergence analysis of a class of LDPC convolutional codes for the erasure channel," in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, 2004, pp. 953–962.

[54] M. Luby, "LT codes," in *Proc. of the 43rd Symp. on Foundations of Comp. Sci.*, Washington, D.C., June 2002, p. 271.

[55] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Springer Verlag, 2005.

[56] A. Shokrollahi and R. Storn, "Design of efficient erasure codes with differential evolution," in *Proc. IEEE Int. Symp. Inform. Theory*, 2000.

APPENDIX A

PROOFS

Proof of Lemma III.1

Consider the term

$$\begin{aligned} H(\boldsymbol{X}|\boldsymbol{Y}) &= H(X_i|\boldsymbol{Y}) + H(\boldsymbol{X}_{\sim i}|X_i, \boldsymbol{Y}) \\ &= H\left(X_i^{[1]}|Y_i, \Phi_i\right) + H\left(X_i^{[2]}|X_i^{[1]}, \boldsymbol{Y}_{\sim i}\right) + H(\boldsymbol{X}_{\sim i}|X_i, \boldsymbol{Y}_{\sim i})\,. \end{aligned}$$

Note that only the first term of the decomposition depends on the first channel at position i. A similar decomposition can be done while taking the derivative with respect to the second channel at position i. So, we get

$$\begin{aligned} \mathsf{g}_i(\alpha_i^{[1]}, \alpha_i^{[2]}) &= \frac{\mathrm{d}}{\mathrm{d}\alpha_i^{[1]}} H\left(X_i^{[1]}|Y_i^{[1]}, \Phi_i^{[1]}\right) + \frac{\mathrm{d}}{\mathrm{d}\alpha_i^{[2]}} H\left(X_i^{[2]}|Y_i^{[2]}, \Phi_i^{[2]}\right) \\ &= \frac{\partial \alpha_i^{[1]}}{\partial \alpha} \frac{\partial}{\partial \alpha} H\left(X_i^{[1]}|Y_i^{[1]}, \Phi_i^{[1]}\right) + \frac{\partial \alpha_i^{[2]}}{\partial \alpha} \frac{\partial}{\partial \alpha} H\left(X_i^{[2]}|Y_i^{[2]}, \Phi_i^{[2]}\right). \end{aligned}$$

Following the standard procedure for single user channels, we have

$$H\left(X_i^{[1]}|Y_i^{[1]}, \Phi_i^{[1]}\right) = \int_{u,v} \mathsf{a}_i(u) \mathsf{a}_{\mathrm{BMSC}(\alpha_i^{[1]})}(v) \log_2(1 + \mathrm{e}^{-u-v}) \mathrm{d}v \mathrm{d}u.$$

So,

$$\mathsf{g}_i(\alpha_i^{[1]}, \alpha_i^{[2]}) = \frac{\partial \alpha_i^{[1]}}{\partial \alpha} \int_u \mathsf{a}_i(u) \kappa(\mathsf{a}_{\mathrm{BMSC}(\alpha_i^{[1]})}, u) \mathrm{d}u + \frac{\partial \alpha_i^{[2]}}{\partial \alpha} \int_u \mathsf{b}_i(u) \kappa(\mathsf{b}_{\mathrm{BMSC}(\alpha_i^{[2]})}, u) \mathrm{d}u.$$

Proof of Lemma III.3

From Lemma III.2, the GEXIT value can be simplified to

$$\mathsf{G}^{\mathrm{BP}}(a,b) = L(g(a))\frac{\partial \alpha^{[1]}}{\partial \alpha} + L(g(b))\frac{\partial \alpha^{[2]}}{\partial \alpha}. \tag{A.1}$$

Since $\alpha^{[1]}$ and $\alpha^{[2]}$ are functions of α, from (2.8) a and b must be functions of some common parameter x and we write $\alpha(x) = \alpha(a,b)$, $a(x)$ and $b(x)$. For example when $a, b \neq 0$, we have $a(x) = x$ and $b(x) = \psi^{-1}(x/\lambda(x))$, where

$$\psi^{-1}(x) = g^{-1}\left(L^{-1}\left(\zeta^{-1}\left(\frac{1}{\gamma}\left(x - (1-\gamma)\alpha^{[1]}(\alpha)\right)\right)\right)\right).$$

All the above functions are well defined for $x \neq 0$. Noting that $\alpha(x)$, $a(x)$ and $b(x)$ are differentiable, we define the trial entropy along the curve $\mathcal{C}$ by

$$P(x) = \int_0^x \mathsf{g}^{\mathrm{BP}}(t)\mathrm{d}\alpha(t).$$

We do not require an explicit characterization of the functions $a(x)$, $b(x)$ and $\alpha(x)$. We first note that

$$\mathrm{d}\alpha^{[1]}(x) = \frac{\partial \alpha^{[1]}}{\partial \alpha}\mathrm{d}\alpha(x),\ \mathrm{d}\alpha^{[2]}(x) = \frac{\partial \alpha^{[2]}}{\partial \alpha}\mathrm{d}\alpha(x),\ \text{and}$$
$$\frac{\mathrm{d}\zeta}{\mathrm{d}x}(x) = p.$$

So, $P(x) = \int_0^x L(g(a(x)))\mathrm{d}\alpha^{[1]}(x) + \int_0^x L(g(b(x)))\mathrm{d}\alpha^{[2]}(x)$. From (2.8), we have

$$\alpha^{[1]}(x) = \frac{1}{1-\gamma}\left[\frac{a(x)}{\lambda(g(a(x)))} - \gamma\zeta(L(g(b(x))))\right] \text{ and}$$
$$\alpha^{[2]}(x) = \frac{1}{1-\gamma}\left[\frac{b(x)}{\lambda(g(b(x)))} - \gamma\zeta(L(g(a(x))))\right].$$

After integration by parts and some algebra, this can be simplified to

$$\begin{aligned}P(x) &= -\frac{p\gamma}{1-\gamma}L(g(a(x)))L(g(b(x))) \\ &+ \frac{1}{1-\gamma}\left(L(g(a(x)))\frac{a(x)}{\lambda(g(a(x)))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-a(x)) - a(x)R'(1-a(x))\right]\right) \\ &+ \frac{1}{1-\gamma}\left(L(g(b(x)))\frac{b(x)}{\lambda(g(b(x)))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-b(x)) - b(x)R'(1-b(x))\right]\right).\end{aligned}$$

Consider the case when $a = 0, b \neq 0$. We then have the parametrization $b(x) = x$, $\alpha(x)$ such that

$$\alpha^{[2]}(x) = \frac{1}{1-\gamma}\left[\frac{b(x)}{\lambda(g(b(x)))}\right].$$

