Code Design for the Noisy Slepian-Wolf Problem

Arvind Yedla, *Student Member, IEEE,* Henry D. Pfister, *Member, IEEE,*

and Krishna R. Narayanan, *Senior Member, IEEE*

arXiv:1201.0409v1 [cs.IT] 2 Jan 2012

Abstract

We consider a noisy Slepian-Wolf problem where two correlated sources are separately encoded (using codes of fixed rate) and transmitted over two independent binary memoryless symmetric channels. The capacity of each channel is characterized by a single parameter which is not known at the transmitter. The goal is to design systems that retain near-optimal performance without channel knowledge at the transmitter.

It was conjectured that it may be hard to design codes that perform well for symmetric channel conditions. In this work, we present a provable capacity-achieving sequence of LDGM ensembles for the erasure Slepian-Wolf problem with symmetric channel conditions. We also introduce a staggered structure which enables codes optimized for single user channels to perform well for symmetric channel conditions.

We provide a generic framework for analyzing the performance of joint iterative decoding, using density evolution. Using differential evolution, we design punctured systematic LDPC codes to max-

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA (email: yarvind@tamu.edu; hpfister@tamu.edu; krn@tamu.edu).

imize the region of achievable channel conditions. The resulting codes are then staggered to further increase the region of achievable parameters. The main contribution of this paper is to demonstrate that properly designed irregular LDPC codes can perform well simultaneously over a wide range of channel parameters.

Index Terms

LDPC codes, LDGM codes, density evolution, correlated sources, non-systematic encoders, joint decoding, differential evolution, area theorem.

I. INTRODUCTION

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 [1], [2]. A sensor network 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. This problem is often referred to as the sensor reachback problem. 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 (MAC) protocol in place, which 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. A simple sensor network consisting of two sensors is shown in Fig. 1. This problem of distributed encoding and transmission over independent channels gives a noisy version of

the celebrated Slepian-Wolf (SW) problem. The SW problem was introduced and solved in the landmark paper [3], 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. A variety of coding schemes have been designed that can achieve the SW bound when channel state information is known at the transmitter.

A. Prior Work

The first practical SW coding scheme was introduced by Wyner and is based on linear error-correcting codes [4]. Chen et al. related the SW (distributed source coding) problem to channel coding via an equivalent channel describing the source correlation [5], [6]. 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 [7].

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 [8]. 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 [9].

The performance of concatenated LDGM codes has been studied in [10] and that of Turbo codes in [9]. Serially concatenated LDPC and convolutional codes were also considered in [11], where the outer LDPC code is used for distributed source coding.

It was conjectured in [12] that LDPC codes do not perform well for the noisy SW problem[1] and that it is hard to design codes that perform well for symmetric channel conditions. In this work, we show a sequence of LDGM codes which approach the SW bound for symmetric channel conditions.

B. Universality

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. [2]. 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. This problem of unknown channel state at the transmitter naturally arises in the context of many multi-user scenarios, including cellular telephony.

For fixed user code rates, reliable communication is theoretically possible over a wide range of channel conditions [13]. We call a system *universal* if it provides good performance for all system parameters that do not violate theoretical limits. 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, the performance can deteriorate markedly as the channel conditions change. So, we design LDPC codes which are robust to variation in channel conditions[2]. Such schemes are 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).

---

[1]The authors consider only systematic LDPC codes

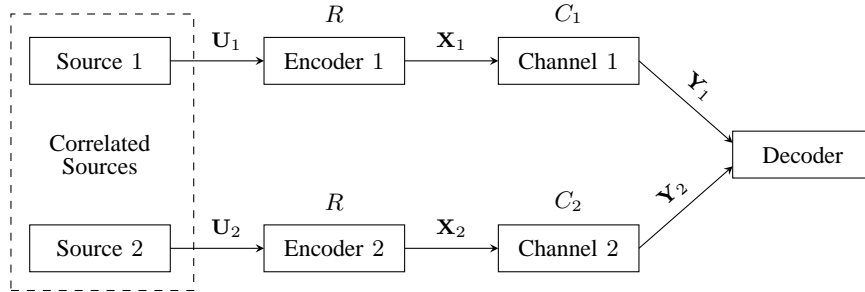[2]Unfortunately, the LDGM codes that achieve the symmetric channel condition are not universal.

