



## **A Refined Scaling Law for Spatially Coupled LDPC Codes Over the Binary Erasure Channel**

Downloaded from: <https://research.chalmers.se>, 2021-08-31 10:49 UTC

Citation for the original published paper (version of record):

Sokolovskii, R., Brännström, F., Graell I Amat, A. (2019)

A Refined Scaling Law for Spatially Coupled LDPC Codes Over the Binary Erasure Channel

2019 IEEE Information Theory Workshop (ITW): 90-94

<http://dx.doi.org/10.1109/ITW44776.2019.8989167>

N.B. When citing this work, cite the original published paper.

# A Refined Scaling Law for Spatially Coupled LDPC Codes Over the Binary Erasure Channel

Roman Sokolovskii, Fredrik Brännström, and Alexandre Graell i Amat

Department of Electrical Engineering, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

**Abstract**—We propose a refined scaling law to predict the finite-length performance in the waterfall region of spatially coupled low-density parity-check codes over the binary erasure channel. In particular, we introduce some improvements to the scaling law proposed by Olmos and Urbanke that result in a better agreement between the predicted and simulated frame error rate. We also show how the scaling law can be extended to predict the bit error rate performance.

## I. INTRODUCTION

Spatially coupled low-density parity-check (SC-LDPC) codes [1], [2] have received a great deal of attention in the last years due to the remarkable threshold saturation effect, i.e., the belief propagation (BP) decoder can achieve the maximum-a-posteriori (MAP) threshold of the underlying uncoupled ensemble. Threshold saturation was first shown in [2] and proved for the binary erasure channel (BEC) in [3]. It was later proved for the general class of binary-input memoryless symmetric channels [4]. Another desirable property of spatial coupling is that it preserves the distance growth properties of the underlying ensemble, i.e., an SC-LDPC ensemble constructed from a regular ensemble preserves the linear growth of the minimum distance with the block length. Spatial coupling is a very general concept, and is not limited to low-density parity-check (LDPC) codes. It has been used, e.g., in the context of turbo [5] and polar [6] codes. Staircase codes [7] can also be viewed as a class of spatially coupled codes.

The BP threshold has become a parameter of substantial importance in modern coding theory, as it allows fast code optimization (e.g., in terms of degree distribution for LDPC codes) via density evolution. Furthermore, a code with better threshold typically also achieves better performance in the waterfall region for finite length. Predicting the probability of error for a fixed code length, however, is also of great practical interest. A finite-length scaling for LDPC code ensembles over the BEC, which yields accurate predictions of the frame error rate (FER) in the waterfall region, was proposed in [8]. The scaling law is based on the analysis of the statistical evolution of the residual graph of the peeling decoder as a function of time. Following this approach, a scaling law was proposed in [9] to predict the FER performance of terminated SC-LDPC code ensembles. It is based on modeling the stochastic process associated with the fraction of degree-one check nodes (CNs) during the peeling decoding by an appropriately chosen Ornstein-Uhlenbeck process. An approximation of the distribution of the first hit time of this process, i.e., of the earliest

time when the fraction of degree-one CNs reaches zero, is used to predict the probability of decoding error. The authors derived a system of coupled differential equations, dubbed *mean and covariance evolution* [8], to estimate the parameters of the Ornstein-Uhlenbeck process. In contrast to the case of block LDPC codes, the results in [9] show a mismatch between the predicted performance and the simulation results, which was attributed in [9] to the approximation of the distribution of the first hit time of the Ornstein-Uhlenbeck process by an exponential distribution.

In this paper, we propose a refined scaling law for terminated SC-LDPC codes over the BEC. In particular, we model the decoding process as two independent Ornstein-Uhlenbeck processes, in correspondence to the two decoding waves that propagate toward the center of the coupled chain for terminated SC-LDPC codes. This is in contrast to the scaling law in [9], which assumes a single process. Accordingly, the probability density function (PDF) of the first hit time can be modeled as the convolution of two exponential PDFs, leading to the PDF of an Erlang distribution, which we use to predict the probability of error. We show that this refined model results in a better agreement with simulations compared to the original scaling law in [9]. We further improve the match between the predicted performance and simulation results by introducing a dependency on the channel parameter of the scaling constants that can be computed from the mean evolution. Finally, we show that the scaling law can be easily adapted to predict the bit error rate (BER).