The trial entropy in this case is given by

$$\begin{aligned}P(x) &= \int_0^x L(g(b(x)))\mathrm{d}\alpha^{[2]}(x) \\ &= \frac{1}{1-\gamma}\left(L(g(b(x)))\frac{b(x)}{\lambda(g(b(x)))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-b(x)) - b(x)R'(1-b(x))\right]\right).\end{aligned}$$

Similarly, for the case when $a \neq 0, b = 0$, we have

$$\begin{aligned}P(x) &= \int_0^x L(g(a(x)))\mathrm{d}\alpha^{[1]}(x) \\ &= \frac{1}{1-\gamma}\left(L(g(a(x)))\frac{a(x)}{\lambda(g(a(x)))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-a(x)) - a(x)R'(1-a(x))\right]\right).\end{aligned}$$

Putting this together, the trial entropy at a fixed point (a, b) is given by

$$\begin{aligned}P(a,b) &= -\frac{p\gamma}{1-\gamma}L(g(a))L(g(b)) \\ &+ \frac{1}{1-\gamma}\left(L(g(a))\frac{a}{\lambda(g(a))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-a) - aR'(1-a)\right]\right) \\ &+ \frac{1}{1-\gamma}\left(L(g(b))\frac{b}{\lambda(g(b))} - \frac{L'(1)}{R'(1)}\left[1 - R(1-b) - bR'(1-b)\right]\right).\end{aligned}$$

Proof of Lemma III.4

The expected check node degree distribution $\tilde{R}_\epsilon(z)$ can be derived similarly to the BEC case.

For the degree distribution of bit nodes, the fraction of unpunctured bits is $1-\gamma$. For these bit nodes to remain in the residual graph, the messages from the channel as well as the from the corresponding check nodes must be erasures. This happen with probability $\alpha^{[1]} \cdot L(y(a))$ and $\alpha^{[2]} \cdot L(y(b))$. The fraction of punctured bits is γ. Among these, for bits which are not connected by correlation nodes (w.p. $1-p$) to remain in the residual graph, the messages from corresponding check nodes must be erasures. This happens with probability $(1-p)L(y(a))$ and $(1-p)L(y(b))$. Meanwhile, every two bits which are connected by a correlation node (this happens with probability p) are merged into a larger bit node with twice the degree. For these larger bit nodes to remain in the residual graph, the messages from check nodes in both sources must be erasures. This happens with probability $\frac{p}{2}L(y(a))L(y(b))$. The results follows immediately.

Proof of Lemma III.7

Suppose that each bit is transmitted through a channel with parameter α_i and consider the term

$$\begin{aligned} H(\boldsymbol{X}|\boldsymbol{Y}) &= H(X_i|\boldsymbol{Y}) + H(\boldsymbol{X}_{\sim i}|X_i, \boldsymbol{Y}) \\ &= H(X_i|Y_i, \Phi_i) + H(\boldsymbol{X}_{\sim i}|X_i, \boldsymbol{Y}_{\sim i})\,. \end{aligned}$$

Note that the second term of the decomposition does not depend on the channel at position i. So, we get

$$\mathsf{g}_i(\alpha_i) = \frac{\mathrm{d}}{\mathrm{d}\alpha_i} H(X_i|Y_i, \Phi_i)\,.$$

We have,

$$\begin{aligned} H(X_i|Y_i,\Phi_i) &= -\iint\limits_{y,\phi} \sum_x p(x,y,\phi) \log_2 \frac{p(x,y,\phi)}{\sum\limits_{x'} p(x',y,\phi)} \mathrm{d}y \mathrm{d}\phi \\ &= \sum_x p(x) \int\limits_{\phi} p(\phi|x) \left(\int\limits_y p(y|x) \log_2 \frac{\sum\limits_{x'} p(x'|\phi) p(y|x')}{p(x|\phi)p(y|x)} \mathrm{d}y \right) \mathrm{d}\phi, \end{aligned}$$

which follows from the fact that

$$p_{X_i,Y_i,\Phi_i}(x,y,\phi) = p(y|x)p(\phi|x)p(x),$$

since $Y_i \to X_i \to \Phi_i$. Taking the derivative and noting that $p(x_i|\phi_i) = p(x_i|\mathbf{y}_{\sim i})$, we obtain[1]

$$\begin{aligned} \mathsf{g}_i(\alpha_i) &= \sum_x p(x) \int\limits_{\phi} p(\phi|x) \Bigg(\int\limits_y \frac{\partial}{\partial \alpha} p(y|x) \\ &\qquad \log_2 \frac{\sum\limits_{x'} p(x'|\phi) p(y|x')}{p(x|\phi)p(y|x)} \mathrm{d}y \Bigg) \mathrm{d}\phi \\ &= \sum_x p(x) \int\limits_{\mathbf{u}} \mathsf{a}_{x,i}(\mathbf{u}) \kappa_x(\mathbf{u}) \mathrm{d}\mathbf{u}, \end{aligned}$$

and the result follows by setting $\alpha_i = \alpha$.

---

[1]The terms obtained by differentiating with respect to the channel inside the log vanish.

Proof of Lemma III.8

Let $\Phi^{\mathrm{BP}} = \mathbf{u}$. Then,

$$\begin{aligned}\mathbf{u} &= \phi(\mathbf{y}_{\sim i}) = \{p(x_i|\mathbf{y}_{\sim i}), x_i \in \mathcal{X}\} \\ &= \{p(\pi_1(x_i)|\mathbf{y}_{\sim i}) \cdot p(\pi_2(x_i)|\mathbf{y}_{\sim i}), x_i \in \mathcal{X}\}.\end{aligned}$$

If we define

$$u \triangleq \log \frac{p(X_i^{[1]} = +1|\mathbf{y}_{\sim i})}{p(X_i^{[1]} = -1|\mathbf{y}_{\sim i})}, v \triangleq \log \frac{p(X_i^{[2]} = +1|\mathbf{y}_{\sim i})}{p(X_i^{[2]} = -1|\mathbf{y}_{\sim i})},$$

then

$$\begin{aligned}\mathbf{u} &= \left(\frac{\boldsymbol{\epsilon}^u}{1+\boldsymbol{\epsilon}^u}\frac{\boldsymbol{\epsilon}^v}{1+\boldsymbol{\epsilon}^v}, \frac{\boldsymbol{\epsilon}^u}{1+\boldsymbol{\epsilon}^u}\frac{1}{1+\boldsymbol{\epsilon}^v}, \frac{1}{1+\boldsymbol{\epsilon}^u}\frac{\boldsymbol{\epsilon}^v}{1+\boldsymbol{\epsilon}^v}, \frac{1}{1+\boldsymbol{\epsilon}^u}\frac{1}{1+\boldsymbol{\epsilon}^v}\right) \\ &\triangleq f(u,v). \end{aligned} \tag{A.2}$$

