Why We Can Not Surpass Capacity: The Matching Condition

Cyril Méasson
EPFL
CH-1015 Lausanne
cyril.measson@epfl.ch,

Andrea Montanari
LPT, ENS
F-75231 Paris
montanar@lpt.ens.fr

Rüdiger Urbanke
EPFL
CH-1015 Lausanne
ruediger.urbanke@epfl.ch

Abstract

We show that iterative coding systems can not surpass capacity *using only quantities which naturally appear in density evolution.* Although the result in itself is trivial, the method which we apply shows that in order to achieve capacity the various components in an iterative coding system have to be perfectly matched. This generalizes the perfect matching condition which was previously known for the case of transmission over the binary erasure channel to the general class of binary-input memoryless output-symmetric channels. Potential applications of this perfect matching condition are the construction of capacity-achieving degree distributions and the determination of the number required iterations as a function of the multiplicative gap to capacity.

Assume we are transmitting over a binary-input memoryless output-symmetric (BMS) channel using sparse graph codes and an iterative decoder. Why can we not surpass capacity with such a set-up? The trivial answer is of course given by the converse to the channel coding theorem. In this paper we give an alternative proof of this fact which uses only quantities which are naturally tied to the setup of iterative coding. We show that in order to achieve capacity the various components in an iterative coding system have to be perfectly matched. This generalizes the perfect matching condition which was previously known for the case of transmission over the binary erasure channel (BEC) to the general class of BMS channels.

The first bound in which capacity was derived explicitly from density evolution was given by Shokrollahi and Oswald [1, 2] for the case of transmission over the BEC. For the same channel, a very pleasing geometric bound using ten Brink's EXIT charts ([3–6]) was later given by Ashikhmin, Kramer and ten Brink using the Area Theorem [7–9]. See also [10–13] for related work and the extension to parallel concatenation.

For general BMS channels, this geometric interpretation is unfortunately no longer valid since the Area Theorem is no longer fulfilled by the component-wise EXIT curves. Motivated by the pleasing geometric statement observed for the BEC, a similar chart, called MSE chart was constructed by Bhattad and Narayanan [14]. Assuming that the input densities to the component codes are Gaussian, this chart again fulfills the Area Theorem. The introduction of this function was motivated by the elegant relationship between mutual information and signal-to-noise observed by Guo, Shamai and Verdu [15, 16]. In order to apply the MSE chart in the context of iterative coding the authors proposed to approximate the intermediate densities which appear in density evolution by "equivalent" Gaussian densities. This was an important first step in generalizing the matching condition to the whole class of BMS channels.

In the following we show how to overcome the need for making the Gaussian approximation by using generalized EXIT (GEXIT) functions [17]. The bound which we derive and its geometric interpretation is quite similar to the one given for the BEC: We represent the "actions" performed by each component code by their respective GEXIT functions. By construction the area under these GEXIT functions is related to the rate of the corresponding codes. Further, we show that if we are transmitting below the threshold of iterative coding then these two curves do not overlap. Using then an argument identical to the one introduced for the BEC it follows that in order to achieve capacity the two individual component curves have to be perfectly matched.

There are two obvious potential applications of the perfect matching condition. First, assuming a perfect matching of the component curves and working backwards, one might be able to exhibit capacity-achieving degree distribution pairs. Secondly, the geometric picture given by the bound seems to provide the natural setting to prove that the number of iterations of iterative codings systems scales at least like $\Theta(1/\delta)$, where δ is the multiplicative gap to capacity, another long standing conjecture of iterative coding.