Fig. 1.   System Model

II. PROBLEM SETUP

Consider the problem of transmitting the outputs of two discrete memoryless correlated sources, (U_1, U_2), to a central receiver through two independent discrete memoryless channels with capacities C_1 and C_2, respectively. The system model is shown in Figure 1. 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 ($\mathbf{U}_1$ and $\mathbf{U}_2$) to n_1 and n_2 output symbols ($\mathbf{X}_1$ and $\mathbf{X}_2$), respectively. The rates of the encoders are given by $R_1 = k/n_1$ and $R_2 = k/n_2$. The decoder receives $(\mathbf{Y}_1, \mathbf{Y}_2)$ and makes an estimate of $(\mathbf{U}_1, \mathbf{U}_2)$.

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 rate R (i.e., $R = k/n, n_1 = n_2 = n$). Reliable transmission over a channel pair (α_1, α_2) is possible as long as the SW conditions (1) are satisfied.

$$\frac{C_1(\alpha_1)}{R} \geq H(U_1|U_2)$$
$$\frac{C_2(\alpha_2)}{R} \geq H(U_2|U_1) \tag{1}$$
$$\frac{C_1(\alpha_1)}{R} + \frac{C_2(\alpha_2)}{R} \geq H(U_1, U_2)$$

For a given pair of encoding functions of rate R and a joint decoding algorithm, a pair of

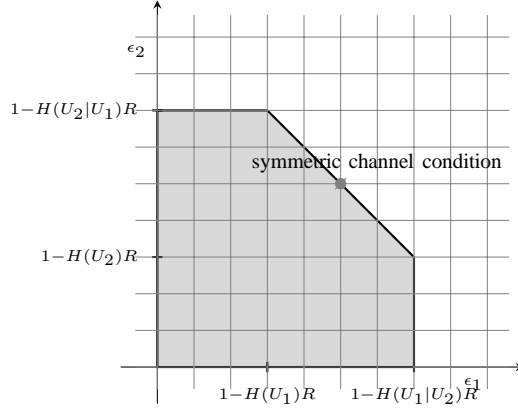Fig. 2.   The SW region for erasure channels, for a fixed rate pair (R, R)

channel parameters (α_1, α_2) is *achievable* if the encoder/decoder combination can achieve an arbitrarily low probability of error for limiting block-lengths (i.e., $k \to \infty$). 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 region* as the set of all channel parameters (α_1, α_2) for which (1) is satisfied. The SW region for the erasure channel family is shown in Figure 2.

In this paper, we consider the following scenarios:

1) The channels are erasure channels and the source correlation is modeled through 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 III admits general correlation models and memoryless channels.

A. *Erasure Correlation*

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

$$(U_1, U_2) = \begin{cases} (X, X') & \text{if } Z = 0 \\ (X, X) & \text{if } Z = 1 \end{cases}.$$

We have $H(U_1|U_2) = H(U_2|U_1) = 1 - p$ and $H(U_1, U_2) = 2 - p$. This correlation model can be incorporated into the Tanner graph (see Section III-A, III-C) 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 [14].

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.

B. *BSC Correlation*

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, $H(U_1|U_2) = H(U_2|U_1) = h_2(p)$ and $H(U_1, U_2) = 1 + h_2(p)$.

This correlation model can be incorporated into the Tanner graph at the decoder (described in Section III-C) as check nodes between the source bits, with a hidden node representing the auxiliary random variable Z (which carries a constant log-likelihood ratio $\log \frac{1-p}{p}$) 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.

C. Existence of Universal codes

In this section, we discuss the existence of universal coding schemes, for the system model considered in Figure 1. Let $I_{\alpha_1}(X_1; Y_1)$ and $I_{\alpha_2}(X_2; Y_2)$ denote the mutual information between the channel inputs and outputs when the channel parameters are given by α_1 and α_2. The following theorem shows the existence of codes which have large ACPRs.

