Predicting Error Floors of Structured LDPC Codes: Deterministic Bounds and Estimates

Lara Dolecek, Pamela Lee, Zhengya Zhang, Venkat Anantharam, Borivoje Nikolic, and Martin Wainwright

Abstract—The error-correcting performance of low-density parity check (LDPC) codes, when decoded using practical iterative decoding algorithms, is known to be close to Shannon limits for codes with suitably large blocklengths. A substantial limitation to the use of finite-length LDPC codes is the presence of an error floor in the low frame error rate (FER) region. This paper develops a deterministic method of predicting error floors, based on high signal-to-noise ratio (SNR) asymptotics, applied to absorbing sets within structured LDPC codes. The approach is illustrated using a class of array-based LDPC codes, taken as exemplars of high-performance structured LDPC codes. The results are in very good agreement with a stochastic method based on importance sampling which, in turn, matches the hardware-based experimental results. The importance sampling scheme uses a mean-shifted version of the original Gaussian density, appropriately centered between a codeword and a dominant absorbing set, to produce an unbiased estimator of the FER with substantial computational savings over a standard Monte Carlo estimator. Our deterministic estimates are guaranteed to be a lower bound to the error probability in the high SNR regime, and extend the prediction of the error probability to as low as 10^{-30}. By adopting a channel-independent viewpoint, the usefulness of these results is demonstrated for both the standard Gaussian channel and a channel with mixture noise.

Index Terms—LDPC codes; belief propagation; hardware emulation; error floor; importance sampling; near-codeword; trapping set; absorbing set; pseudocodeword.

I. INTRODUCTION

THE CLASS of low-density parity check (LDPC) codes was first introduced by Gallager [14], and has been the focus of intensive study over the past decade (e.g., [18], [22], [23]). An attractive feature of these codes is their outstanding error-correction performance, even when decoded using practical iterative algorithms; in particular, the performance of suitably designed LDPC codes of sufficiently large blocklength is known to be very close to Shannon limits [22]. Various analytical techniques, including density evolution [22] and EXIT charts [26], have been developed for predicting the performance of iteratively decoded LDPC codes. Based on asymptotic approximations, these methods are very accurate for large blocklengths. However, for moderate blocklengths—i.e., those on the order of hundreds to thousands—these methods can yield inaccurate results, and thus there remain

various open questions regarding the performance of specific finite-length LDPC codes.

A particular issue with structured LDPC codes of moderate blocklength is the presence of an *error floor*—that is, a significant flattening in the curve that relates signal-to-noise ratio (SNR) to the frame error rate (FER). Many coding applications—among them satellite communications [1], Ethernet transmission [2], and data storage applications [27]—require very low error rates, so that an important problem is the development of practical tools for predicting error floors and evaluating the performance of LDPC codes in the low frame error rate region.

The main contribution of this paper is the development of a method for predicting and understanding low error rate performance for structured LDPC codes. We describe the method, which exploits a channel-independent perspective on the decoding process to establish deterministic bounds on the error rate. The results are shown to be a lower bound to those obtained through a stochastic method based on importance sampling that produces quick yet accurate estimates of the low probability of error. The error floor is commonly attributed to the suboptimality of the iterative decoding algorithms on graphs with cycles, and past work has studied concepts such as near-codewords [19], trapping sets [21], pseudocodewords [13], and elementary trapping sets [16]. In our own previous work, we have introduced the notion of *(fully) absorbing sets* as the main cause of the error floor of structured LDPC codes. These absorbing sets are a specific type of near-codewords [19] or trapping sets [21], and in particular fully absorbing sets are stable under bit-flipping operations. An absorbing set is a combinatorial object associated with a code, defined independently of the particular decoding scheme or channel noise model. Consequently, the structural properties of absorbing sets can be studied analytically, and their cardinalities can be computed in closed form for certain structured LDPC codes [9]. It can be shown [8] that the factor graphs associated with certain structured LDPC codes contain absorbing sets which have strictly fewer bits than the minimum codeword weight.

The performance of an iterative decoding algorithm in the low FER region is predominantly dictated by the number and the structure of the smallest (fully) absorbing sets, in contrast to the performance of a maximum-likelihood decoder, which is governed by the minimum distance codewords.

In early work on error floors, Richardson [21] developed a fast-simulation method, based on using simulation traces of a hardware emulator to extract trapping set candidates, and then using an approximate integration technique to estimate the associated error probability. This method involves compu-

tations for a sequence of possible channel noise models, which interpolates along a curve between the actual channel noise and noise centered at the trapping set. Subsequent follow-up work [5], [7], [29], [30] has applied importance sampling, which involves drawing samples from a channel noise distribution suitably "biased" towards a candidate trapping set.

The analytic method for approximating the error probability yields lower bounds for sufficiently high SNR. This method is based on approximating the *absorbing region*—namely, the set of algorithm inputs that lead to convergence to a given absorbing set. This absorbing region is defined in terms of the code and the decoder, but does not depend on the particular channel model. The channel model enters only in assessing a type of "distance" to the absorbing region, which we do with Chernoff-type error bounds. Once the absorbing regions have been approximated, our method is easily applied to any channel (without any further substantial computation), as we illustrate by generating predictions for both the standard additive white Gaussian noise (AWGN) channel and also a non-standard Gaussian-mixture channel. The Gaussian-mixture channel is practically important as well, as it is commonly used to model noise in a channel with fading [3], [20], [28]. The results are compared to the method based on mean-shifted importance sampling, suitably applied to absorbing sets of a structured LDPC code [10].

We consider the performance of various structured LDPC codes, different quantized forms of sum-product decoding, as well as different channel models. Section II of this paper contains the relevant background on LDPC codes, absorbing sets, and iterative decoding. The analytic method which produces guaranteed lower bounds on the FER in the high SNR regime is presented in Section III. The channel-independent set-up enables the performance prediction for different channel models without a need to recompute the relevant distance parameters for each application. A brief discussion of the stochastic, importance sampling based method we previously developed in [10], [17] is given in Section IV. The experimental results based on our bounding technique are presented in Section V. The comparison with both the experimental results obtained on a hardware emulator as well as the stochastic approach show excellent agreement. In Section VI we conclude the paper and propose future extensions of the work presented here.

II. BACKGROUND

A. Low-density parity check codes and absorbing sets

Letting $H \in \{0,1\}^{m \times n}$ be a parity check matrix; it defines a binary code of blocklength n, corresponding to the set of all strings $x \in \{0,1\}^n$ that satisfy the relation $Hx = 0$ in modulo two arithmetic. The term *low-density* refers to parity check matrices in which the number of 1s per row and column remains bounded, independent of the blocklength n and number of checks m. A convenient graphical representation of any low-density parity check (LDPC) code is in terms of its associated *Tanner graph* [25] or *factor graph* [15]: it is a bipartite graph $G = (V, F, E)$ in which the subset $V = \{1, \ldots, n\}$ indexes the code bits, the subset $F = \{1, \ldots, m\}$ indexes the code checks, and the edges $E = \{(j,i) \mid H(j,i) = 1\}$ join checks $j \in F$ to bits $i \in V$.
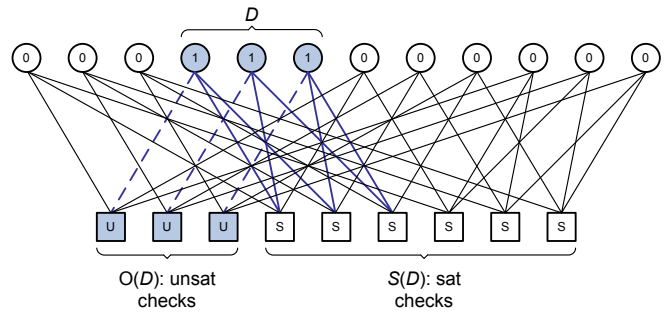

Fig. 1.   An example of a (3,3) absorbing set.

Next we define the notion of an *absorbing set* associated with an LDPC code [8], [31]. For a subset D of V, let $\mathbb{O}(D)$ be the set of neighboring vertices of D in F in the graph G with odd degree with respect to D. Given an integer pair (a, b), an (a, b) *absorbing set* is a subset D of V of size a, with $\mathbb{O}(D)$ of size b and with the property that each element of D has strictly fewer neighbors in $\mathbb{O}(D)$ than in $F \backslash \mathbb{O}(D)$. We say that an (a, b) absorbing set D is an (a, b) *fully absorbing set*, if in addition, all bit nodes in $V \backslash D$ have strictly more neighbors in $F \backslash \mathbb{O}(D)$ than in $\mathbb{O}(D)$ [8]. Therefore fully absorbing sets are stable under the bit-flipping operations. An example of an (a, b) fully absorbing set with $a = 3$, $b = 3$ is given in Fig. 1. Notice that all bit nodes have strictly more satisfied than unsatisfied checks. In the remainder of the paper, all of the discussed absorbing sets are in fact fully absorbing sets, and unless otherwise noted, this subclass will be implicitly assumed.