1 EXIT Charts and the Matching Condition for BEC

To start, let us review the case of transmission over the BEC($\mathtt{h}$) using a degree distribution pair (λ, ρ). In this case density evolution is equivalent to the EXIT chart approach and the condition for successful decoding under BP reads

$$c(x) \triangleq 1 - \rho(1-x) \leq \lambda^{-1}(x/\mathtt{h}) \triangleq v_{\mathtt{h}}^{-1}(x).$$

This is shown in Fig. 1 for the degree distribution pair $(\lambda(x) = x^3, \rho(x) = x^4)$. The area

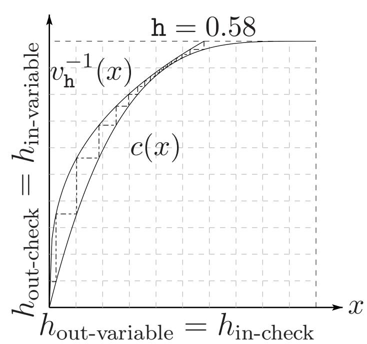

Figure 1: The EXIT chart method for the degree distribution $(\lambda(x) = x^3, \rho(x) = x^4)$ and transmission over the BEC($\mathtt{h} = 0.58$).

under the curve $c(x)$ equals $1 - \int \rho$ and the area to the left of the curve $v_{\mathtt{h}}^{-1}(x)$ is equal to $\mathtt{h} \int \lambda$. By the previous remarks, a necessary condition for successful BP decoding is that these two areas do not overlap. Since the total area equals 1 we get the necessary condition $\mathtt{h} \int \lambda + 1 - \int \rho \leq 1$. Rearranging terms, this is equivalent to the condition

$$1 - C_{\text{Sh}} = \mathtt{h} \leq \frac{\int \rho}{\int \lambda} = 1 - r(\lambda, \rho).$$

In words, the rate $r(\lambda, \rho)$ of any LDPC ensemble which, for increasing block lengths, allows successful decoding over the BEC($\mathtt{h}$), can not surpass the Shannon limit $1 - \mathtt{h}$. As pointed out in the introduction, an argument very similar to the above was introduced by Shokrollahi and Oswald [1,2] (albeit not using the language and geometric interpretation of EXIT functions and applying a slightly different range of integration). It was the first bound on the performance of iterative systems in which the Shannon capacity appeared explicitly using only quantities of density evolution. A substantially more general version of this bound can be found in [7–9] (see also Forney [18]).

Although the final result (namely that transmission above capacity is not possible) is trivial, the method of proof is well worth the effort since it shows how capacity enters in the calculation of the performance of iterative coding systems. By turning this bound around, we can find conditions under which iterative systems achieve capacity: In particular it shows that the two component-wise EXIT curves have to be matched perfectly. Indeed, all currently known capacity achieving degree-distributions for the BEC can be derived by starting with this perfect matching condition and working backwards.

2 GEXIT Charts and the Matching Condition for BMS Channels

Let us now derive the equivalent result for general BMS channels. As a first ingredient we show how to interpolate the sequence of densities which we get from density evolution so as to form a complete family of densities.

Definition 1 (Interpolating Channel Families) *Consider a degree distribution pair* (λ, ρ) *and transmission over the BMS channel characterized by its L-density* c. *Let* $\mathsf{a}_{-1} = \Delta_0$ *and* $\mathsf{a}_0 = \mathsf{c}$ *and set* a_α, $\alpha \in [-1, 0]$, *to* $\mathsf{a}_\alpha = -\alpha \mathsf{a}_{-1} + (1+\alpha)\mathsf{a}_0$. *The* interpolating density evolution families $\{\mathsf{a}_\alpha\}_{\alpha=-1}^{\infty}$ *and* $\{\mathsf{b}_\alpha\}_{\alpha=0}^{\infty}$ *are then defined as follows:*

$$\mathsf{b}_\alpha = \sum_i \rho_i \mathsf{a}_{\alpha-1}^{\boxast(i-1)}, \qquad \alpha \geq 0,$$
$$\mathsf{a}_\alpha = \sum_i \lambda_i \mathsf{c} \star \mathsf{b}_\alpha^{\star(i-1)}, \qquad \alpha \geq 0,$$

where $\star$ *denotes the standard convolution of densities and* $\mathsf{a} \boxast \mathsf{b}$ *denotes the density at the output of a check node, assuming that the input densities are* a *and* b, *respectively.*