Theorem 1. *Consider encoders with rate pair* (R, R)*. For a fixed pair of channel conditions* (α_1, α_2)*, which are not known at the transmitter, random coding with typical-set decoding at the receiver can achieve an average probability of error* $\bar{P}_{e, \alpha_1, \alpha_2}$ *bounded above by* $2^{-n\gamma(\alpha_1, \alpha_2)}$*, where*

$$\gamma(\alpha_1, \alpha_2) = \min \left\{ I_{\alpha_1}(X_1; Y_1) - RH(U_1 \mid U_2), \right.$$

$$I_{\alpha_2}(X_2; Y_2) - RH(U_2 \mid U_1),$$

$$\left. I_{\alpha_1}(X_1; Y_1) + I_{\alpha_2}(X_2; Y_2) - RH(U_1, U_2) \right\}.$$

Hence, there exists an encoder for which the probability of error

$$P_{e, \alpha_1, \alpha_2} \leq 2^{-n\gamma(\alpha_1, \alpha_2)}.$$

Proof: This follows from extending the proofs in [15] to the SW problem. ∎

Remark 1. *A simple application of Fano's inequality shows that any pair of channel parameters for which $\gamma(\alpha_1, \alpha_2) < 0$ are not achievable (the probability of error is strictly bounded away from zero). For binary memoryless symmetric (BMS) channels, the condition $\gamma(\alpha_1, \alpha_2) > 0$ translates to the conditions in (1). So, the conditions in (1) are both necessary and sufficient for transmission over BMS channels.*

Remark 2. *For BMS channels, the achievable channel parameter region for a random code is a dense subset of the entire SW region for limiting block-lengths. This follows by using Theorem 1 and applying the Markov inequality. This result is also easily extended to random linear codes.*

We conclude that, for a given rate pair (R, R), a single encoder/decoder pair suffices to communicate the sources over all pairs of BMS channels in the SW region. Thus, one can obtain optimal performance even without knowledge of (α_1, α_2) at the transmitter. We refer to such encoder/decoder pairs as being *universal*. This means that random codes with typical-set decoding are universal for BMS channels.

While random codes with typical-set decoding are universally good, encoding and decoding is known to be impractical due to its large complexity. This motivates the search for low complexity encoding/decoding schemes which are universal.

III. ANALYSIS

A. LDGM Codes

Assume that the sequences $\mathbf{U}_1$ and $\mathbf{U}_2$ are encoded using LDGM codes with a degree distribution pair (λ, ρ). Based on standard notation [16], we let $\lambda(x) = \sum_i \lambda_i x^{i-1}$ be the degree distribution (from an edge perspective) corresponding to the variable nodes and $\rho(x) = \sum_i \rho_i x^{i-1}$ be the degree distribution (from an edge perspective) of the parity-check nodes in the decoding graph. The coefficient λ_i (resp. ρ_i) gives the fraction of edges that connect to the variable nodes
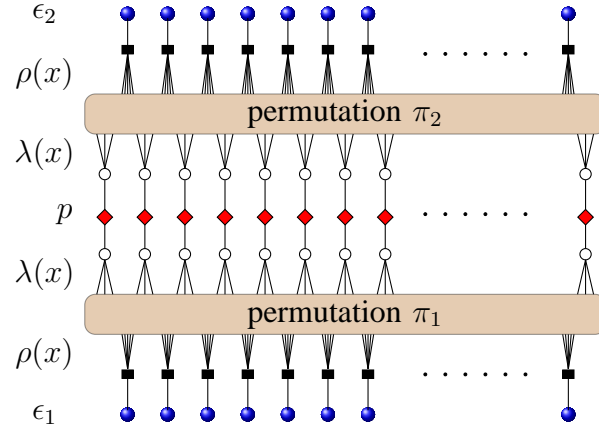
Fig. 3.   Tanner Graph of an LDGM (LT) Code with erasure correlation between the sources

(resp. parity-check nodes) of degree i. Likewise, L_i (resp. R_i) is the fraction of variable nodes

(resp. check nodes) with degree i.

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 $L(x) = \mathsf{e}^{\mathsf{m}(x-1)}$. Throughout

this section, we consider the erasure correlation model described in Section II-A.

The Tanner graph [16] for the code is shown in Fig. 3. 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 [16] in terms of the variable-node

to check-node messages can be written as

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

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

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

$$x_{\ell+1} = \left[(1-p) + p\lambda\left(1 - (1-\epsilon)\rho(1-x_\ell)\right)\right]\lambda\left(1 - (1-\epsilon)\rho(1-x_\ell)\right). \tag{2}$$

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

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

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

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

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 [17].