## II. PRELIMINARIES

We consider the  $(d_v, d_c, L, M)$  SC-LDPC code ensemble introduced in [9], where a sequence of  $L$   $(d_v, d_c)$ -regular LDPC codes of length  $M$ , with variable node (VN) degree  $d_v$  and check node degree  $d_c$ , is coupled according to the following rule: every VN at position  $i$  is connected to one randomly chosen CN at each of the positions in the range  $[i, \dots, i + d_v - 1]$ . The *terminated* ensemble is obtained from such a chain by appending additional  $d_v - 1$  positions containing CNs only. Note that while this coupling rule prescribes VNs to be connected to each of the  $d_v$  different positions in  $[i, \dots, i + d_v - 1]$ , it does not enforce any particular connectivity on the CNs within this range. Thus, a given CN can be connected to VNs at the same position. This “semi-structured” ensemble was chosen in [9] instead of the conventional ensemble with *smoothing* parameter [3] to simplify the analysis. Besides the terminated

ensemble, we also consider the *truncated* ensemble, where the coupled chain is truncated after  $L$  positions.

We consider decoding using the peeling decoding algorithm, which has identical performance to that of BP decoding over the BEC, but makes the analysis more tractable. At the initialization step of the peeling decoding, all non-erased VNs and adjacent edges are removed from the Tanner graph. Next, at each iteration, one degree-one CN is randomly chosen. Since its neighbor VN is known (i.e., the code bit can be recovered), the selected CN can be deleted, as well as the neighbor VN and all its adjacent edges. This results in a sequence of residual graphs indexed by the iteration number  $\ell$ . The decoding process stops when there are no degree-one CNs left, which occurs either because the whole message has been successfully recovered or because decoding fails.

In the waterfall region, the primary contributor to decoding failures are large (linear sized with respect to the code length) stopping sets [8]. Thus, the scaling law aims to estimate the probability of a linear sized number of bits remaining unrecovered after the termination of the peeling decoding.

#### A. Finite-Length Scaling of SC-LDPC Ensembles in [9]

The scaling law in [9] is based on analyzing the first hit time of the stochastic process associated with the fraction of degree-one CNs in the residual graphs during peeling decoding [10],

$$r_1(\tau) \triangleq \frac{1}{M} \sum_u R_{1,u}(\tau), \quad (1)$$

where  $\tau \triangleq \ell/M$  can be viewed as normalized time of the peeling decoding process, and  $R_{1,u}(\tau)$  is the number of degree-one CNs at position  $u$  of the residual graph at iteration  $\ell$ . It was shown in [8], [9] that the distribution of  $r_1(\tau)$  converges to a Gaussian distribution as  $M \rightarrow \infty$ .

The first hit time  $\tau_0$  is defined as the earliest time of the peeling decoding for which the process  $r_1(\tau)$  hits zero, i.e.,

$$\tau_0 \triangleq \min\{\tau : r_1(\tau) = 0\}. \quad (2)$$

We denote the PDF of  $\tau_0$  as  $f_{\tau_0}$ .

The expected number of degree-one CNs,  $\bar{r}_1(\tau) \triangleq \mathbb{E}[r_1(\tau)]$ , where the expectation is taken over the channel and ensemble realizations, exhibits a steady-state phase where it remains essentially constant. We denote the range of  $\tau$  corresponding to the steady state as  $[\alpha, \beta]$ . The underlying assumption of the scaling law in [9] is that the decoding failure is only possible at the steady state, hence the FER is approximated as

$$P_f \approx \int_{\alpha}^{\beta} f_{\tau_0}(x) dx. \quad (3)$$

The decoding process  $r_1(\tau)$  in the steady state is characterized by the following parameters.