Let $\mathsf{a}(u)$ denote the density of U conditioned on $X_i^{[1]} = +1$ and $\mathsf{b}(v)$ be the density of V conditioned on $X_i^{[2]} = +1$. Then, $\mathsf{a}(-u)$ is the density of U conditioned on $X_i^{[1]} = -1$ and $\mathsf{b}(-v)$ is the density of V conditioned on $X_i^{[2]} = -1$. In the limit $n \to \infty$ and taking expectation these densities are given by the fixed point $(\mathsf{a}_{\mathrm{BAWGNMA}(\alpha)}, \mathsf{a}, \mathsf{b})$. Let $\mathsf{F}_x[\mathsf{a},\mathsf{b}](u,v)$ be the density of Φ_i^{BP} conditioned on $(X_i^{[1]} = \pi_1(x), X_i^{[2]} = \pi_2(x))$. Then,

$$\mathsf{F}_x[\mathsf{a},\mathsf{b}](u,v) = \mathsf{a}\left(\pi_1(x)u\right)\mathsf{b}\left(\pi_2(x)v\right). \tag{A.3}$$

For example $\mathsf{F}_0[\mathsf{a},\mathsf{b}](u,v) = \mathsf{a}(u)\mathsf{b}(v), \mathsf{F}_1[\mathsf{a},\mathsf{b}](u,v) = \mathsf{a}(u)\mathsf{b}(-v)$ and so on. The result follows by the definition of the GEXIT curve. The kernels $\kappa_x(u,v)$ are defined in the sense of (A.2).

APPENDIX B

LDGM CODES FOR THE ESW PROBLEM*

In this appendix, we consider the design of LDGM codes for the SWE problem. Assume that the sequences $\mathbf{U}_1$ and $\mathbf{U}_2$ are encoded using LDGM codes with a degree distribution pair (λ, ρ). Since the encoded variable nodes are are attached to the check nodes randomly, the degree of each variable node is a Poisson random variable whose mean is given by the average number of edges attached to each check node. This mean is given by $\mathsf{m} = R'(1)$, where $R'(1)$ is the average check degree. Therefore, the resulting degree distribution is $L(x) = \mathrm{e}^{\mathsf{m}(x-1)}$. Throughout this section, we consider the erasure correlation model described in Section a.

The Tanner graph for the code is shown in Fig. 33. Code 1 corresponds to the bottom half of the graph, code 2 corresponds to the top half and both the codes are connected by correlation nodes at the source variable nodes. One can verify that the computation graph for decoding a particular bit is asymptotically tree-like, for a fixed number of iterations as the blocklength tends to infinity. This enables the use of density evolution to compute the performance of the joint iterative decoder.

Let x_ℓ and y_ℓ denote the average erasure probability of the variable nodes at iteration ℓ for users 1 and 2 respectively. The density evolution equations in terms of

---

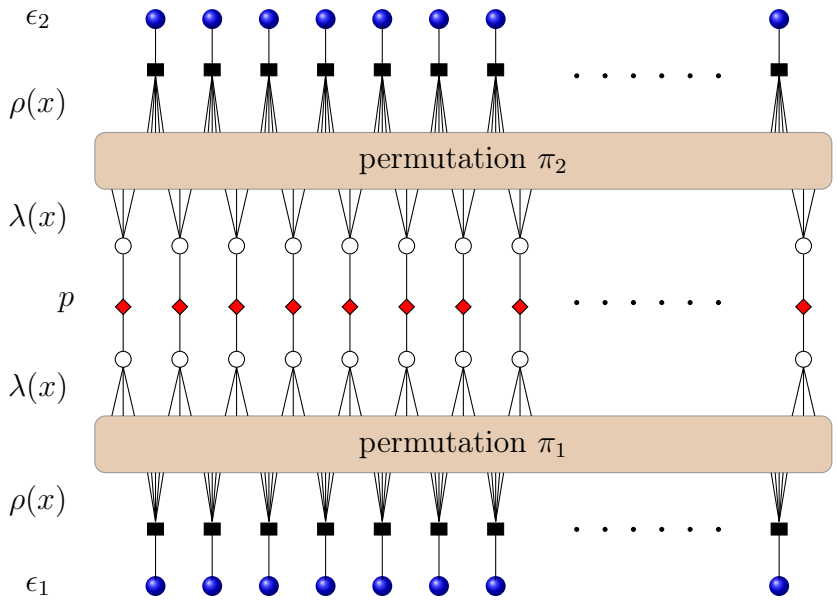

Fig. 33. Tanner Graph of an LDGM (LT) Code with erasure correlation between the sources

the variable-node to check-node messages can be written as

$$x_{\ell+1} = [(1-p) + pL\left(\varrho(\epsilon_2, y_\ell)\right)]\,\lambda\left(\varrho(\epsilon_1, x_\ell)\right)$$
$$y_{\ell+1} = [(1-p) + pL\left(\varrho(\epsilon_1, x_\ell)\right)]\,\lambda\left(\varrho(\epsilon_2, y_\ell)\right),$$

where $\varrho(\epsilon, x) = 1 - (1-\epsilon)\rho(1-x)$. Notice that, for LT codes, the variable-node degree distribution from the edge perspective is given by $\lambda^{(i)}(x) = L^{(i)}(x)$ because $\lambda(x) \triangleq L'(x)/L'(1) = L(x)$, when $L(x)$ is Poisson. With this simplification, the density evolution for symmetric channel conditions ($\epsilon_1 = \epsilon_2 = \epsilon$) can be written as

$$x_{\ell+1} = \big[(1-p) + p\lambda\big(1-(1-\epsilon)\rho(1-x_\ell)\big)\big]\,\lambda\big(1-(1-\epsilon)\rho(1-x_\ell)\big). \tag{B.1}$$

This recursion can be solved analytically, resulting in the unique non-negative $\rho(x)$ which satisfies

$$x = \big[(1-p) + p\lambda\big(1-(1-\epsilon)\rho(1-x)\big)\big]\lambda\big(1-(1-\epsilon)\rho(1-x)\big).$$