Theorem 2. *Consider transmission over erasure channels with parameters $\epsilon_1 = \epsilon_2 = \epsilon$. For $N \in \mathbb{N}$ and $\mu > 0$, define*

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

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

$$\lambda(x) = e^{\mathsf{m}(x-1)} \; and \; \rho^N(x) = \frac{\mu + \sum_{i=1}^{N} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i + x^N}{\mu + G_N(p) + 1}, \tag{4}$$

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: See Appendix A.   ∎

From Theorem 2, we conclude that the optimized ensemble $\text{LDGM}\big(n, \lambda(x), \rho^N(x)\big)$ can achieve the extremal symmetric point of the capacity region. Unfortunately, one can show that this ensemble cannot simultaneously achieve both the extremal symmetric point and the corner points of the SW region. In Figure 4, this can also be observed numerically via the density evolution ACPR (DE-ACPR) of this ensemble for $N = 2048$.

B. Puncturing and LDPC Codes

In [18], it is shown that correlated codes are suboptimal when transmitting correlated sources over independent channels. The conditions in (1) implicitly assume the use of uncorrelated codes
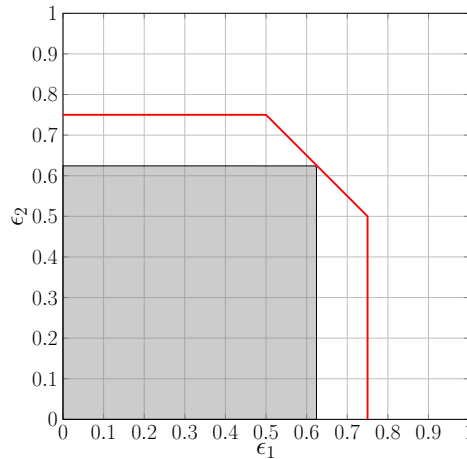

Fig. 4.   ACPR (Density Evolution threshold) of the optimized (erasure channel) LT Code with $N = 2048$
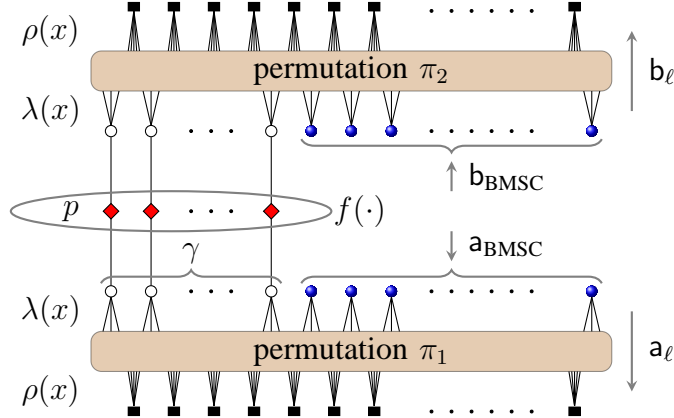
Fig. 5.   Tanner Graph of an LDPC Code with source correlation

i.e., we require the average mutual information (over the code ensemble) $I(X_1; X_2) = 0$.

This condition is clearly not satisfied when we use a systematic LDPC ensemble. This also explains the loss in performance of systematic LDPC codes when compared to Turbo codes, as shown in [12]. To ensure the independence of the transmitted symbols, we use LDPC ensembles with punctured systematic encoders.

C. Density Evolution for LDPC codes

Assume that the sequences $\mathbf{U}_1$ and $\mathbf{U}_2$ 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 γ.

The Tanner graph [16] for the joint decoder is shown in Figure 5. 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 a_ℓ and b_ℓ denote the density[3] of the messages emanating from the variable nodes at iteration ℓ, corresponding to codes 1 and