The notion of the absorbing set is being used in this work to disambiguate it from the definitions of objects used for describing the error floors. The original definition of the trapping set by Richardson [21] is semi-empirical and decoder-dependent. Subsequent work offered an alternative definition of a trapping set as a fixed point of the decoder [6]. A related structure is an (a, b) elementary trapping set [7], [16], which is defined as a trapping set for which all check nodes in the induced subgraph have either degree one or two, and there are exactly b degree-one check nodes. In contrast, the absorbing set is defined as a combinatorial object, and is decoder-independent. An absorbing set can be understood as a special type of a trapping set [21], in which each variable node is connected to strictly more satisfied than unsatisfied checks. In contrast to an elementary trapping set, an absorbing set does not impose any degree constraint on check nodes. We utilize this specific definition of the absorbing set in the body of this work.

Array-based LDPC code constructions [12] are a representative class of high-performing structured LDPC codes. These codes have subsequently been proposed for a number of applications, including digital subscriber lines [11] and magnetic recording [27]. It is known [8] that these codes have absorbing sets that are strictly smaller in size than the minimum distance of the code; moreover, results from hardware emulation show that their low FER performance and the error floor are indeed dominated by these absorbing sets [31]. These codes will be used in the remainder of the paper to conveniently illustrate the methodology developed here.

B. Iterative decoding

The analysis of this paper focuses on the sum-product decoding algorithm, and binary-phase-shift-keyed (BPSK) signalling (under the mapping $0 \mapsto 1$ and $1 \mapsto -1$). As with other iterative algorithms, the sum-product algorithm relies on the exchange of messages between bit nodes and check nodes to achieve correct bit decisions. Suppose the Tanner graph consists of n bit nodes and m check nodes. In the first step, bit nodes x_i, $i = 1, 2, \ldots n$, are initialized with the prior log likelihood ratios given in (1) using the channel outputs y_i, $i = 1, 2, \ldots, n$

$$\ell_i = \log \frac{\mathbb{P}(x_i = 0 | y_i)}{\mathbb{P}(x_i = 1 | y_i)} = \frac{2y_i}{\sigma^2}, \qquad (1)$$

where σ denotes the standard deviation of noise in this Gaussian channel. Bit nodes first send the prior LLR messages to the neighboring check nodes along the edges of the Tanner graph, and the subsequent message exchange is governed by the bit-to-check message $Q_{i \to j}$ and the check-to-bit message $R_{j \to i}$, where $N(i)$ refers to the neighborhood of the node i, where

$$Q_{i \to j} = \ell_i + \sum_{k \in N(i) \setminus j} R_{k \to i} \quad \text{and,}$$

$$R_{j \to i} = \prod_{l \in N(j) \setminus i} \mathrm{sgn}(Q_{l \to j}) \Phi^{-1} \left(\sum_{l \in N(j) \setminus i} \Phi(|Q_{l \to j}|) \right), \qquad (2)$$

where $\Phi(x) := -\log \left[\tanh \left(\frac{x}{2} \right) \right]$ for $x \geq 0$. The posterior log-likelihood ratio at each bit node is then computed as

$$LLR_i^{\mathrm{post}} = \ell_i + \sum_{j \in N(i)} R_{j \to i} \qquad (3)$$

The message passing algorithm is typically allowed to run for a fixed number of iterations, both because convergence is not guaranteed when many cycles are present, and due to practical (delay) constraints. Based on the posterior LLR, a bit-wise hard decision is made: '0' if $LLR_i^{\mathrm{post}} \geq 0$, and '1' otherwise. For practical hardware implementations, the real-valued messages in (2)–(3) are necessarily quantized, and we present results for various quantization schemes.

In terms of channels, we consider additive noise models, in that the channel input $x_i \in \{-1, 1\}$, corresponding to the ith bit in the transmitted codeword, is received as $Y_i = x_i + W_i$, where W_i is observation noise. We consider only memoryless channels meaning that W_i and W_j are independent for $i \neq j$. In the additive white Gaussian noise (AWGN) channel, the noise W_i is a zero-mean Gaussian signal with variance σ^2. For the Gaussian mixture model, the noise takes the form $W_i = U_i V_i$ where $V_i \sim N(0, \sigma^2)$, and U_i is a binary variable, taking value 1 with probability r, and value $a > 1$ with probability $1 - r$.

III. DETERMINISTIC BOUNDS ON ERROR PROBABILITIES

Well-designed LDPC codes of moderate blocklength can yield excellent performance when decoded with suboptimal iterative message-passing algorithms. Due to analytic intractability, the performance of an iteratively-decoded LDPC code is typically reported as the (empirical) probability of error for a certain SNR value, where the total number of decoding errors over a set of trials is counted. For high SNR values, the probability of error is very small, so that a very large number

of trials needs to be run in order to obtain reliable estimates. This explosion in complexity renders such naïve Monte-Carlo approaches unfeasible for estimating low probabilities of error. Instead, modified Monte-Carlo methods using variants of importance sampling (IS) have been used [5], [7], [10], [29], [30]. Applying a fast simulation method requires substantially less computation than direct simulation, but the computation must be re-performed each time that the channel parameters are changed.

In this section, we describe an alternative analytical procedure that provides deterministic lower bounds on the error probability, and has much lower computational cost than even fast simulation techniques like importance sampling. We begin in Section III-A by defining the *absorbing region* associated with any given absorbing set and a particular decoder. Although the nature of this set depends strongly on the decoder (e.g., quantization, saturation levels, etc.), it is independent of the channel noise model. In general, the absorbing region lies in $\mathbb{R}^n$, where n is the blocklength of the code. In Section III-B, we describe a low-dimensional approximation to the absorbing region that is easily computed, and illustrate it for different absorbing sets and decoders.

A. Absorbing regions for decoders

Consider some fixed decoder (e.g., floating-point sum-product, quantized sum-product, or a bit-flipping decoder) that operates on an LDPC code of blocklength n. On any given trial, the decoder is initialized with some vector $\ell \in \mathbb{R}^n$, corresponding to the log-likelihood ratios assumed at each bit node. After each iteration (up to some maximum number), the estimated LLRs of the decoder are thresholded, yielding a $\{0, 1\}^n$ sequence that is an estimate of the transmitted codeword. Accordingly, any decoder is characterized in terms of its quantization levels, saturation points, and maximum number of iterations. Given any such decoder, the associated *absorbing region* $\mathcal{R}(A)$ of a given absorbing set A is the set of input vectors $\ell \in \mathbb{R}^n$ for which the decoder outputs the indicator vector of the absorbing set as its estimate within the maximum number of iterations.

Two properties of this absorbing region are important. First, it is a *channel-independent* quantity, since it is only a function of the initializing LLR vector $\ell \in \mathbb{R}^n$. Although we frequently model the likelihood ratio L as being drawn from a probabilistic channel, when conditioned on a particular initialization $L = \ell$, the decoder's behavior is purely deterministic, and hence channel-independent. Second, it varies as features of the decoder—quantization schemes, number of iterations, etc.—are changed. Indeed, the relative size of the absorbing region is a measure of its impact on a particular decoder.

B. Approximations of absorbing regions

Exact computation of the absorbing region is, unfortunately, prohibitively expensive, since it involves testing the decoder over an n-dimensional space. For instance, discretizing each dimension to m locations yields the complexity $\mathcal{O}(m^n)$. In practice, we are forced to seek approximations to the absorbing region. Here we describe a particular low-dimensional approximation to the absorbing region; later, we describe how it can be exploited to obtain rigorous lower bounds on the absorbing probability.