The solution is given by

$$\begin{aligned}\rho(x) &= \frac{-1}{\alpha(1-\epsilon)} \cdot \log\left(\frac{\sqrt{(1-p)^2+4p(1-x)}-(1-p)}{2p}\right) \\ &= \frac{1}{\alpha(1-\epsilon)}\sum_{i=1}^{\infty}\frac{\sum_{k=0}^{i-1}\binom{2i-1}{k}p^k}{i(1+p)^{2i-1}}x^i,\end{aligned}$$

which is not a valid degree distribution because it has infinite mean. To overcome this, we define a truncated version of the check degree distribution via

$$\begin{aligned}\rho^N(x) &= \frac{\mu + \sum_{i=1}^{N}\frac{\sum_{k=0}^{i-1}\binom{2i-1}{k}p^k}{i(1+p)^{2i-1}}x^i + x^N}{\mu + G_N(p) + 1} \\ G_N(p) &= \sum_{i=1}^{N}\frac{\sum_{k=0}^{i-1}\binom{2i-1}{k}p^k}{i(1+p)^{2i-1}},\end{aligned} \tag{B.2}$$

for some $\mu > 0$ and $N \in \mathbb{N}$. This is a well defined degree distribution as all the coefficients are non-negative and $\rho^N(1) = 1$. The parameter μ increases the number of degree one generator nodes and is introduced in order to overcome the stability problem at the beginning of the decoding process [54].

Theorem VI.1. *Consider transmission over erasure channels with parameters $\epsilon_1 = \epsilon_2 = \epsilon$. For $N \in \mathbb{N}$ and $\mu > 0$, define*

$$G_N(p) = \sum_{i=1}^{N}\frac{\sum_{k=0}^{i-1}\binom{2i-1}{k}p^k}{i(1+p)^{2i-1}}, \text{ and } \mathsf{m} = \frac{\mu + G_N(p) + 1}{1-\epsilon}.$$

Then, in the limit of infinite blocklengths, the ensemble $LDGM\big(n, \lambda(x), \rho^N(x)\big)$, where

$\lambda(x) = e^{\mathsf{m}(x-1)}$ *and*

$$\rho^N(x) = \frac{\mu + \sum_{i=1}^{N} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i + x^N}{\mu + G_N(p) + 1}, \tag{B.3}$$

enables transmission at a rate $R = \frac{(1-\epsilon)(1-e^{-\mathsf{m}})}{\mu+1-p/2}$*, with a bit error probability not exceeding* $1/N$*.*

Proof. We will use the following Lemma to show that the density evolution equations converge to zero at the extremal symmetric point.

Lemma VI.2.

$$\rho^N(x) > \frac{\mu + \rho(x)}{\mu + G_N(p) + 1}, \ \textit{for } 0 \le x < 1 - \frac{1}{N}.$$

Proof. For $0 \le x < 1 - \frac{1}{N}$, we have

$$\begin{aligned}
\rho^N(x) &= \frac{\mu + \sum_{i=1}^{N} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i + x^N}{\mu + G_N(p) + 1} \\
&= \frac{\mu + \rho(x) + x^N}{\mu + G_N(p) + 1} - \frac{\sum_{i=N+1}^{\infty} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i}{\mu + G_N(p) + 1} \\
&> \frac{\mu + \rho(x)}{\mu + G_N(p) + 1}.
\end{aligned} \tag{B.4}$$

(B.4) follows from the fact that

$$\sum_{i=N+1}^{\infty} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i < \sum_{i=N+1}^{\infty} \frac{x^i}{i} < \frac{1}{N+1} \sum_{i=N+1}^{\infty} x^i = \frac{1}{N+1} \cdot \frac{x^{N+1}}{1-x} < x^N.$$

The last step follows from explicit calculations, taking into account that $0 \le x < 1 - \frac{1}{N}$. ■

From (B.1), the convergence criteria for the density evolution equation is given

by

$$x > \left[(1-p) + p\bar{\lambda}^N(\epsilon, x)\right] \bar{\lambda}^N(\epsilon, x),$$

where $\bar{\lambda}^N(\epsilon, x) = \lambda\left(1 - (1-\epsilon)\rho^N(1-x)\right)$. Therefore, we have

$$\begin{aligned}
\bar{\lambda}^N(\epsilon, x) &= e^{-\mathsf{m}(1-\epsilon)\cdot\rho^N(1-x)} \\
&\leq e^{-\mathsf{m}(1-\epsilon)\frac{\mu+\rho(1-x)}{\mu+G_N(p)+1}}, \text{ if } x \geq \frac{1}{N} \qquad \text{(B.5)} \\
&< e^{-\mu} \cdot \frac{\sqrt{(1-p)^2+4px} - (1-p)}{2p} \\
&< \frac{\sqrt{(1-p)^2+4px} - (1-p)}{2p},
\end{aligned}$$

where (B.5) follows from Lemma VI.2. The polynomial $f(y) = py^2 + (1-p)y - x$ is a convex function of y, with the only positive root at $y = \frac{\sqrt{(1-p)^2+4px}-(1-p)}{2p}$. So, if $y < \frac{\sqrt{(1-p)^2+4px}-(1-p)}{2p}$, then $f(y) < 0$. Hence, $\left[(1-p) + p\bar{\lambda}(\epsilon, x)\right] \bar{\lambda}(\epsilon, x) - x < 0$ and the density evolution equation converges, as long as $x \geq \frac{1}{N}$. So, the probability of erasure is upper bounded by $1/N$.

Note that $\int_0^1 \rho^{(N)}(x)\,\mathrm{d}x$ is a monotonically increasing sequence, upper bounded by $1 - \frac{p}{2}$. So, in the limit of infinite blocklengths the design rate is given by

$$R = \lim_{N\to\infty} \frac{\int_0^1 \lambda(x)\,\mathrm{d}x}{\int_0^1 \rho^{(N)}(x)\,\mathrm{d}x} = \frac{(1-\epsilon)(1-e^{-\alpha})}{\mu + (1-\frac{p}{2})}. \qquad \blacksquare$$

From Theorem VI.1, we conclude that the code ensemble $\mathrm{LDGM}\big(n, \lambda(x), \rho^N(x)\big)$ can achieve the extremal symmetric point of the capacity region. Unfortunately, one can show (e.g., see Theorem VI.3) that this ensemble cannot simultaneously achieve both the extremal symmetric point and the corner points of the SW region. In Fig. 34, this can also be observed numerically via the density evolution ACPR (DE-ACPR) of this ensemble for $N = 2048$.