---

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

2. The density evolution equations [16] can be written as follows

$$
\begin{aligned}
\mathsf{a}_{\ell+1} &= \left[\gamma f\Big(L\left(\rho(\mathsf{b}_\ell)\right) \Big) + (1-\gamma)\mathsf{a}_{\text{BMSC}} \right] \circledast \lambda(\rho(\mathsf{a}_\ell)) \\
\mathsf{b}_{\ell+1} &= \left[\gamma f\Big(L\left(\rho(\mathsf{a}_\ell)\right) \Big) + (1-\gamma)\mathsf{b}_{\text{BMSC}} \right] \circledast \lambda(\rho(\mathsf{b}_\ell)),
\end{aligned}
\tag{5}
$$

where $\lambda(\mathsf{a}) = \sum_i \lambda_i \mathsf{a}^{\circledast(i-1)}$, $L(\mathsf{a}) = \sum_i L_i \mathsf{a}^{\circledast(i-1)}$, $\rho(\mathsf{a}) = \sum_i \rho_i \mathsf{a}^{\boxtimes(i-1)}$, a_{BMSC} and b_{BMSC} are the densities of the log-likelihood ratios received from the channel. The function f at the correlation nodes depends on the equivalent channel corresponding to the correlation model, as described in [5]. 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.

First consider the BSC correlation model. By symmetry of the problem, we can assume that user 1 transmits the all-zero codeword and the second user transmits a typical codeword. 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 [16, 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 i.e., $f(\mathsf{a}) = \mathsf{a}_{\text{BSC}(p)} \boxtimes \mathsf{a}$. This simplification enables us to proceed with density evolution assuming the transmission of an all-zero codeword for both the users.

Note that such a simplification is not necessary for the erasure correlation model. 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, so $f(\mathsf{a}) = (1-p) + p\mathsf{a}$.

The residual error probability at iteration ℓ, (e_1^ℓ, e_2^ℓ), is computed using the error functional

$\mathfrak{E}(\cdot)$ defined in [16, p. 201]:

$$e_1^\ell = \mathfrak{E}\left(\left[\gamma f\left(L\left(\rho(\mathsf{b}_\ell)\right)\right) + (1-\gamma)\mathsf{a}_{\mathrm{BMSC}}\right] \circledast L(\rho(\mathsf{a}_\ell))\right)$$

$$e_2^\ell = \mathfrak{E}\left(\left[\gamma f\left(L\left(\rho(\mathsf{a}_\ell)\right)\right) + (1-\gamma)\mathsf{b}_{\mathrm{BMSC}}\right] \circledast L(\rho(\mathsf{b}_\ell))\right).$$

D. Staggered Block Codes

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 3. *Consider transmission over erasure channels with erasure rates (ϵ_1, ϵ_2) using capacity approaching punctured $(n-k, k)$ LDPC codes. The staggered 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 $\mathrm{BEC}(\epsilon_1)$ channel and the source bits see an effective $\mathrm{BEC}(1 - \beta)$ channel (assuming no information comes from the
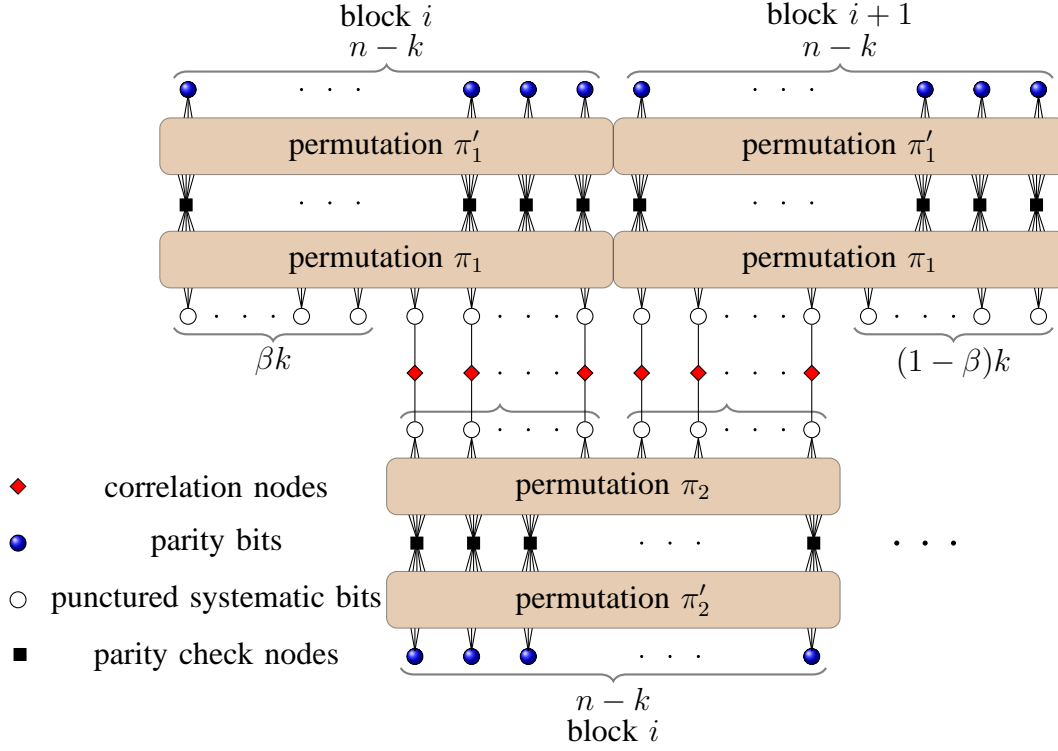
Fig. 6.   Decoder structure for staggered codes