- 1) Expectation constant  $\gamma$ . The value of  $\bar{r}_1(\tau)$  at the steady state is approximated by

$$\bar{r}_1(\tau) \approx \gamma(\epsilon^* - \epsilon),$$

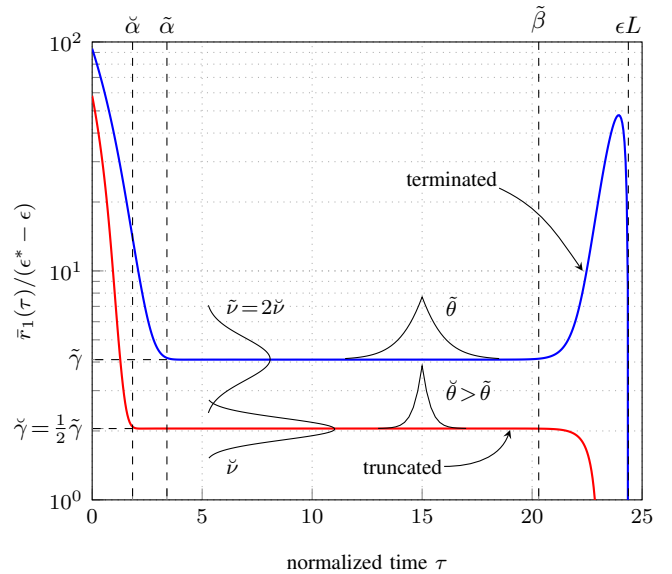


Fig. 1. The evolution of the expected fraction of degree-one CNs  $\bar{r}_1(\tau)$  during peeling decoding, normalized by the distance to the BP threshold, for the (5, 10,  $L=50$ ) ensemble with  $\epsilon^* = 0.4994$  at  $\epsilon = 0.4875$ .

where  $\epsilon$  is the channel erasure probability and  $\epsilon^*$  denotes the BP decoding threshold of the given  $(d_v, d_c, L)$  SC-LDPC ensemble, which can be computed via density evolution.

- 2) Variance constant  $\nu$ . The variance of  $r_1(\tau)$  is also nearly constant at the steady state and is approximated as

$$\text{Var}[r_1(\tau)] \approx \frac{\nu}{M}.$$

- 3) Correlation decay constant  $\theta$ . For two time instants  $\tau, \zeta \in [\alpha, \beta]$ , the covariance of the decoding process along iterations of the peeling decoding is assumed to be

$$\mathbb{E}[r_1(\tau)r_1(\zeta)] - \bar{r}_1(\tau)\bar{r}_1(\zeta) \approx \frac{\nu}{M}e^{-\theta|\zeta-\tau|}.$$

Overall, apart from the BP threshold  $\epsilon^*$ , the scaling law requires the five parameters  $(\alpha, \beta, \gamma, \nu, \theta)$ . The meaning of these parameters is illustrated in Fig. 1 for the (5, 10,  $L=50$ ) SC-LDPC code ensemble at  $\epsilon = 0.4875$ . In the following, we denote the variables associated with the terminated ensembles with a tilde, e.g.,  $\tilde{\gamma}$ , and those associated with the truncated ensembles with a breve, e.g.,  $\breve{\gamma}$ .

In [9], the scaling parameters  $(\tilde{\gamma}, \tilde{\nu}, \tilde{\theta})$  were estimated for a particular value of  $\epsilon$ , namely  $\epsilon = \epsilon^* - 0.04$ , by solving numerically a system of partial differential equations called covariance evolution, adapting the approach proposed in [8] for uncoupled ensembles. A lower bound on  $\tilde{\alpha}$ , denoted as  $\tilde{\alpha}_{\text{LB}}$ , was obtained from the density evolution of the underlying uncoupled  $(d_v, d_c)$  LDPC code ensemble by calculating the fraction of recovered bits at the threshold  $\epsilon^*$ . The end of the steady state was approximated by  $\tilde{\beta} = \epsilon L$ .