Theorem VI.3. *LT codes cannot simultaneously achieve the extremal symmetric point and a corner point of the Slepian-Wolf region, under iterative decoding.*

Proof. A corner point is given by the channel condition

$$(\epsilon_1, \epsilon_2) = (1 - (1-p)R, 1 - (1 - p/2)R)\,.$$

The density evolution equations are

$$\begin{aligned} x_{i+1} &= \left[(1-p) + p\bar{\lambda}^N(\epsilon_2, y_i)\right] \bar{\lambda}^N(\epsilon_1, x_i) \\ y_{i+1} &= \left[(1-p) + p\bar{\lambda}^N(\epsilon_1, x_i)\right] \bar{\lambda}^N(\epsilon_2, y_i), \end{aligned} \tag{B.6}$$

where $\bar{\lambda}^N(\epsilon, x) = \lambda\left(1 - (1-\epsilon)\rho^N(1-x)\right)$. To analyze the convergence of the ensem-

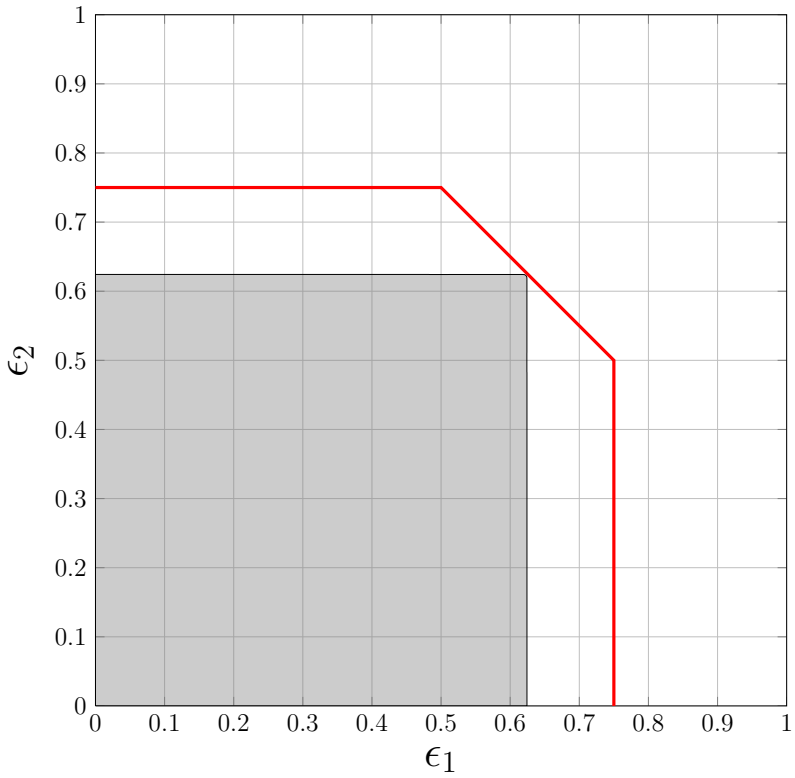

Fig. 34. ACPR (Density Evolution threshold) of the optimized (erasure channel) LT Code with $N = 2048$

ble LDGM$\big(n, \lambda(x), \rho^{(N)}(x)\big)$, consider the functions

$$f(x,y) = \left[(1-p) + p\bar{\lambda}^N(\epsilon_2, y)\right] \bar{\lambda}^N(\epsilon_1, x) - x$$
$$g(x,y) = \left[(1-p) + p\bar{\lambda}^N(\epsilon_1, x)\right] \bar{\lambda}^N(\epsilon_2, y) - y.$$

The condition for convergence of the density evolution equations are given by $f(x,y) < 0$ and $g(x,y) < 0$. When $\epsilon_1 < \epsilon_2$, we can approximately characterize the convergence by analyzing the condition $g(0,y) < 0$. We have

$$\begin{aligned} g(0,y) &= \left[(1-p) + p\lambda(\epsilon_1)\right] \lambda\left(1 - (1-\epsilon_2)\rho^N(1-y)\right) - y \\ &< \left[(1-p) + p\lambda(\epsilon_1)\right] \lambda\left(1 - (1-\epsilon_2)\rho(1-y)\right) - y \\ &= k\left(\sqrt{1+ay} - 1\right)^{\beta} - y, \end{aligned}$$

where

$$k = \left(\frac{\mathrm{e}^{-\mu}(1-p)}{2p}\right)^{\frac{1-\epsilon_2}{1-\epsilon_0}} \left[(1-p) + p\mathrm{e}^{-\alpha(1-\epsilon_1)}\right],$$
$$\beta = \frac{1-\epsilon_2}{1-\epsilon_0} \text{ and } a = \frac{4p}{(1-p)^2}$$

The fixed point of $g(0,y)$ can be found by solving

$$y = k\left(\sqrt{1+ay} - 1\right)^{\beta}, \text{ i.e.,}$$
$$\sqrt{1+ay} = 1 + k^{-1/\beta} y^{1/\beta}$$

This equation is of the form

$$k^{-2/\beta} y^{(2/\beta - 1)} + 2k^{-1/\beta} y^{(1/\beta - 1)} - a = 0,$$

the root of which is approximately equal to the root of the quadratic

$$k^{-2/\beta} z^2 + 2k^{-1/\beta} z - a,$$

where $z = y^{(1/\beta - 1/2)}$. The positive root of the quadratic is given by $z = \frac{-1+\sqrt{1+a}}{k^{-1/\beta}}$. So, the fixed point of density evolution is $y \approx \left(\frac{2p}{(1-p)k^{-1/\beta}}\right)^{\frac{2\beta}{2-\beta}} = \left(\frac{2p}{(1-p)k^{-1/\beta}}\right)^{2(1-p)} = \left(\mathrm{e}^{-\mu}\left[(1-p) + p\mathrm{e}^{-\alpha(1-\epsilon_1)}\right]\right)^{2(1-p)} > 0$.

Due to the presence of a constant fixed point, which does not approach 0 even in the limit of infinite maximum degree, the residual erasure rate is always bounded away from 0. So, the ensemble $\mathrm{LDGM}\big(n, \lambda(x), \rho^{(N)}(x)\big)$ cannot converge at a corner point of the capacity region. ■

APPENDIX C

LDPC CODE DESIGN FOR THE SW PROBLEM*

In this appendix, we present a way to construct near-universal codes, without using spatial-coupling.

Staggering

It is well known that single-user codes perform well at the corner points of the SW region. Although single-user codes do not perform well for symmetric channel conditions, they can be used to construct staggered codes that perform well at the corner points and for symmetric channel conditions. Consider 2 sources with $Lk + (1-\beta)k$ bits each. Without loss of generality, add βk zeros at the beginning for source $U^{[1]}$ and add βk zeros at the end for source $U^{[2]}$, to get $(L+1)k$ bits. We call β the staggering fraction. Next encode each block of k bits using a punctured $(n-k, k)$ LDPC code. The rate loss incurred by the addition of βk zeros can be made arbitrarily small by increasing the number of blocks L. At the decoder, one has the following structure: The performance of this staggered structure can be understood by considering the erasure case in the limit $L \to \infty$.