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 ϵ_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 $\epsilon_2 \leq 1 - R(1-p(1-\beta))$. The decoding continues by alternating between blocks of U_1 and U_2. This proves the claim.   ∎

Corollary 1. *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. ■

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:

$$
\begin{aligned}
\mathsf{a}_{\ell+1}^{(i)} &= \left[\gamma \left(\beta f \left(L \left(\rho(\mathsf{b}_\ell^{(i-1)}) \right) \right) + (1 - \beta) f \left(L \left(\rho(\mathsf{b}_\ell^{(i)}) \right) \right) \right) + (1 - \gamma) \mathsf{a}_{\mathrm{BMSC}} \right] \circledast \lambda(\rho(\mathsf{a}_\ell)) \\
\mathsf{b}_{\ell+1}^{(i)} &= \left[\gamma \left((1 - \beta) f \left(L \left(\rho(\mathsf{a}_\ell^{(i)}) \right) \right) + \beta f \left(L \left(\rho(\mathsf{a}_\ell^{(i+1)}) \right) \right) \right) + (1 - \gamma) \mathsf{b}_{\mathrm{BMSC}} \right] \circledast \lambda(\rho(\mathsf{b}_\ell)).
\end{aligned}
\tag{6}
$$

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

E. 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 [19], for a design rate R_d. 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 $x = [x_\lambda, x_\rho]$[4], 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.

---

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

Let C be a finite subset of channel parameters (α_1, α_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[5] (using joint density evolution as described in Section III-C) for each decoder, for a pair of codes with degree distribution x (i.e., (x_λ, x_ρ)), when transmitted over channels with parameters (α_1, α_2). We use discretized density evolution [21][6] 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 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 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 [20]. The resulting codes are then staggered as described in Section III-D.

IV. RESULTS AND CONCLUDING REMARKS

This paper shows that the SW conditions are necessary and sufficient for communication of correlated sources through independent BMS channels, without channel state information at the transmitter. This implies that a single random code is sufficient to communicate with vanishing probability of error, for the entire SW region. We showed the achievability of the symmetric

---

[5]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}.

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

channel condition under message passing by providing a sequence of LDGM ensembles which can achieve an arbitrarly low probability of error.

We designed punctured systematic LDPC codes for the scenarios described in Section II. 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 Figure 7 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
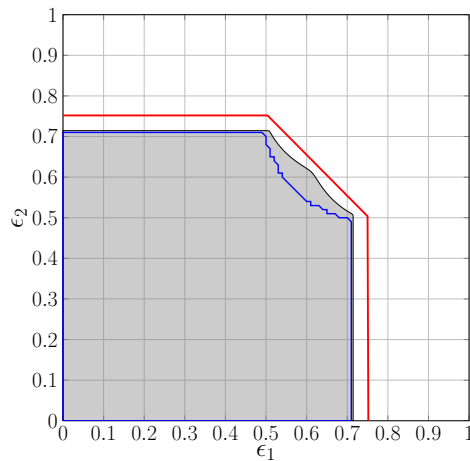

Fig. 7. ACPR (Density Evolution threshold) of an optimized (erasure channel) LDPC Code of rate 0.3308 is shown in blue. The grey area is the ACPR after staggering.
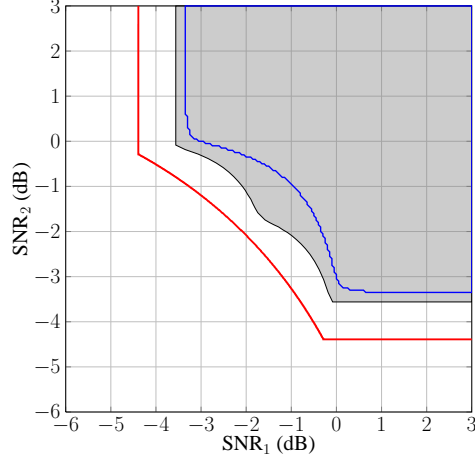
Fig. 8. ACPR (Density Evolution threshold) of an optimized (AWGN channel) LDPC Code of rate 0.323 is shown in blue. The grey area is the ACPR after staggering.