The decoding process  $r_1(\tau)$  at the steady state is modeled in [9] by an Ornstein-Uhlenbeck process. Consequently, the distribution of  $\tau_0$  in the steady state is approximated by the

distribution of the first hit time of the corresponding Ornstein-Uhlenbeck process, which converges to an exponential distribution with mean  $\mu_0$  as  $M \rightarrow \infty$ ,

$$\mu_0 = \frac{\sqrt{2\pi}}{\theta} \int_0^{\gamma\sqrt{M/\nu}(\epsilon^* - \epsilon)} \Phi(z) e^{\frac{1}{2}z^2} dz, \quad (4)$$

where  $\Phi(z)$  is the cumulative distribution function (CDF) of the Gaussian distribution. Thus, the PDF of  $\tau_0$  in the steady state is approximated by an exponential PDF with scale parameter  $\mu_0$ , shifted by  $\alpha$  to account for the initial transient period,

$$f_{\tau_0}(x) \approx f_{\tau_0}^{(1)}(x) \triangleq \mu_0^{-1} \exp\left(-\frac{x-\alpha}{\mu_0}\right) H(x-\alpha), \quad (5)$$

where  $H(x)$  is the Heaviside step function.

The FER of a terminated ( $d_v, d_c, L, M$ ) SC-LDPC code ensemble is then estimated by using the approximation (5) in (3), resulting in [9]

$$P_f \approx 1 - \exp\left(-\frac{\beta - \alpha}{\mu_0}\right), \quad (6)$$

with scaling parameters ( $\tilde{\alpha}_{\text{LB}}, \tilde{\beta} = \epsilon L, \tilde{\gamma}, \tilde{\nu}, \tilde{\theta}$ ).

### III. REFINED SCALING LAW

We propose modeling the steady state of the stochastic process  $r_1(\tau)$  for terminated SC-LDPC ensembles as the sum of two identical independent Ornstein-Uhlenbeck processes, to mimic the two decoding waves moving toward the center of the chain that characterize these ensembles. Each Ornstein-Uhlenbeck process is the same as the equivalent process for the truncated ensemble, where only one decoding wave is present. The decoding is successful if the two decoding waves meet, otherwise a decoding failure occurs.

The motivation behind this model, as opposed to the model in [9] based on a single Ornstein-Uhlenbeck process, is provided in the following. In Fig. 2 (top), we depict the simulated CDF of the first hit time of the peeling decoding of the terminated ( $5, 10, L=50, M=2000$ ) SC-LDPC ensemble for  $\epsilon = 0.4875$  (blue curve), and compare it with the simulated CDF of the first hit time of the single Ornstein-Uhlenbeck process with appropriately chosen parameters (green dotted curve). We also plot the CDF of the exponential distribution corresponding to the analytical approximation  $f_{\tau_0}^{(1)}(x)$  (see (5)) proposed in [9] (red dashdotted curve). The corresponding PDFs are shown in Fig. 2 (bottom). The figure shows a significant disagreement between the distributions of the first hit time of the Ornstein-Uhlenbeck process and those of the first hit time of the peeling decoding. Hence, a single Ornstein-Uhlenbeck process appears to be inadequate as a model for  $r_1(\tau)$  in the steady state. In particular, it can be clearly seen in Fig. 2 (bottom) that the simulated distribution of the first hit time of the peeling decoding is not exponential. On the other hand, the agreement between the red dashdotted and green dotted curves indicates that the exponential distribution approximates well the distribution of the first hit time of the Ornstein-Uhlenbeck process.