To approximate the exact absorbing region, we first divide the bit nodes of the absorbing set into groups of nodes with the same number of satisfied and unsatisfied checks when all bit nodes in the absorbing set are incorrect. Here and in the remainder, the notation $(s : u)$ indicates that a bit node in an absorbing set is connected to s satisfied and u unsatisfied checks, and we then say that the bit node is of type $(s : u)$. For example, the $(8, 6)$ absorbing set in the $(2209, 1978)$ array code, shown in Fig. 2b, is made up of six variable nodes of type $(4 : 1)$ and two variable nodes of type $(5 : 0)$. Each of these groups of nodes is associated with an axis. All bit nodes are initially assigned value '1' (the all-zeros codeword under BPSK). In the two-dimensional case, the region is found by trying all combinations of shifts for the two groups of nodes, where each axis ranges from the absorbing set (centered at -1, which corresponds to a shift of -2) to the all-zeros codeword (centered at $+1$, which corresponds to no shift), separated by increments of size 0.01. The decoder is run for 50 iterations, and if the hard decision at the end of these iterations is the absorbing set, then we include this combination of shifts of the absorbing set nodes in a set S. The approximation to the projection of the absorbing region, discretized to this level of granularity, is this set S of points that decode to the absorbing set. Its calculation requires running the decoder a total of $(201)^2$ times, one for each point in the discretization.

Figure 3 illustrates some of the approximations of the projections of the absorbing regions, in the sense above, for different decoders and different absorbing sets. Each panel is a two-dimensional plot, with the upper right $(+1, +1)$ point corresponding to receiving the all-zeros codeword (without any noise), and the lower left $(-1, -1)$ centered on the absorbing set. The marked contours correspond to the boundary between not decoding to and decoding to the appropriate absorbing set (towards lower left). Panel (a) shows regions for the $(4, 8)$ absorbing set, which has the structure shown in Fig. 2a, taken from the Tanner graph of the $(2209, 1978)$ array-based LDPC code, for three different quantized forms of sum-product (details of the quantization choices are in [31]). An $(x.y)$ fixed-point quantization scheme uses x bits to represent the integer portion of the number and y bits to represent the fractional portion. The four variable nodes in the $(4, 8)$ absorbing set are all of the type $(3 : 2)$, so the four nodes are randomly divided into two pairs in order to show a two-dimensional plot, which highlights the symmetry of variable nodes of the same type and thus supports the method of grouping nodes of one type into one axis. Note how the size of the approximation to the projection of the absorbing region shrinks as the quantization is improved.

The effect of better quantization is also seen in panel (b), which shows the approximations of the projection of the absorbing region for the $(8, 6)$ absorbing set of the same code under three different quantization schemes. The effect of a finer quantization scheme is more pronounced when the bits in the absorbing set have only marginally more neighboring satisfied versus unsatisfied checks since the additional bits used to represent the messages can more easily help favorable messages overpower the unfavorable messages, as is the case for the $(4, 8)$ absorbing set.

C. Lower bounds on absorbing probabilities

Although the absorbing region is a channel-independent quantity, the probability of falling within it depends strongly on the channel. We now discuss a method for obtaining lower bounds on this probability, one which uses the approximate absorbing regions defined in the previous section. At first, these approximate regions may seem like poor surrogates to the full n-dimensional absorbing region, since they only consider a very limited number of bits. However, as we prove here, these approximate regions capture the decay in error probability as the SNR increases.

We begin by defining an alternative "channel" under which the probability of landing in the full absorbing region is the same as landing in the approximate absorbing region. In particular, for a fixed absorbing set A, let $\mathbb{Q}_A$ denote the joint distribution over received sequences $(Y_1, \ldots, Y_n)$ in which $Y_i \sim N(1, \sigma^2)$ for all $i \in A$, and $Y_i = 1$ for all $i \notin A$. As before, we let $L = \frac{2}{\sigma^2} Y$ denote the associated log likelihood ratios, corresponding to the decoder input. Our method is based on the fact that it is straightforward to compute $\mathbb{Q}_A[L \in \mathcal{R}(A)]$, since the distribution $\mathbb{Q}_A$ involves randomness only on the absorbing set.

The following result shows for any given class $\mathcal{A}_{a,b}$ of absorbing sets, the probability $\mathbb{Q}_A[L \in \mathcal{R}(A)]$, combined with total number $\mathrm{card}(\mathcal{A}_{a,b})$ of such absorbing sets, can be used to lower bound the error probability. Recalling that we are using quantized algorithms, we let $\mathcal{I}_{\max}$ be the interval of numbers that are quantized to M_q, the maximum quantization level in the decoder.

Theorem 1 (Lower bounds on error probability): For any noise variance σ^2 such that $\frac{2}{\sigma^2} \in \mathcal{I}_{\max}$ and any class $\mathcal{A}_{a,b}$ of absorbing sets, the probability of error under the AWGN channel is lower bounded as

$$\mathbb{P}[\mathcal{E}(L)] \geq (1 - o_\sigma(1)) \, \mathrm{card}(\mathcal{A}_{a,b}) \, \mathbb{Q}_A[L \in \mathcal{R}(A)], (4)$$

where A is any member of the absorbing class $\mathcal{A}_{a,b}$, and $o_\sigma(1)$ denotes a term that converges to zero as $\sigma^2 \to 0$.

Proof: The overall error probability $\mathbb{P}[\mathcal{E}(L)]$ is lower bounded by the probability $\mathbb{P}[\mathcal{E}_\mathcal{A}(L)] = \cup_{A \in \mathcal{A}_{a,b}} \mathbb{P}[L \in \mathcal{R}(A)]$ of decoding incorrectly to some absorbing set $A \in \mathcal{A}$. The events in this union are disjoint, so that we have

$$\begin{aligned}\mathbb{P}[\mathcal{E}(L)] \geq \mathbb{P}[\mathcal{E}_\mathcal{A}(L)] &= \sum_{A \in \mathcal{A}} \mathbb{P}[L \in \mathcal{R}(A)] \\ &= \mathrm{card}(\mathcal{A}_{a,b}) \, \mathbb{P}[L \in \mathcal{R}(A)],\end{aligned}$$
(5)

where the final equality holds since, by exchangeability (invariance to permutations) of the distribution $\mathbb{P}$, the absorption probability is independent of the particular absorbing set considered. Consequently, it suffices to lower bound $\mathbb{P}[L \in \mathcal{R}(A)]$. Next we observe that $\mathbb{Q}_A[L \in \mathcal{R}(A)] = \mathbb{P}[L \in \mathcal{R}(A) \mid L_i \in \mathcal{I}_{\max}, \forall i \notin A]$, which follows from the assumption that $\frac{2}{\sigma^2} \in \mathcal{I}_{\max}$, and $L_i = 2/\sigma^2$ for all $i \notin A$. Consequently, we have the lower bound

$$\begin{aligned}&\mathbb{P}[L \in \mathcal{R}(A)] \\ &\geq \mathbb{P}[L \in \mathcal{R}(A) \mid L_i \in \mathcal{I}_{\max}, \forall i \notin A] \, \mathbb{P}[L_i \in \mathcal{I}_{\max}, \forall i \notin A] \\ &= \mathbb{Q}_A[L \in \mathcal{R}(A)] \, \mathbb{P}[L_i \in \mathcal{I}_{\max}, \forall i \notin A].\end{aligned}$$
(6)