Theorem VI.4. *Consider transmission over erasure channels with erasure rates* $(\epsilon^{[1]}, \epsilon^{[2]})$ *using capacity approaching punctured* $(n-k, k)$ *LDPC codes. The staggered*

*Copyright 2010 IEEE. Reprinted, with permission, from A. Yedla, H. D. Pfister, and K. R. Narayanan, "LDPC code design for transmission of correlated sources across noisy channels without CSIT," in *Proc. Int. Symp. on Turbo Codes & Iterative Inform. Proc.*, Brest, France, Sept. 2010, pp. 474–478. For more information, go to `http://thesis.tamu.edu/forms/IEEE\%20permission\%20note.pdf/view`.

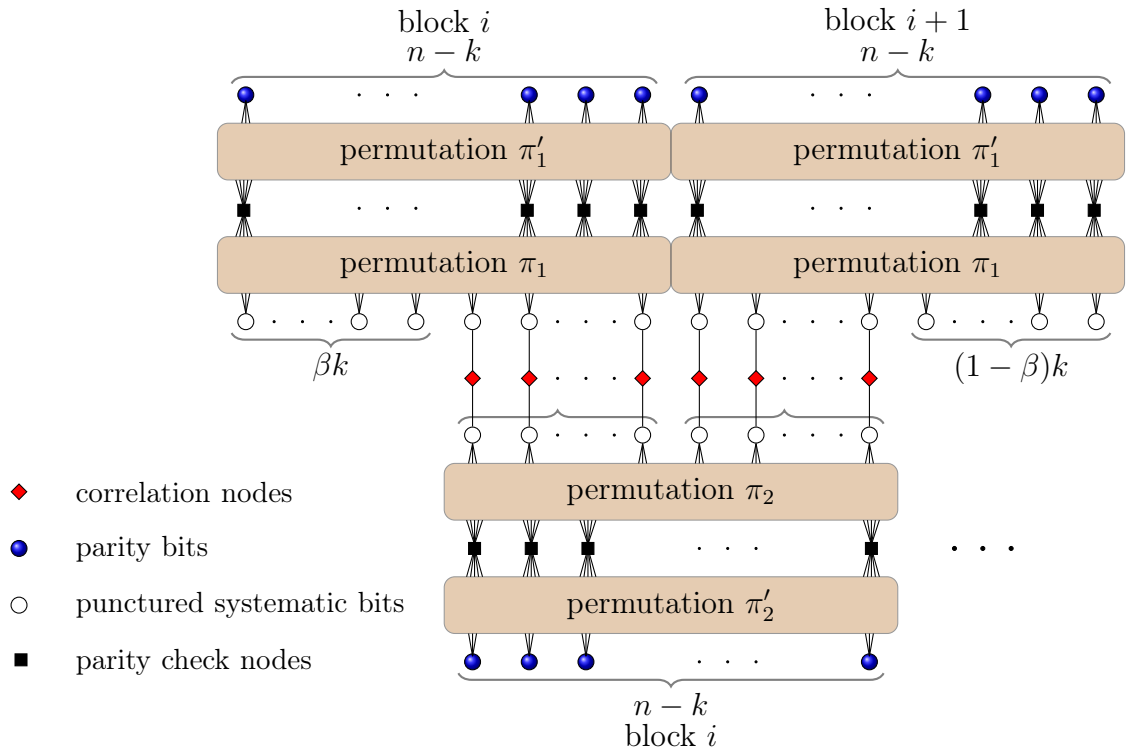

Fig. 35. Decoder structure for staggered codes

block code (with staggering fraction β) allows reliable communication for channel parameters

$$\epsilon^{[1]} \leq \min\{1 - R(1-\beta), 1 - R(1-p\beta)\}, \text{ and}$$

$$\epsilon_2 \leq 1 - R(1 - p(1-\beta)),$$

where $R = k/(n-k)$ is the design rate of the code.

Proof. Consider the first block for source $U^{[1]}$. The parity bits see a BEC($\epsilon^{[1]}$) channel and the source bits see an effective BEC($1-\beta$) channel (assuming no information comes from the decoder on the other side). So the effective erasure rate at the first block is $(1-R')\epsilon_1 + R'(1-\beta)$ ($R' = k/n$ is the rate of the code before puncturing). The code can decode as long as $R' \leq 1-((1-R')\epsilon_1 + R'(1-\beta))$ i.e., $\epsilon^{[1]} \leq 1 - R(1-\beta)$.

Suppose the first block of $U^{[1]}$ can decode successfully, then the source bits in the first block of $U^{[2]}$ see an effective channel of $(1-\beta)(1-p)+\beta$. The parity bits see a channel with erasure probability $\epsilon^{[2]}$. So, the effective channel seen by the first block of the second code is $(1-R')\epsilon^{[2]} + R'(1-p(1-\beta))$. So this block can be decoded as long as $\epsilon^{[2]} \leq 1 - R(1-p(1-\beta))$. Now proceed to the second block of $U^{[1]}$. The effective channel seen by the source bits of this code is $\beta(1-p)+(1-\beta)$. The parity bits see a channel with erasure probability $\epsilon^{[1]}$. So, the effective channel seen by the second block of the first code is $R'(\beta(1-p)+(1-\beta)) + (1-R')\epsilon^{[1]}$. This block can be decoded as long as $\epsilon^{[1]} \leq 1 - R(1-p\beta)$. The decoding continues by alternating between blocks of $U^{[1]}$ and $U^{[2]}$. This proves the claim. ■

Corollary VI.5. *Consider transmission over erasure channels using capacity approaching punctured $(n-k,k)$ LDPC codes. The staggered block code (with staggering fraction $\beta = 1/2$) allows reliable communication at both the corner points and the symmetric channel condition.*

Proof. The proof follows by matching the conditions of the previous theorem to a corner point and the extremal symmetric point of the SW region. Consider the extremal symmetric channel condition in the Slepian-Wolf region. The channel parameters are given by $(\epsilon^{[1]}, \epsilon^{[2]}) = (1-(1-\frac{p}{2})R, 1-(1-\frac{p}{2})R)$. For successful decoding at the extremal symmetric channel condition, we obtain the condition $\beta = 1/2$. A corner point of the Slepian-Wolf region is given by $(\epsilon^{[1]}, \epsilon^{[2]}) = (1-R, 1-(1-p)R)$. Successful decoding at this point requires that $0 \leq \beta \leq 1$. So, for $\beta = 1/2$, the above staggered structure allows successful communication at both the corner points and the symmetric channel condition. ■

Remark VI.1. Note that staggered capacity-approaching codes can be used to communicate at a corner point and any other point on the dominant face of the SW region

(using different values of β).