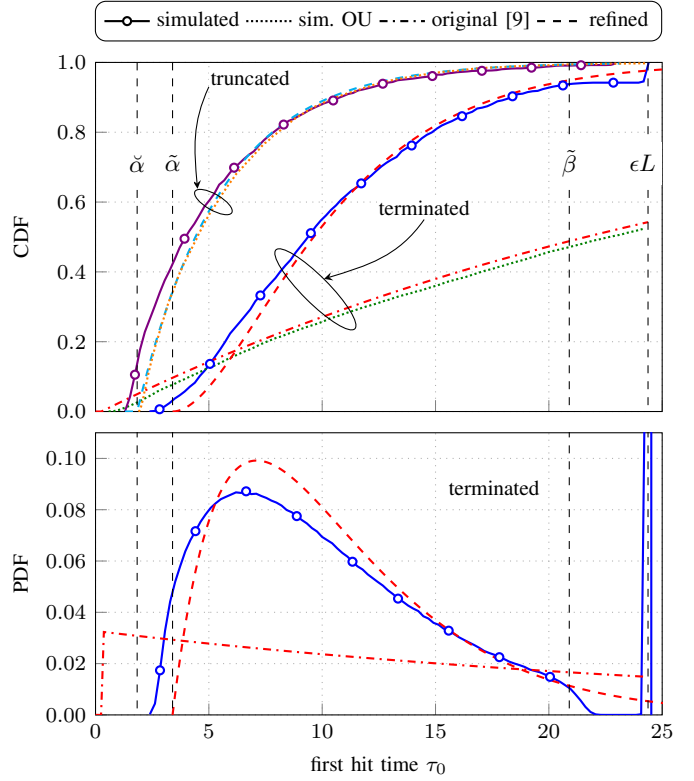


Fig. 2. Comparison of the simulated and approximated CDFs of the first hit time for the terminated and truncated ( $5, 10, L=50, M=2000$ ) ensembles at  $\epsilon = 0.4875$  (top). At the bottom plot, the corresponding PDFs for the terminated ensemble are shown.

In the figure, we also plot the simulated CDF of the first hit time of the peeling decoding (purple curve), the simulated CDF of the first hit time of the Ornstein-Uhlenbeck process (orange dotted curve) and its exponential approximation (cyan dashed curve) for the corresponding truncated ensemble. In this case, a much better match is observed. Hence, a single Ornstein-Uhlenbeck process models well the behavior of the single-wave decoding process but not that of the two-wave decoding process, motivating the proposed model based on two Ornstein-Uhlenbeck processes for the terminated case.

#### A. Decoding Process as Two Independent Ornstein-Uhlenbeck Processes

Since the PDF of the first hit time for the truncated ensemble (featuring a single decoding wave) is well approximated by an exponential, we model the PDF of the first hit time for the terminated ensemble, which is characterized by two decoding waves, as the convolution of two exponential PDFs. This results in an Erlang PDF with shape parameter 2 and scale parameter  $\mu_0$ ,

$$f_{\tau_0}^{(2)}(x) \triangleq \mu_0^{-2} (x - \alpha) \exp\left(-\frac{x - \alpha}{\mu_0}\right) H(x - \alpha), \quad (7)$$

where  $\mu_0$  is given in (4). As in (5), the PDF is shifted by  $\alpha$  so that the initial transient period before the establishment of the steady-state regime is taken into account.

Thus, for the terminated ensemble, we approximate the PDF of  $\tau_0$  in the steady state as

$$f_{\tau_0}(x) \approx f_{\tau_0}^{(2)}(x). \quad (8)$$

Using (7)–(8) in (3), the FER of the terminated SC-LDPC code ensemble can then be approximated as

$$P_f \approx 1 - \left(1 + \frac{\beta - \alpha}{\mu_0}\right) \exp\left(-\frac{\beta - \alpha}{\mu_0}\right). \quad (9)$$

We remark that in the refined scaling law (9), the scaling parameters  $(\gamma, \nu, \theta)$  are those corresponding to a single decoding wave, hence they should be estimated for the truncated ensemble. On the other hand, the time interval  $[\alpha, \beta]$  of the steady state determines the times when the two-wave regime is established and ended, thus  $\alpha$  and  $\beta$  must be estimated from the evolution of  $\bar{r}_1(\tau)$  for the terminated ensemble. To summarize, the FER in (9) should be evaluated using the scaling parameters  $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\nu}, \tilde{\theta})$ .

It is important to emphasize that the authors of [9] mentioned that considering the decoding process as two processes corresponding to the two decoding waves would affect the scaling constants  $\nu, \gamma$ , and  $\theta$ , from which the constants for the combined process (still with the first hit time modeled as an exponential distribution) could be obtained. Here, however, we argue that the separate treatment of the two decoding waves leads not only to a change in the scaling constants, but also to the change of the distribution of the first hit time of the combined process from an exponential to an Erlang distribution.