To complete the proof, we need to show that $\mathbb{P}[L_i \in \mathcal{I}_{\max}, \forall i \notin A] = 1 - o_\sigma(1)$ as $\sigma \to 0$. Under the distri-
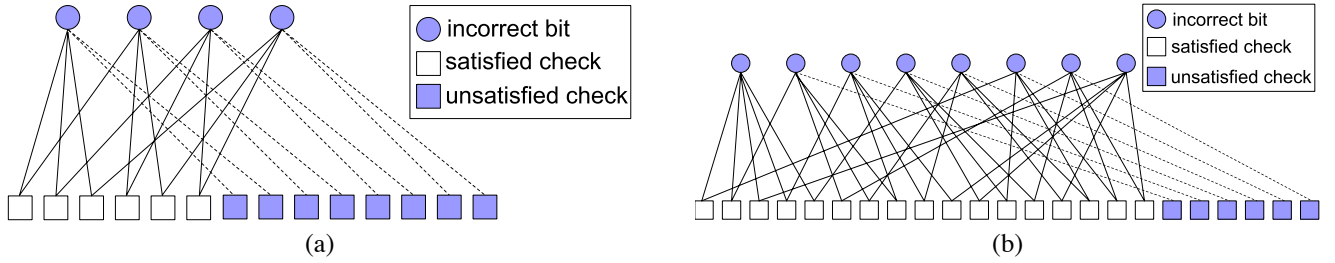
Fig. 2.   These plots show the structure of absorbing sets in the $(2209, 1978)$ array code, where (a) shows a $(4, 8)$ absorbing set, and (b) shows an $(8, 6)$ absorbing set.
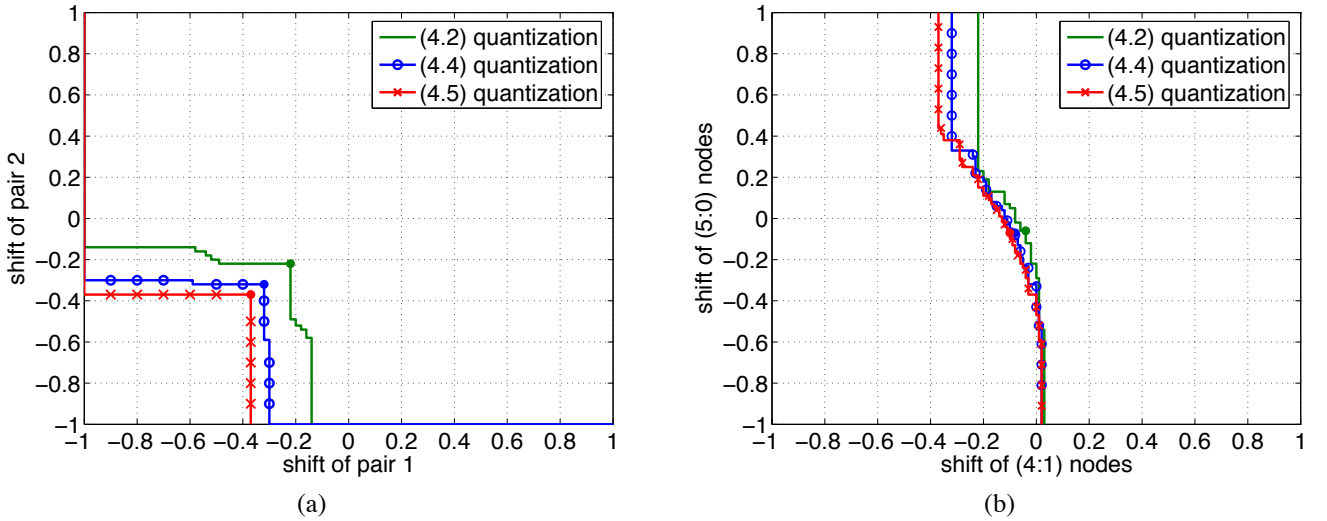

Fig. 3.   Subsets of the signal space regions for which the decoder converges to the absorbing set. These plots show particular projections of these absorbing regions. (a) Absorbing region for a $(4, 8)$ absorbing set in the $(2209, 1978)$ array code. The plot shows three different levels of quantization: (4.2), (4.4), and (4.5) fixed-point quantization [31]. Note how the absorbing region contracts as the quantization scheme improves. (b) Absorbing region for a $(8, 6)$ absorbing set in the $(2209, 1978)$ array code. Here the benefits of improved quantization schemes are only minor.

bution $\mathbb{P}$, the random variable $L_i = \frac{2}{\sigma^2} Y_i$ has a $N(\frac{2}{\sigma^2}, \frac{4}{\sigma^2})$ distribution. Letting M_q be the upper quantization level, we have

$$1 - \mathbb{P}[L_i \in \mathcal{I}_{\max}, \ \forall i \notin A]$$
$$\leq \mathbb{P}[L_i \geq M_q \text{ for some } i \notin A] \ \leq \ (n - |A|) \, \mathbb{P}[L_i \geq M_q],$$
$$\tag{7}$$

where the last step applies union bound. Applying Gaussian tail bounds [24], we conclude that

$$
\begin{aligned}
1 - \mathbb{P}[L_i &\in \mathcal{I}_{\max}, \ \forall i \notin A] \\
&\leq \ 2(n - |A|) \, \exp(-\tfrac{\sigma^2}{8}(\tfrac{2}{\sigma^2} - M_q)^2) \\
&\leq \ 2(n - |A|) \exp(M_q/2) \exp(-\tfrac{1}{2\sigma^2}) \\
&= \ o_\sigma(1)
\end{aligned}
\tag{8}
$$

as claimed.   ∎
Remark: An examination of the proof shows the same result applies more generally to non-Gaussian channels. All that is required is that the probability $\mathbb{P}[L_i \in \mathcal{I}_{\max}, \ \forall i \notin A]$ converge to one as the SNR parameter is taken to infinity. We pursue this idea in application to a time-varying channel later.

We now discuss different methods to approximate or lower bound $\mathbb{Q}_A[L \in \mathcal{R}(A)]$, which generally cannot be easily computed directly since $\mathcal{R}(A)$ may not have a simply defined shape. To find a lower bound for $\mathbb{Q}_A[L \in \mathcal{R}(A)]$, we can find a simple inner bound to the approximate absorbing region and compute the probability of falling into this simple region. For example, for the $(4, 8)$ absorbing set under (4.2) quantization, we see from Fig. 3 that a natural inner bound

$\hat{\mathcal{R}}(A)$ to the approximate absorbing region is a box from $(-.22, -.22)$, the point closest to $(1, 1)$; that is, the region defined by the intersection of the halfspaces $x \leq -.22$ and $y \leq -.22$, where x corresponds to the horizontal axis and y corresponds to the vertical axis. The computation of $\mathbb{Q}_A[L \in \hat{\mathcal{R}}(A)]$ is straightforward, since $\hat{\mathcal{R}}(A)$ is the region where all four bit nodes in the absorbing set have values less than or equal to $-.22$. Therefore, $\mathbb{Q}_A[L \in \hat{\mathcal{R}}(A)] = \mathbb{P}(Y_i \leq -.22)^4$, where $Y_i \sim N(1, \sigma^2)$ represents the value of a bit node in the absorbing set, so in the AWGN case $\mathbb{Q}_A[L \in \hat{\mathcal{R}}(A)] = \Phi(\frac{-.22-1}{\sigma})^4$, where Φ is the CDF of a standard normal random variable. Since $\hat{\mathcal{R}}(A)$ is contained in $\mathcal{R}(A)$, $\mathbb{Q}_A[L \in \hat{\mathcal{R}}(A)] \leq \mathbb{Q}_A[L \in \mathcal{R}(A)]$.

For the approximate absorbing region of the $(8, 6)$ absorbing set, a box would provide only a loose inner bound, and therefore we instead use the intersection $\mathcal{H}$ of two half-spaces. Two lines can form a good inner approximation of the border of the absorbing region, so $\mathcal{H}$ is contained in the absorbing region and closely approximates the absorbing region. Each axis is scaled by the square root of the number of nodes it represents (so for the $(8, 6)$ absorbing set, the x-axis is scaled by $\sqrt{6}$ and the y-axis is scaled by $\sqrt{2}$) to ensure that the minimum distance from the all-zeros codeword to the absorbing region is correctly factoring in the number of nodes grouped to each axis. We then find $\mathbb{Q}_A[L \in \mathcal{H}]$ by numerically integrating the joint pdf of the noise in the channel (two independent Gaussian random variables in the AWGN

case or two independent Gaussian mixture random variables in the mixture noise case) over $\mathcal{H}$. Since $\mathcal{H}$ is an inner bound to $\mathcal{R}(A)$, $\mathbb{Q}_A[L \in \mathcal{H}] \leq \mathbb{Q}_A[L \in \mathcal{R}(A)]$.

To estimate $\mathbb{Q}_A[L \in \mathcal{R}(A)]$ rather than lower bounding the probability, we find the probability of L falling in a simple region that approximates $\mathcal{R}(A)$ (but is not contained in $\mathcal{R}(A)$, which is required in the lower bound case). For example, for the $(8,6)$ absorbing set, we can find a halfspace $\mathcal{H}_{\text{est}} = \{z | a^T z \geq d_{\min}\}$ with $\|a\|^2 = 1$, where $d_{\min}$ is the minimum distance from $(1,1)$ to $\mathcal{R}(A)$. Since this halfspace is an approximation for $\mathcal{R}(A)$, $\mathbb{Q}_A[L \in \mathcal{H}_{\text{est}}] \approx \mathbb{Q}_A[L \in \mathcal{R}(A)]$. For the AWGN channel, $\mathbb{Q}_A[L \in \mathcal{H}_{\text{est}}] = \mathbb{P}(a^T(X_1,...,X_8)^T \geq d_{\min}) = \Phi(\frac{-d_{\min}}{\sigma})$, where $X_i \sim N(0,\sigma^2)$, for $i = 1,...,8$ represents the AWGN noise at the absorbing set bit nodes.

Experimental results demonstrating the tightness of these estimates and bounds for various absorbing sets in different codes under several quantization schemes in the AWGN channel and the mixture channel are shown in Section V.