For general channels we can analyze the performance of the staggered code using density evolution. Let $i \in \{1, \ldots, L\}$ and $\mathsf{a}_\ell^{(i)}$ and $\mathsf{b}_\ell^{(i)}$ denote the density of the messages emanating from the variable nodes at iteration ℓ, corresponding to codes 1 and 2 in block i. The DE equations can be written as follows:

$$\begin{aligned}
\mathsf{a}_{\ell+1}^{(i)} &= \left[\gamma\left(\beta f\left(L\left(\rho(\mathsf{b}_\ell^{(i-1)})\right)\right) + (1-\beta) f\left(L\left(\rho(\mathsf{b}_\ell^{(i)})\right)\right)\right) + (1-\gamma)\mathsf{a}_{\mathrm{BMSC}}\right] \\
&\quad \circledast \lambda(\rho(\mathsf{a}_\ell)) \\
\mathsf{b}_{\ell+1}^{(i)} &= \left[\gamma\left((1-\beta) f\left(L\left(\rho(\mathsf{a}_\ell^{(i)})\right)\right) + \beta f\left(L\left(\rho(\mathsf{a}_\ell^{(i+1)})\right)\right)\right) + (1-\gamma)\mathsf{b}_{\mathrm{BMSC}}\right] \\
&\quad \circledast \lambda(\rho(\mathsf{b}_\ell)).
\end{aligned} \tag{C.1}$$

Here, $\mathsf{a}_\ell^{(i)}, \mathsf{b}_\ell^{(i)} = \Delta_{+\infty}$ (the delta function at ∞) for $i \notin \{1, \ldots, L\}$.

Differential Evolution

Throughout this section, we use x to denote an element of $\mathbb{R}^n$ for some $n \in \mathbb{N}$, and x_i to denote its ith component. Let $\mathcal{V} = \{i \,|\, \lambda_i \neq 0\}$ and $\mathcal{P} = \{i \,|\, \rho_i \neq 0\}$ be the support sets of the variable and parity-check degree distributions respectively, which are assumed to be known. The correlation parameter p is fixed. We design LDPC codes for this scenario using differential evolution [55], for a design rate R_d. In an n-dimensional search space, a fixed number of vectors are randomly initialized and then evolved over time, exploring the search space, to locate the minima of the objective function. Let

$$\Delta^{n-1} = \left\{ x \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} x_i = 1, x_i \geq 0, i = 1, \cdots, n \right\}$$

denote the unit simplex and $n_v = |\mathcal{V}|$, $n_p = |\mathcal{P}|$. Then, the search space for all variable (check) degree profiles is Δ^{n_v-1} (Δ^{n_p-1}). The optimization is performed over

the search space $\mathcal{S} = \Delta^{n_v-1} \times \Delta^{n_p-1}$, with parameter vectors $x = [x_\lambda, x_\rho]$,[2] where $x_\lambda \in \Delta^{n_v-1}, x_\rho \in \Delta^{n_p-1}$. In our optimization procedure, we expand the search space to $\mathcal{S}' = \{x \in \mathbb{R}^{n_v+n_p}, \sum_i (x_\lambda)_i = 1, \sum_i (x_\rho)_i = 1\}$, for simplicity in the crossover stage. We generate an initial population of trial degree distributions by uniformly sampling the degree distributions from the unit simplex.

For the optimization to work well, differential evolution requires an initial population of trial vectors which are spread out uniformly across the search space. To obtain a sample x uniformly from Δ^{n-1}, we generate uniform random variables $u_i \sim U[0,1], i = 1, 2, \cdots, n-1$. Define $u_0 = 0, u_n = 1$ and let π_u be the permutation that sorts (u_i) in ascending order i.e., if $i \leq j$, then $u_{\pi_u(i)} \leq u_{\pi_u(j)}$. For $i = 1, \cdots, n$, define $x_i = u_{\pi_u(i)} - u_{\pi_u(i-1)}$, and $x = (x_i)$. Then x has a uniform distribution over Δ^{n-1}.

Let C be a finite subset of channel parameters $(\alpha^{[1]}, \alpha^{[2]})$ that correspond to the sum rate constraint of the SW conditions for a design rate R_d. Let $\Gamma : \mathcal{S}' \times \mathsf{C} \to [0,1] \times [0,1]$, $(x, \alpha^{[1]}, \alpha^{[2]}) \mapsto (e_1, e_2)$ be the function that gives the residual error probability[3] (using joint density evolution as described in Section 2) for each decoder, for a pair of codes with degree distribution x (i.e., (x_λ, x_ρ)), when transmitted over channels with parameters $(\alpha^{[1]}, \alpha^{[2]})$. We use discretized density evolution [41][4] to compute the performance of an ensemble.

For our design, we want the code to achieve an arbitrarily low probability of error on C and we want the rate of the code $R(x)$ to be as close to the design rate R_d as

[2] $(x_\lambda, \mathcal{V})$ and $(x_\rho, \mathcal{P})$ correspond to the variable and parity node degree profiles respectively.

[3] We set the maximum number of iterations to 100 for all the designs considered in this paper. Density evolution is stopped when the maximum number of iterations is reached or the difference in the residual error probability between successive iterations is less than 10^{-8}.

[4] A 9 bit linear quantization is used over a likelihood ratio range $[-20, 20]$

possible. So, we define the cost function,

$$\mathcal{F}(x) = a \cdot \left(\sum_{(\alpha_1,\alpha_2)\in\mathcal{C}} \left(1 - \mathbb{1}_{\{(\alpha_1,\alpha_2)|\Gamma(x,\alpha_1,\alpha_2)\preceq(\tau,\tau)\}}\right) \right) + b \cdot (R_d - R(x)),$$

if $x \in \mathcal{S}$ and $\mathcal{F}(x) = \infty$, if $x \in \mathcal{S}'\backslash\mathcal{S}$. The rate of the code $R(x) = R(x_\lambda, x_\rho)$ is computed as in (2.6). The constants a and b are chosen through trial and error. The parameters chosen for the designs considered in this paper are $\tau = 10^{-5}, a = 10$ and $b = 30$. The optimization is then setup as $\min_{x\in\mathcal{S}'} \mathcal{F}(x)$.

We use a variant of differential evolution, with the mutation and recombination scheme given in [56]. The resulting codes are then staggered as described in Section C.