### B. Dependence of the Scaling Parameters on the Channel Parameter

The scaling law in [9] models the scaling parameters  $\alpha, \gamma, \nu$ , and  $\theta$  as constants independent of the channel parameter  $\epsilon$ . However, their estimation for different values of  $\epsilon$  yields different parameters, which indicates that they are, in fact, dependent on  $\epsilon$ . Thus, it is preferable to calculate these parameters for each  $\epsilon$  separately.

The parameters  $\alpha, \beta$ , and  $\gamma$  depend on the evolution of  $\bar{r}_1(\tau)$ , and can be obtained from the mean evolution in [9]. We remark that the range of the steady state  $[\alpha, \beta]$  depends not only on  $\epsilon$  but also on the length of the coupled chain  $L$ . For each  $(d_v, d_c, L)$  SC-LDPC ensemble, we estimate  $\alpha, \beta$ , and  $\gamma$  from the evolution of  $\bar{r}_1(\tau)$  for a number of channel parameters  $\epsilon$  and obtain the intermediate values by linear interpolation. Introducing the dependence of  $\alpha, \beta$ , and  $\gamma$  on  $\epsilon$  slightly improves the prediction of the FER.

Estimating  $\nu$  and  $\theta$  as a function of  $\epsilon$ , on the other hand, requires numerically solving the full covariance evolution for each  $\epsilon$ , which is significantly more complex than the mean evolution and thus renders the approach infeasible. Therefore, as in [9] we treat the variance parameter  $\nu$  and the covariance decay parameter  $\theta$  as ensemble-dependent constants and use the same values of  $\nu$  and  $\theta$  for all channel parameters  $\epsilon$ . In this work, we resort to Monte-Carlo simulations of the peeling decoding process to estimate the constants  $\nu$  and  $\theta$ . We set

$M = 10^4$  and the highest  $\epsilon$  for which the system operates in an effectively error-free regime.

In Fig. 2, we plot the approximation of the CDF and the PDF of the first hit time for the terminated ensemble (red dashed line) by the Erlang distribution with parameters  $\tilde{\alpha}, \tilde{\beta}$ , and  $\tilde{\gamma}$  computed for  $\epsilon = 0.4875$ . The approximation is in good agreement with the simulated distributions of the first hit time of the peeling decoding, which supports the proposed model.

We remark that making  $\alpha, \beta$ , and  $\gamma$  dependent on  $\epsilon$  does not improve the prediction of the original scaling law in [9], further indicating that the main disagreement comes from modeling the decoding process with a single Ornstein-Uhlenbeck process instead of two processes.

### C. Scaling Law to Predict the Bit Error Rate

It is possible to apply the described framework to estimate the BER performance of an SC-LDPC code ensemble. Suppose the peeling decoder halted at normalized time  $\tau_0 = x$ . In that case it would have performed  $xM$  iterations before the decoding failure, so approximately  $\epsilon LM - xM$  out of  $LM$  bits would remain unrecovered. Accordingly, the BER can be approximated as

$$P_b \approx \int_{\alpha}^{\beta} \left(\epsilon - \frac{x}{L}\right) f_{\tau_0}(x) dx. \quad (10)$$

Using the approximation (7)–(8) in (10), we obtain the predicted BER for the terminated ensemble as

$$P_b \approx \exp\left(-\frac{\beta - \alpha}{\mu_0}\right) \cdot \frac{\beta^2 + \alpha\epsilon L - (\epsilon L + \alpha - 2\mu_0)(\beta + \mu_0)}{\mu_0 L} + \frac{\epsilon L - \alpha - 2\mu_0}{L}. \quad (11)$$

In summary, the prediction of the FER and BER for the terminated  $(d_v, d_c, L, M)$  SC-LDPC code ensemble is given by (9) and (11), respectively, with parameters  $(\tilde{\alpha}_\epsilon, \tilde{\beta}_\epsilon, \tilde{\gamma}_\epsilon, \tilde{\nu}, \tilde{\theta})$ . The dependence of the parameters on  $\epsilon$  is highlighted with the subscript  $\epsilon$ .