Moreover, this same approach can also be applied to compute analytical approximations for other channels, which could be used to model time-varying behavior or model uncertainty. Consider a mixture channel, in which the noise on per-symbol basis is either $N(0,\sigma^2)$ with probability $1/2$ ("good" channel state), or $N(0,4\sigma^2)$ with probability $1/2$ ("bad" channel state).

D. A heuristic argument for accuracy of error probability calculations under $\mathbb{Q}_A$

In this section we sketch a heuristic argument for the accuracy of error probability calculations under $\mathbb{Q}_A$.

Let n be the number of bits in the code and $\mathcal{A}$ a class of absorbing sets which dominate the error events when the all-zeros codeword is transmitted. We will assume that no set in $\mathcal{A}$ is contained in another in $\mathcal{A}$. $Y_1,\ldots,Y_n$ are i.i.d. $N(1,\sigma^2)$. $L_1,\ldots,L_n$ are i.i.d. $N(\frac{2}{\sigma^2},\frac{4}{\sigma^2})$. Here $L_i = \frac{2}{\sigma^2}Y_i$. Fix $A \in \mathcal{A}$. We couple the distribution $\mathbb{Q}_A$ to the true noise distribution by defining $\bar{L}_1,\ldots,\bar{L}_n$ as

$$\bar{L}_i = \begin{cases} \frac{2}{\sigma^2} & \text{if } i \notin A \\ L_i & \text{if } i \in A . \end{cases} \quad (9)$$

We will assume that what the sum-product algorithm does with the log-likelihood input $(l_1,\ldots,l_n)$ that results from the transmission of the all-zeros codeword is to decode to the maximizing vertex of the polytope that is convex hull of the vectors $(1,\ldots,1)$ (corresponding to decoding to the all-zeros codeword) and the vectors $((-1, i \in B), (1, i \notin B))$, $B \in \mathcal{A}$, corresponding respectively to decoding to the absorbing set B. This heuristic is motivated by the assumption that the class $\mathcal{A}$ of absorbing sets dominates the error events.

Under this heuristic, the "true" probability of "error by decoding to the absorbing set A when the all-zeros codeword is transmitted" can be written as

$$\pi := \mathbb{P}(-\sum_{i \in B} L_i + \sum_{i \in B^c} L_i \leq -\sum_{i \in A} L_i + \sum_{i \in A^c} L_i ,$$
$$\forall B \in \mathcal{A},\, B \neq A \text{ and } \sum_i L_i \leq -\sum_{i \in A} L_i + \sum_{i \in A^c} L_i)$$
$$= \mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \sum_{i \in B \cap A^c} L_i ,$$
$$\forall B \in \mathcal{A},\, B \neq A \text{ and } \sum_{i \in A} L_i \leq 0)$$
$$= \mathbb{P}(\sum_{i \in A} L_i \leq 0)\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \sum_{i \in B \cap A^c} L_i ,$$
$$\forall B \in \mathcal{A},\, B \neq A \mid \sum_{i \in A} L_i \leq 0) . \quad (10)$$

Similarly, writing $\bar{\pi}$ for the "true" probability of "error by decoding to A when the all zeros codeword is transmitted over A and the log-likelihoods outside A are pinned to $\frac{2}{\sigma^2}$", which would be the probability of such error under $\mathbb{Q}_A$, we have

$$\bar{\pi} = \mathbb{P}(\sum_{i \in A} L_i \leq 0) \times$$
$$\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \frac{2}{\sigma^2} \mid B \cap A^c \mid$$
$$\text{for all } B \in \mathcal{A},\, B \neq A \mid \sum_{i \in A} L_i \leq 0) . \quad (11)$$

Note that we continue to write $\mathbb{P}$ for probabilities because we have coupled $\mathbb{Q}_A$ to $\mathbb{P}$ by pinning down the log-likelihoods outside A.

We now argue that in each of the expressions for π and $\bar{\pi}$ above the conditional probability asymptotically approaches 1 as $\sigma^2 \to 0$. This shows that $\bar{\pi}$ can serve as an upper bound to π within any desired factor $1 + \delta$. It also asymptotically serves as a lower bound for π within any desired constant factor $1 - \delta$.

For $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \sum_{i \in B \cap A^c} L_i$ for all $B \in \mathcal{A}$, $B \neq A \mid \sum_{i \in A} L_i \leq 0)$ consider $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \sum_{i \in B \cap A^c} L_i \mid \sum_i L_{i \in A} \leq 0)$ for any $B \in \mathcal{A}$, $B \neq A$. The conditioning is irrelevant for the term $\sum_{i \in B \cap A^c} L_i$ which is Gaussian centered at $\frac{2}{\sigma^2} \mid B \cap A^c \mid$, while the conditioning makes the term $\sum_{i \in B^c \cap A} L_i$ have conditional mean negative[1], so the claim that $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \sum_{i \in B \cap A^c} L_i \mid \sum_{i \in A} L_i \leq 0)$ converges to 1 as $\sigma^2 \to 0$ should only involve routine variance estimates.

For $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \frac{2}{\sigma^2} \mid B \cap A^c \mid$ for all $B \in \mathcal{A}$, $B \neq A \mid \sum_{i \in A} L_i \leq 0)$ consider $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \frac{2}{\sigma^2} \mid B \cap A^c \mid \mid \sum_{i \in A} L_i \leq 0)$ for any $B \in \mathcal{A}$, $B \neq A$. By assumption $\mid B \cap A^c \mid > 0$, and we again have that the conditioning makes the term $\sum_{i \in B^c \cap A} L_i$ have conditional mean negative, so the claim that $\mathbb{P}(\sum_{i \in B^c \cap A} L_i \leq \frac{2}{\sigma^2} \mid B \cap A^c \mid \mid \sum_{i \in A} L_i \leq 0)$ converges to 1 as $\sigma^2 \to 0$ should again only involve routine variance estimates.

IV. COMPARISON WITH STOCHASTIC SIMULATION

A. Importance Sampling

Importance sampling (IS) is a particular type of Monte Carlo method which uses statistical sampling to approximate analytic expressions of probabilities. The basic idea is to perform simulation under a tilted distribution so as to make the event of interest more likely and hence reduce the computational cost; the averages are then re-weighted to compensate for the tilting. Supposing without loss of generality that the all-zeros codeword is transmitted, let $Y^{(1)},\ldots,Y^{(M)}$ be a set of M trials, each $Y^{(i)} \in \mathbb{R}^n$ sampled in an i.i.d. manner from a biased distribution f_{bias}. Consider a particular absorbing set $A_{a,b}$ of type (a,b) and let $\mathcal{R}(A)$ be its associated absorbing region—that is, the set of decoder inputs $Y \in \mathbb{R}^n$ for which the decoder converges to the absorbing set. The associated IS estimate of the absorbing probability $p(A_{a,b}) := \mathbb{P}[Y \in \mathcal{R}(A_{a,b})]$ is given by

$$\hat{p}_{\text{IS}}(A_{a,b}) := \frac{1}{M}\sum_{i=1}^{M} \mathbb{I}_{\text{err}}(Y^{(i)} \in \mathcal{R}(A_{a,b}))w(Y^{(i)}), \quad (12)$$

[1]Each conditional expectation $E[L_i | \sum_{i \in A} L_i]$ equals $\frac{1}{|A|}\sum_{i \in A} L_i$ so for any B: $E[\sum_{i \in B^c \cap A} L_i | \sum_{i \in A} L_i]$ equals $\frac{|B^c \cap A|}{|A|}\sum_{i \in A} L_i$.

where $\mathbb{I}_{err}$ is a $0-1$-valued indicator function for whether the decoder converges to the given absorbing set on trial $Y^{(i)}$, and $w(Y^{(i)}) = \frac{f(Y^{(i)})}{f_{bias}(Y^{(i)})}$ is the appropriate weighting function to produce an unbiased estimate [4].