Results

The design was performed to maximize the ACPR, in contrast to previous work. For the erasure correlation model, the optimization was performed for a design rate of $R_d = 0.57$ after puncturing and source correlation $p = 0.5$. The resulting degree profile

$$\lambda(x) = 0.3633x + 0.2834x^2 + 0.2315x^6 + 0.1217x^{19},$$
$$\rho(x) = 0.531776x^3 + 0.468224x^5,$$

has a design rate of 0.3308 and transmission rate 0.4962. The ACPR for this code is shown in Fig. 36 along with the SW region for the rate pair $(0.4962, 0.4962)$. This shows optimized ensembles can achieve a large portion of the SW region.

The BSC source correlation parameter was $p = 0.9$ and the optimization was

performed for a design rate $R_d = 0.5$ after puncturing. The resulting degree profile

$$\lambda(x) = 0.26725x + 0.26823x^2 + 0.07557x^3 + 0.212x^6 + 0.027898x^7 + 0.0061593x^8 + 0.0011654x^{14} + 0.14173x^{19},$$

$$\rho(x) = 0.37856x^3 + 0.56211x^5 + 0.0080803x^9 + 0.028448x^{14} + 0.0095319x^{19} + 0.013267x^{24},$$

has a design rate of 0.323 and transmission rate 0.476. The ACPR for this code is shown in Fig. 37 along with the SW region for the rate pair $(0.476, 0.476)$. These results show that ensembles optimized using differential evolution can achieve almost

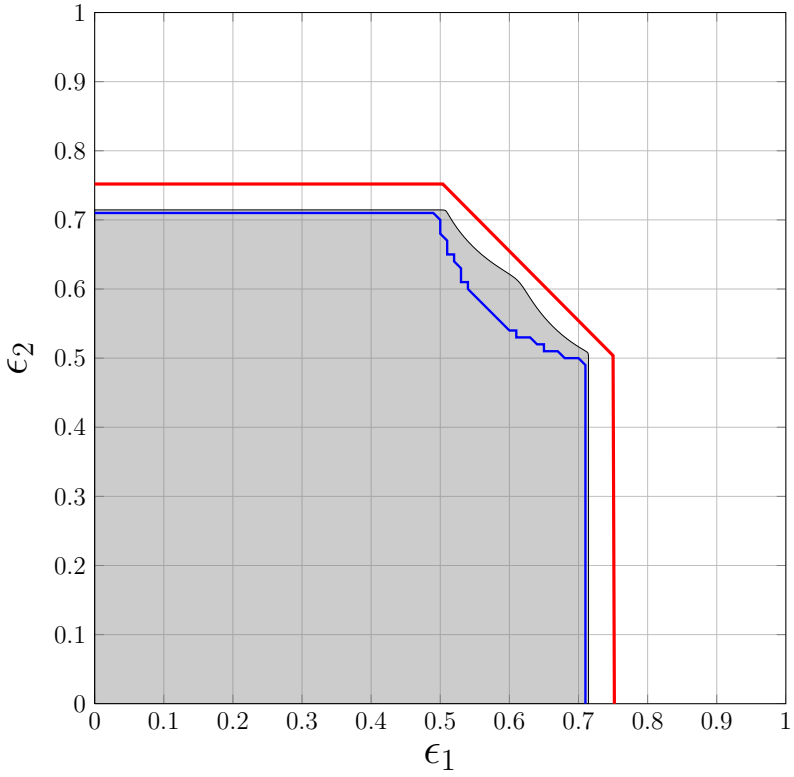

Fig. 36. ACPR (Density Evolution threshold) of an optimized (erasure channel) LDPC Code of rate 0.3308 is shown in blue. The grey area is the ACPR after staggering.

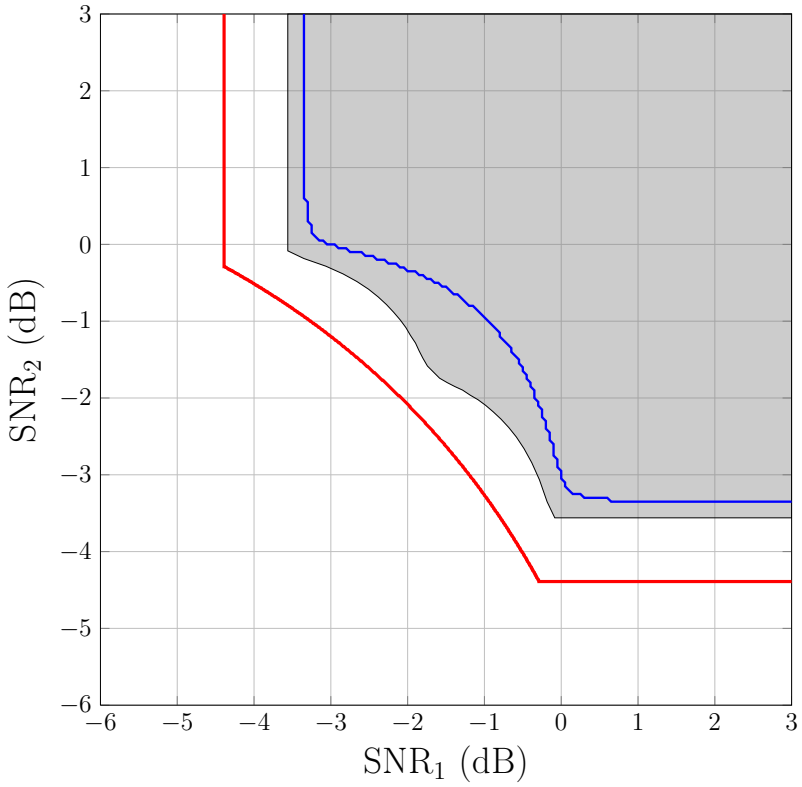

Fig. 37. ACPR (Density Evolution threshold) of an optimized (AWGN channel) LDPC Code of rate 0.323 is shown in blue. The grey area is the ACPR after staggering.

the entire SW region.

VITA

Arvind Yedla received his Bachelors of Technology degree in electrical engineering from the Indian Institute of Technology, Madras, India in 2005. He joined the electrical and computer engineering graduate program at Texas A & M University, College Station, Texas, to pursue his doctoral studies and graduated with his Ph.D. in August 2012.

Dr. Yedla was an engineer at Tejas Networks India Ltd. from 2005 to 2006 in Bangalore, India. He joined Samsung Information Systems America in San Diego, California, as a senior engineer in 2012. His research interests include information theory and channel coding. He can be reached at the Zachry Engineering Center, TAMU, College Station, TX or at arvind.yedla@gmail.com.

The typist for this dissertation was Arvind Yedla.