## IV. NUMERICAL RESULTS

In Fig. 3, we compare the simulated FER and BER performance with the analytical approximations in (9) and (11) for the terminated  $(5, 10, L=50, M)$  SC-LDPC code ensemble with  $M = 500, 1000$ , and  $2000$ . For comparison purposes, we also include the FER predicted using the scaling law in [9] (see Section II-A). We observe that the proposed refined scaling law yields a significantly improved prediction of the FER performance. The BER performance is also very well predicted. We remark that, similar to the approach used in [8], to remove the effect of the error floor we consider an expurgated ensemble by ignoring all failures involving only size-2 stopping sets in the calculation of the simulated error rates.

In Fig. 4, we show the FER and BER performance of the terminated  $(4, 8, L=50, M)$  SC-LDPC code ensemble. Similar to the  $(5, 10, L=50, M)$  ensemble, a very good match is observed between the predicted curves and the simulation results.

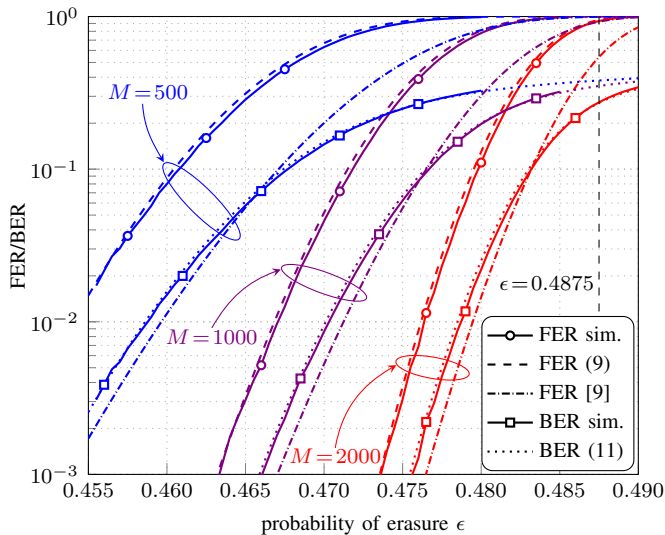


Fig. 3. Simulated FER and BER curves and the corresponding analytical approximations for the terminated  $(5, 10, L = 50, M)$  ensemble with  $\epsilon^* = 0.4994$  for different  $M$ . The parameters  $\tilde{\nu}$  and  $\tilde{\theta}$  are estimated at  $\epsilon = 0.485$  to be  $\tilde{\nu} \approx 0.424$  and  $\tilde{\theta} \approx 1.64$ .

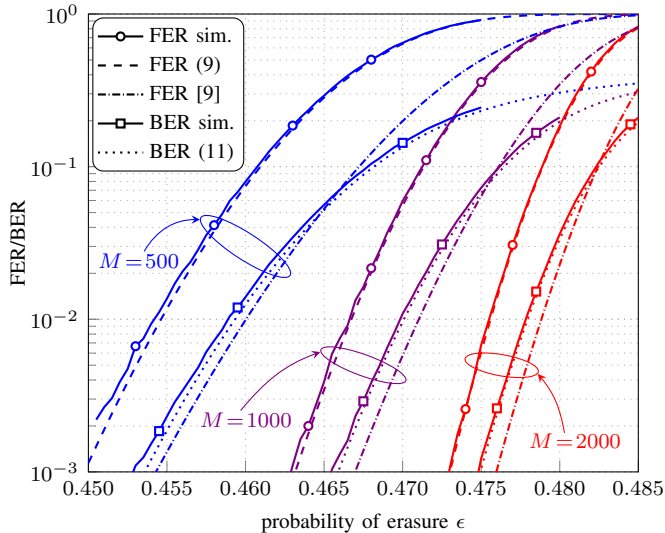


Fig. 4. Simulated FER and BER curves and the corresponding analytical approximations for the terminated  $(4, 8, L = 50, M)$  ensemble with  $\epsilon^* = 0.4977$  for different  $M$ . The parameters  $\tilde{\nu}$  and  $\tilde{\theta}$  are estimated at  $\epsilon = 0.48$  to be  $\tilde{\nu} \approx 0.406$  and  $\tilde{\theta} \approx 1.47$ .