B. Estimates by biasing towards absorbing sets

Suppose that we have a procedure for generating an estimate $\widehat{p}_{IS}(A_{a,b})$ of the probability of error associated with a particular (a,b) absorbing set. Due to the symmetry of the code and channel, the probability of error of any fixed (a,b) absorbing set $A_{a,b}$ is equal to that of any other exemplar having the same absorbing set structure. Since the associated events are disjoint, the error probability $p(\text{all } A_{a,b})$ associated with *all* (a,b) absorbing sets of the given structure is equal to $p(\text{all } A_{a,b}) = \text{card}(A_{a,b}) p(A_{a,b})$ where $\text{card}(A_{a,b})$ is the total number of (a,b) absorbing sets of the same structure, and $p(A_{a,b})$ is the probability of the decoder converging to any single absorbing set of that structure. Using this decomposition, the associated IS estimate of $p(\text{all } A_{a,b})$ is given by

$$\widehat{p}_{IS}(\text{all } A_{a,b}) := \text{card}(A_{a,b}) \, \widehat{p}_{IS}(A_{a,b}) \qquad (13)$$

where $\widehat{p}_{IS}(A_{a,b})$ is the IS estimate of $p(A_{a,b})$ from equation (12). For array-based LDPC codes, the total number of (a,b) absorbing sets of a given structure, $\text{card}(A_{a,b})$, can be found using analytical methods [8]. The basic idea in [8] is to use the structure of the parity check matrix to establish a system of equations that reflects the relationship among the nodes in the absorbing set and their neighboring check nodes. The set of solutions to this system of equations then produces the total count of the absorbing sets of a particular type. In codes with different or limited structure, absorbing sets have been identified through hardware emulation [21].

The final step is to note that in general, the event of error on any absorbing set of any type is equal to a disjoint union over all (a,b) of all possible types (a,b) of absorbing sets—that is, the overall error probability can be written as $p = \sum_{(a,b)} p(\text{all } A_{a,b})$, where the sum ranges over all integer pairs (a,b) that lead to absorbing sets, and for each such (a,b) ranges over all types of (a,b) absorbing sets. This decomposition leads to the final IS-based estimate of the overall probability of error:

$$\widehat{p}_{IS} = \sum_{a,b} \widehat{p}_{IS}(\text{all } A_{a,b}) = \sum_{a,b} \text{card}(A_{a,b}) \, \widehat{p}_{IS}(A_{a,b}). \quad (14)$$

In certain cases, the error floor is dominated by a particular isomorphic sub-class of (a^*, b^*) absorbing sets, so that the overall probability of error is dominated by the contribution $p(\text{all } A_{a^*,b^*})$. In other cases, it is necessary to take into account more than one absorbing set class.
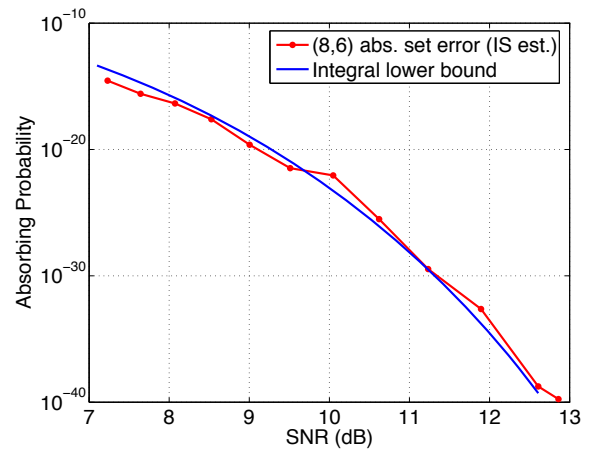
In the case of the all-zeros codeword being transmitted in a BPSK-modulated Gaussian channel, the original density f is an n-variate Gaussian $N(\vec{1}_n, \sigma^2 I_{n\times n})$. A suitable choice for f_{bias} is the mean-shifted Gaussian $N(\nu(\mu), \sigma^2 I_{n\times n})$, where $\nu(\mu)_k = 1 - \mu$ for elements inside the absorbing set, and $\nu(\mu)_k = 1$ otherwise. With this choice, the IS weight w is
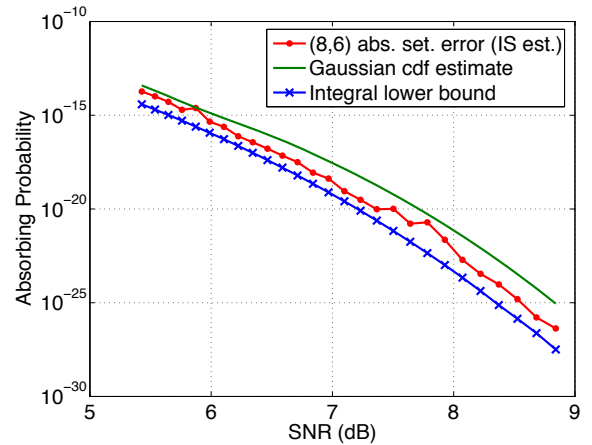
(a)
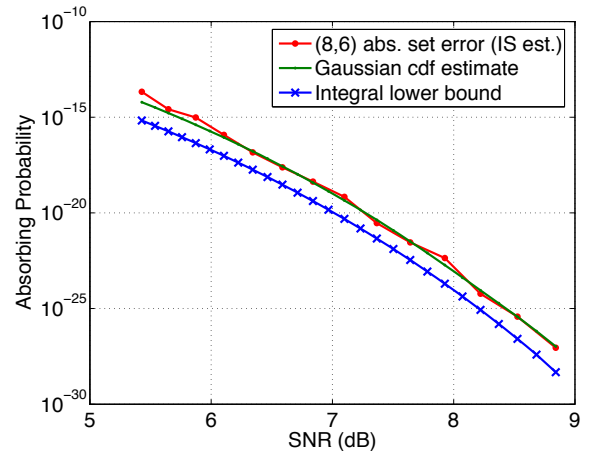
(b)

(c)

Fig. 4. IS result, deterministic estimate, and deterministic lower bound of abs. probability for the $(8, 6)$ abs. sets of the $(2209, 1978)$ array-based LDPC code in the standard AWGN channel under (a) (4.2) fixed-point quantization (b) (4.4) fixed-point quantization. In (c), IS result and deterministic lower bound of abs. probability for the $(8, 6)$ abs. sets of the same code in a mixture channel (noise given by $N(0, \sigma^2)$ with probability $\frac{1}{2}$, and $N(0, 4\sigma^2)$ with probability $\frac{1}{2}$).

given by

$$w(Y^{(i)}|\mu, \sigma^2) = \frac{\exp(-\frac{1}{2\sigma^2}[\sum_{k=1}^{a}(Y_k^{(i)} - 1)^2])}{\exp(-\frac{1}{2\sigma^2}[\sum_{k=1}^{a}(Y_k^{(i)} - (1-\mu))^2])}, \qquad (15)$$

where in the notation $Y_k^{(i)}$, i denotes the index of the sampling

$$w(Y^{(i)}|\mu, \sigma^2) =$$

$$\prod_{k \in V(A)} \frac{r \exp\left(\frac{-(Y_k^{(i)}-1)^2}{2\sigma^2}\right) + \frac{1-r}{a} \exp\left(\frac{-(Y_k^{(i)}-1)^2}{2a^2\sigma^2}\right)}{r \exp\left(\frac{-(Y_k^{(i)}-(1-\mu))^2}{2\sigma^2}\right) + \frac{1-r}{a} \exp\left(\frac{-(Y_k^{(i)}-(1-\mu))^2}{2a^2\sigma^2}\right)} \tag{16}$$

run and k ranges over an enumeration of the nodes in the absorbing set. Likewise, for the noise associated with the mixture Gaussian channel, cf. Subsection II-B, the importance sampling weight is (16) where $V(A)$ is the set of nodes in the absorbing set A.

The paper [10] provides further discussion on different choices for biasing densities, including variance scaling and various mean-shift choices.

In order to evaluate the accuracy of the importance sampling error curves, we calculate 95% confidence intervals for points on the curves. The standard scale for comparing FER curves is $\log_{10}(\widehat{p}_{\text{IS}})$ and $\log_{10}(p)$ versus SNR; accordingly, we first find the variance of $\log_{10}(\widehat{p}_{\text{IS}})$. The variance of the estimator $\widehat{p}_{\text{IS}}$ is given by $\text{var}(\widehat{p}_{\text{IS}}) = \mathbb{E}[(\widehat{p}_{\text{IS}})^2] - p^2$. A way to compute an approximate confidence interval based on the delta-method [24] and Chebyshev's inequality is developed in [17].

V. EXPERIMENTAL RESULTS

We now present experimental results in the AWGN channel and a mixture channel, comparing the deterministic bounds and estimates with importance sampling and hardware emulation curves, for various absorbing sets, codes, and quantization schemes.

We illustrate the deterministic estimate and the deterministic lower bound obtained by our procedure for $(8,6)$ absorbing sets of the $(2209, 1978)$ array-based LDPC code in the AWGN channel under different quantization schemes in Fig. 4. The figure also gives an example of our lower bound results for a Gaussian mixture channel. The deterministic estimates show close agreement with the IS curves and, as predicted theoretically, the deterministic bounds are lower bounds for the IS curves.