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 Figure 8 along with the SW region for the rate pair $(0.476, 0.476)$. These results show that ensembles optimized using differential evolution almost achieve the entire SW region.

APPENDIX A

PROOF OF THEOREM 2

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

Lemma 1.

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

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

$$
\begin{aligned}
\rho^N(x) &= \frac{\mu + \sum_{i=1}^{N} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i + x^N}{\mu + G_N(p) + 1} \\
&= \frac{\mu + \rho(x) + x^N}{\mu + G_N(p) + 1} - \frac{\sum_{i=N+1}^{\infty} \frac{\sum_{k=0}^{i-1} \binom{2i-1}{k} p^k}{i(1+p)^{2i-1}} x^i}{\mu + G_N(p) + 1} \\
&> \frac{\mu + \rho(x)}{\mu + G_N(p) + 1}.
\end{aligned}
\tag{7}
$$

(7) 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 (2), 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 $\bar{\lambda}^N(\epsilon, x) = \lambda \left(1 - (1-\epsilon)\rho^N(1-x) \right)$. We have,

$$
\begin{aligned}
\bar{\lambda}^N(\epsilon, x) &= e^{-\mathsf{m}(1-\epsilon) \cdot \rho^N(1-x)} \\
&\le e^{-\mathsf{m}(1-\epsilon)\frac{\mu+\rho(1-x)}{\mu+G_N(p)+1}}, \text{ if } x \ge \frac{1}{N} \\
&< e^{-\mu} \cdot \frac{\sqrt{(1-p)^2 + 4px} - (1-p)}{2p} \\
&< \frac{\sqrt{(1-p)^2 + 4px} - (1-p)}{2p},
\end{aligned}
\tag{8}
$$

where (8) follows from Lemma 1. 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 $x \ge \frac{1}{N}$. 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

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

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications Magazine, IEEE*, vol. 40, no. 8, pp. 102–114, 2002. 2

[2] C. Chong and S. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003. 2, 4

[3] 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. 3

[4] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. 20, no. 1, pp. 2–10, 1974. 3

[5] 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. 3, 14

[6] 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. 3

[7] 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. 3

[8] J. Barros and S. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 155–170, 2006. 3

[9] J. Garcia-Frias, "Joint source-channel decoding of correlated sources over noisy channels," in *Data Comp. Conf.*, 2001, pp. 283–292. 3

[10] 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.*, vol. 2005, no. 6, pp. 942–953, 2005. 3

[11] 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.*, vol. 1. IEEE, 2004, pp. 51–55. 3

[12] M. Martalo, G. Ferrari, A. Abrardo, M. Franceschini, and R. Raheli, "Density evolution-based analysis and design of LDPC codes with a priori information," in *Proc. Annual Workshop on Inform. Theory and its Appl.*, San Diego, CA, Feb. 2010. 4, 13

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. Wiley, 1991. 4

[14] 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. 7

[15] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *Ann. Math. Stats.*, pp. 1229–1241, 1959. 8

[16] T. J. Richardson and R. L. Urbanke, *Modern Coding Theory*. Cambridge, 2008. 9, 10, 13, 14, 15

[17] M. Luby, "Lt codes," in *Proc. of the 43rd Symp. on Foundations of Comp. Sci.*, Washington, D.C., June 2002, p. 271. 11

[18] J. Barros and S. Servetto, "The sensor reachback problem," 2003, submitted to *IEEE Trans. on Inform. Theory*. 12

[19] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Springer Verlag, 2005. 17

[20] A. Shokrollahi and R. Storn, "Design of efficient erasure codes with differential evolution," in *Proc. IEEE Int. Symp. Inform. Theory*, 2000. 18

[21] S. 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. 18