Finally, in Fig. 5 we give results for the terminated  $(3, 6, L = 50, M)$  SC-LDPC ensemble. In this case, while the refined scaling law yields a significantly better prediction than the scaling law in [9], a gap between the predicted and simulated curves remains. In general, we observed that, irrespective of the code rate, for ensembles with VN degree  $d_v \geq 4$  the predicted error rates are very accurate, whereas for  $d_v = 3$  a gap appears, albeit significantly smaller than that for the original scaling law in [9].

## V. CONCLUSION

We proposed a refined finite length scaling law for terminated SC-LDPC codes over the BEC. The proposed scaling

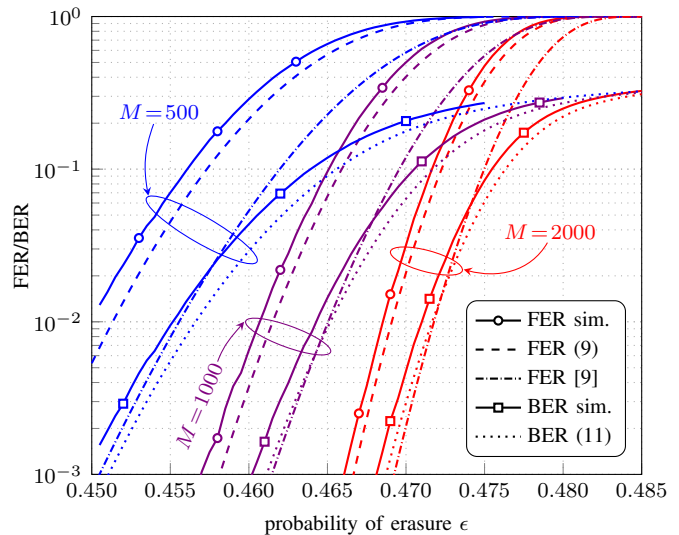


Fig. 5. Simulated FER and BER curves and the corresponding analytical approximations for the terminated  $(3, 6, L = 50, M)$  ensemble with  $\epsilon^* = 0.4881$  for different  $M$ . The parameters  $\tilde{\nu}$  and  $\tilde{\theta}$  are estimated at  $\epsilon = 0.475$  to be  $\tilde{\nu} \approx 0.338$  and  $\tilde{\theta} \approx 1.28$ .

models the decoding process as two independent Ornstein-Uhlenbeck processes, corresponding to the two decoding waves moving from the boundaries toward the center of the coupled chain. This model accurately predicts the frame and bit error rate performance of terminated SC-LDPC ensembles with VN degree  $d_v \geq 4$ . Closing the small gap in the case of  $d_v = 3$  is an interesting research problem, which may be partially addressed by taking into account the possibility of a decoding failure outside of the steady state region.

## REFERENCES

- [1] A. Jimenéz Feltström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2181–2191, Sep. 1999.
- [2] M. Lentmaier, A. Sridharan, D. J. Costello, and K. S. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5274–5289, Oct. 2010.
- [3] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.
- [4] S. Kudekar, T. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.
- [5] S. Moloudi, M. Lentmaier, and A. Graell i Amat, "Spatially coupled turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6199–6215, Oct. 2017.
- [6] K. Wang, W. Hou, S. Lu, P. Wu, Y. Ueng, and J. Cheng, "Improving polar codes by spatial coupling," in *Proc. IEEE Int. Symp. Inf. Theory Appl. (ISITA)*, Singapore, Oct. 2018, pp. 432–436.
- [7] B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightw. Technol.*, vol. 30, no. 1, pp. 110–117, Jan. 2012.
- [8] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 473–498, Feb. 2009.
- [9] P. M. Olmos and R. L. Urbanke, "A scaling law to predict the finite-length performance of spatially-coupled LDPC codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3164–3184, Jun. 2015.
- [10] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 569–584, Feb. 2001.