In Fig. 5, we show the deterministic lower bound described previously, scaled by the cardinality of the $(4,8)$ absorbing sets in the code, along with the scaled importance sampling results (with the green points representing the 95% confidence interval for each of the points on the importance sampling curve) and the hardware emulation FER curve for both the AWGN channel and a mixture channel. The plots show that for both channels, the deterministic bound appears to lie very close to the importance sampling and hardware results and lower bounds the IS curves for high SNR values. Note that with importance sampling and the deterministic bounds, we are able to extend the error probability curve much farther than with hardware emulation. The deterministic bounds give a vast improvement in computational cost over importance sampling and hardware emulation, as the only step where any substantial computation is required is in finding the absorbing regions, and the time required for this step is only a fraction of the time needed to generate one point on the importance sampling or hardware emulation curves.

We demonstrate our techniques on a different code $((2209, 2024)$ array-based LDPC code) in both the AWGN channel and a Gaussian mixture channel in Fig. 6, which shows the hardware emulation FER curve, the importance sampling results (scaled by the cardinality of the $(6,4)$ absorbing sets in the code, and with the 95% confidence intervals for each point shown by a dotted pair), and a scaled deterministic estimate. The deterministic estimate approximates the importance sampling curve very well and the deterministic bound becomes a lower bound in the high SNR region while still closely following the slope of the importance sampling curve.

VI. CONCLUSION

LDPC codes have recently generated a lot of interest due to their excellent performance. While the infinite blocklength regime is better understood, less is known about the performance of LDPC codes for finite blocklengths. Since the performance of finite blocklength LDPC codes for low FER rates cannot be estimated reasonably fast using software based Monte Carlo simulations, and there is a lack of finite-length theoretical analysis, the deployment of LDPC codes has so far been somewhat limited.

Our method produces deterministic estimates, based on computing projections of absorbing regions and then using asymptotics to estimate associated error probabilities. These deterministic estimates are guaranteed to lower bound the true error probability in the high SNR regime. We have put forth a channel-independent viewpoint which enables efficiently estimating the probability of error for various channel models; the technique is demonstrated for both a pure Gaussian model and a mixture model.

Moreover, our method was compared against the experimental results collected on a hardware emulator as well as experimental results obtained using an importance sampling based approach. The results showed a very close agreement with both sets of experiments, thereby confirming the validity and the computational powers of the proposed technique. An interesting future direction is whether techniques similar to those here can yield matching upper bounds on error probabilities.

REFERENCES

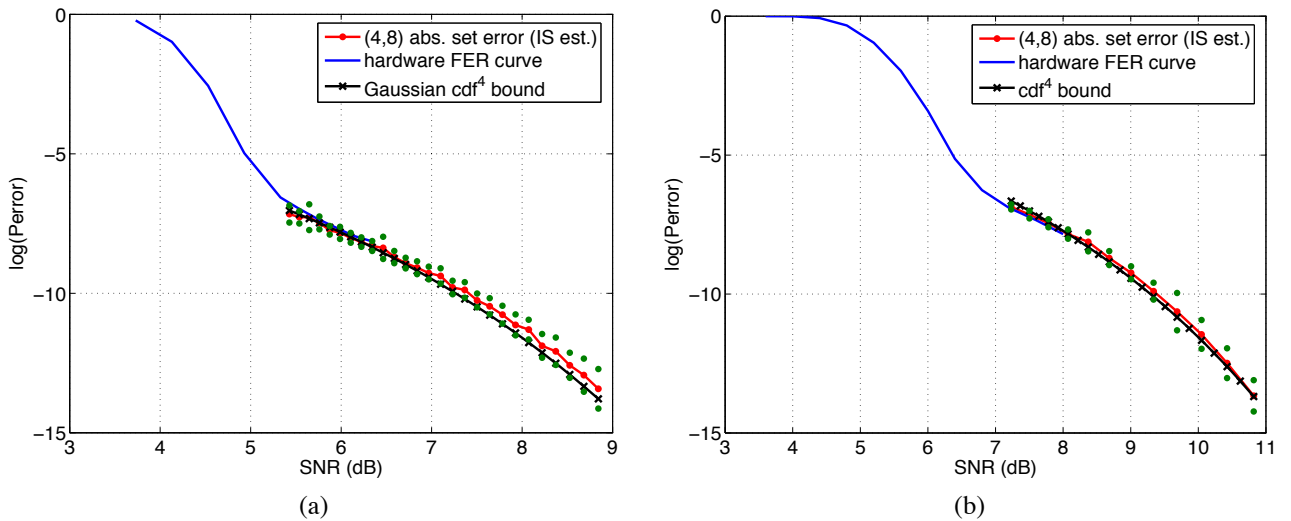[1] Digital Video Broadcasting project. Available at http://www.dvb.org.

Fig. 5. Hardware FER curve, IS result (with the paired points corresponding to approximate 95% confidence interval), and deterministic lower bound of abs. probability for the $(4, 8)$ abs. sets of the $(2209, 1978)$ array-based LDPC code under (4.2) fixed-point quantization in (a) the standard AWGN channel (b) Mixture channel: noise given by $N(0, \sigma^2)$ with probability $\frac{1}{2}$, and $N(0, 4\sigma^2)$ with probability $\frac{1}{2}$.
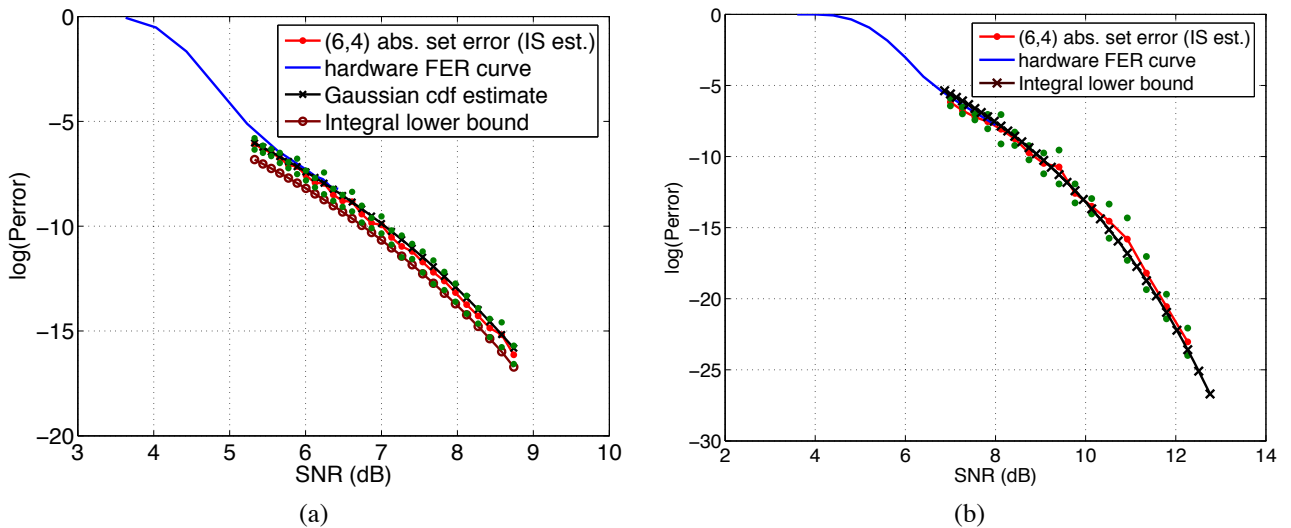

Fig. 6. Hardware FER curve, IS result (with the paired points corresponding to approximate 95% confidence interval), and deterministic estimate of abs. probability for the $(6, 4)$ abs. sets of the $(2209, 2024)$ array-based LDPC code under (4.5) fixed-point quantization in (a) the standard AWGN channel (b) Mixture channel: noise given by $N(0, \sigma^2)$ with probability $\frac{1}{2}$, and $N(0, 4\sigma^2)$ with probability $\frac{1}{2}$.

[2] IEEE Standard 802.3an. Available at http://ieeexplore.ieee.org/servlet/opac?punumber=4039890.

[3] R.S. Blum, R.J. Kozick, and B.M. Sadler. An adaptive spatial diversity receiver for non-Gaussian interference and noise. *IEEE Trans. Signal Processing*, pages 2100–2111, Aug. 1999.

[4] J. Bucklew. *Introduction to Rare Event Simulation*. Springer, 2004.

[5] E. Cavus, C. Haymes, and B. Daneshrad. A highly efficient importance sampling method for performance evaluation of LDPC codes at very low bit errors rates. *Submitted to IEEE Trans. Commun.*

[6] S.K. Chilappagari, S. Sankaranarayanan, and B. Vasic. Error floors of LDPC codes on the binary symmetric channel. In *IEEE Int. Conf. on Comm.*, June 2006.

[7] C. A. Cole, S. G. Wilson, E. K. Hall, and T. R. Giallorenzi. A general method for finding low error rates of LDPC codes. *Submitted to IEEE Trans. Inform. Theory*, 2006.

[8] L. Dolecek, Z. Zhang, V. Anantharam, M. Wainwright, and B. Nikolic. Analysis of absorbing sets of array-based LDPC codes. In *IEEE Int. Conf. Comm.*, June 2007.

[9] L. Dolecek, Z. Zhang, V. Anantharam, M. Wainwright, and B. Nikolic. Analysis of absorbing sets and fully absorbing sets of array-based LDPC codes. *Submitted to IEEE Trans. Inform. Theory*, 2008.

[10] L. Dolecek, Z. Zhang, M. Wainwright, V. Anantharam, and B. Nikolic. Evaluation of the low frame error rate performance of LDPC codes using importance sampling. In *IEEE Info. Theory Workshop*, Sept. 2007.

[11] E. Eleftheriou and S. Ölçer. Low density parity check codes for digital subscriber lines. In *IEEE Int. Conf. on Comm.*, April 2002.

[12] J. Fan. Array-codes as low-density parity-check codes. In *2nd Int. Symp. on Turbo Codes*, Sept. 2000.

[13] B. J. Frey, R. Koetter, and A. Vardy. Skewness and pseudocodewords in iterative decoding. In *IEEE ISIT*, August 1998.

[14] R. G. Gallager. *Low-density parity check codes*. MIT Press, Cambridge, MA, 1963.

[15] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Inform. Theory*, 47(2):498–519, February 2001.

[16] S. Laendner and O. Milenkovic. Algorithmic and combinatorial analysis of trapping sets in structured LDPC codes. In *WirelessComm*, June 2005.

[17] P. Lee, L. Dolecek, Z. Zhang, V. Anantharam, M. Wainwright, and B. Nikolic. Error floors in LDPC codes: fast simulation, bounds and hardware emulation. In *IEEE ISIT*, July 2008.

[18] M. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. Spielman. Improved low-density parity check codes using irregular graphs. *IEEE Trans. Inform. Theory*, 47:585–598, February 2001.

[19] D. MacKay and M. Postol. Weaknesses of Margulis and Ramanujan-Margulis low-density parity-check codes. *Elec. Notes in Theo. Comp. Science*, 74, 2003.

[20] A. Nezampour, A. Nasri, R. Schober, and Y. Ma. Asymptotic BEP and

SEP of differential EGC in correlated Ricean fading and non-Gaussian noise. In *IEEE GLOBECOM*, Nov. 2007.

[21] T. Richardson. Error floors of LDPC codes. In *Allerton*, Oct. 2003.

[22] T. Richardson and R. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Inform. Theory*, 47(2), Feb. 2001.

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

[24] R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Statistics. Wiley, 1980.

[25] R. M. Tanner. A recursive approach to low complexity codes. *IEEE Trans. Inform. Theory*, 27:533–547, September 1980.

[26] S. ten Brink. Convergence of iterative decoding. *Electronics Letters*, 35:806 – 808, May 1999.

[27] B. Vasic and E. Kurtas. *Coding and Signal Processing for Magnetic Recording Systems*. CRC Press, 2005.

[28] X. Wang, G. Yue, and K.R. Narayanan. Optimization of LDPC-coded CDMA systems. *IEEE Trans. Signal Processing*, pages 1500–1510, April 2005.

[29] B. Xia and W. E. Ryan. Estimating LDPC codeword error rates via importance sampling. In *IEEE ISIT*, July 2004.

[30] B. Xia and W. E. Ryan. Importance sampling for Tanner trees. *IEEE Trans. Inform. Theory*, 51:2183 – 2189, June 2005.

[31] Z. Zhang, L. Dolecek, V. Anantharam, and M. Wainwright B. Nikolic. Quantization effects of low-density parity-check decoders. In *IEEE Int. Conf. on Comm.*, June 2007.

Venkat Anantharam received the B.Tech in Electronics in 1980 from the Indian Institute of Technology, Madras (IIT-M) and the M.A. and C.Phil degrees in Mathematics and the M.S. and Ph.D. degrees in Electrical Engineering in 1983, 1984, 1982 and 1986 respectively, from the University of California at Berkeley (UCB). From 1986 to 1994 he was on the faculty of the School of EE at Cornell University. From 1994 he has been on the faculty of the EECS department at UCB. Anantharam received the Philips India Medal and the President of India Gold Medal from IIT-M in 1980, and an NSF Presidential Young Investigator award during the period 1988-1993. He is a co-recipient of the 1998 Prize Paper award of the IEEE Information Theory Society (with S. Verdu) and a co-recipient of the 2000 Stephen O. Rice Prize Paper award of the IEEE Communications Theory Society (with N. Mckeown and J. Walrand). He received the Distinguished Alumnus Award from IIT-M in 2008. He is a Fellow of the IEEE.

Lara Dolecek is a post-doctoral researcher with the Massachusetts Institute of Technology. She holds a B.S., M.S. and Ph.D. degrees in Electrical Engineering and Computer Sciences, as well as an M.A. degree in Statistics, all from the University of California, Berkeley. For her dissertation she received the 2007 David J. Sakrison Memorial Prize for the most outstanding doctoral research in the Department of Electrical Engineering and Computer Sciences at UC Berkeley. Her research interests span information and probability theory, graphical models, statistical algorithms, and computational methods, with applications to complex systems for data processing, communication, and storage.

Pamela Lee received a B.S. degree from the University of California, Berkeley in 2008. She is currently a graduate student at Berkeley in the Department of Electrical Engineering and Computer Sciences. Her research interests include coding and information theory and graphical models.
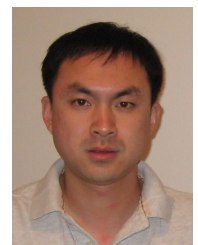
Zhengya Zhang received the B.A.Sc. degree in computer engineering from University of Waterloo, Canada, and the M.S. degree in electrical engineering from University of California, Berkeley. He is currently a Ph.D. candidate in Electrical Engineering at University of California, Berkeley, where he is a member of the Berkeley Wireless Research Center. His research interest is in the design of signal processing and computation systems which require a spectrum of optimizations from algorithm to architecture and implementation. He is the recipient of the Analog Devices Outstanding Student Designer Award and the Vodafone U.S. Foundation Fellowship for his graduate research. He has held multiple internships with Nvidia and Nortel in the past.

Borivoje Nikolic received the Dipl.Ing. and M.Sc. degrees in electrical engineering from the University of Belgrade, Serbia, in 1992 and 1994, respectively, and the Ph.D. degree from the University of California at Davis in 1999. He lectured electronics courses at the University of Belgrade from 1992 to 1996. He spent two years with Silicon Systems, Inc., Texas Instruments Storage Products Group, San Jose, CA, working on disk-drive signal processing electronics. In 1999, he joined the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, where he is now a Professor. His research activities include digital and analog integrated circuit design and VLSI implementation of communications and signal processing algorithms. He is co-author of the book Digital Integrated Circuits: A Design Perspective, 2nd ed, Prentice-Hall, 2003. Dr. Nikolic received the NSF CAREER award in 2003, College of Engineering Best Doctoral Dissertation Prize and Anil K. Jain Prize for the Best Doctoral Dissertation in Electrical and Computer Engineering at University of California at Davis in 1999, as well as the City of Belgrade Award for the Best Diploma Thesis in 1992. For work with his students and colleagues he received the Best Paper Award at the ACM/IEEE International Symposium of Low-Power Electronics in 2005, and the 2004 Jack Kilby Award for the Outstanding Student Paper at the IEEE International Solid-State Circuits Conference.

Martin Wainwright is currently an assistant professor at University of California at Berkeley, with a joint appointment between the Department of Statistics and the Department of Electrical Engineering and Computer Sciences. He received his Ph.D. degree in Electrical Engineering and Computer Science (EECS) from Massachusetts Institute of Technology (MIT) in 2002. His research interests include statistical signal processing, coding and information theory, statistical machine learning, and high-dimensional statistics. He has been awarded an Alfred P. Sloan Foundation Fellowship, an NSF CAREER Award. the George M. Sprowls Prize for his dissertation research (EECS department, MIT), a Natural Sciences and Engineering Research Council of Canada 1967 Fellowship, the IEEE Signal Processing Society Best Paper Award, and several outstanding conference paper awards.