Discussion: First note that a_ℓ (b_ℓ), $\ell \in \mathbb{N}$, represents the sequence of L-densities of density evolution emitted by the variable (check) nodes in the ℓ-th iteration. By starting density evolution not only with $\mathsf{a}_0 = \mathsf{c}$ but with all possible convex combinations of Δ_0 and c, this discrete sequence of densities is completed to form a continuous family of densities ordered by physical degradation. The fact that the densities are ordered by physical degradation can be seen as follows: note that the computation tree for a_α can be constructed by taking the standard computation tree of $\mathsf{a}_{\lceil\alpha\rceil}$ and independently erasing the observation associated to each variable leaf node with probability $\lceil\alpha\rceil - \alpha$. It follows that we can convert the computation tree of a_α to that of $\mathsf{a}_{\alpha-1}$ by erasing all observations at the leaf nodes and by independently erasing each observation in the second (from the bottom) row of variable nodes with probability $\lceil\alpha\rceil - \alpha$. The same statement is true for b_α. If $\lim_{\ell\to\infty} H(\mathsf{a}_\ell) = 0$, i.e., if BP decoding is successful in the limit of large blocklengths, then the families are both complete.

Example 1 (Density Evolution and Interpolation) *Consider transmission over the BSC(0.07) using a (3,6)-regular ensemble. Fig. 2 depicts the density evolution process for this case. This process gives rise to the sequences of densities* $\{\mathsf{a}_\ell\}_{\ell=0}^{\infty}$, *and* $\{\mathsf{b}_\ell\}_{\ell=1}^{\infty}$.

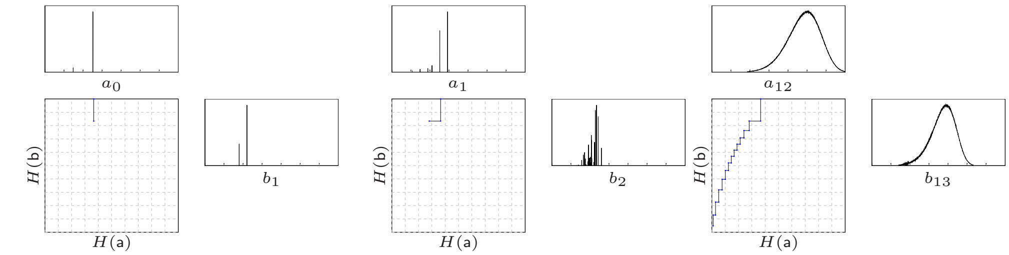

Figure 2: Density evolution for (3, 6)-regular ensemble over BSC(0.07).

Fig. 3 shows the interpolation of these sequences for the choices $\alpha = 1.0, 0.95, 0.9$ *and* 0.8 *and the complete such family.*

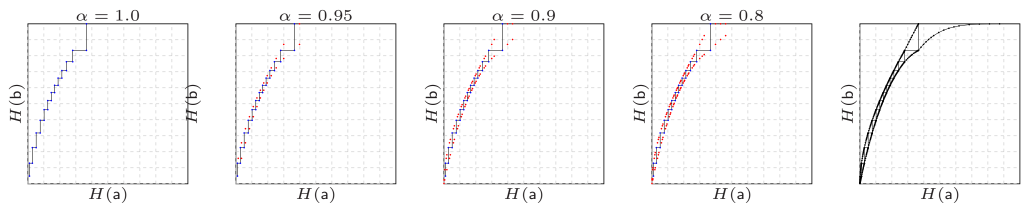

Figure 3: Interpolation of densities.

As a second ingredient we recall from [17] the definition of GEXIT functions. These GEXIT functions fulfill the Area Theorem for the case of general BMS channels. Up to date, GEXIT functions have been mainly used to derive upper bounds on the MAP threshold of iterative coding systems, see e.g., [17, 19]. Here we will apply them to the components of LDPC ensembles.

Definition 2 (The GEXIT Functional) *Given two families of L-densities* $\{\mathsf{c}_\epsilon\}$ *and* $\{\mathsf{a}_\epsilon\}$ *parameterized by* ϵ *define the GEXIT functional* $G(\mathsf{c}_\epsilon, \mathsf{a}_\epsilon)$ *by*

$$G(\mathsf{c}_\epsilon, \mathsf{a}_\epsilon) = \int_{-\infty}^{\infty} \mathsf{a}_\epsilon(z) l^{\mathsf{c}_\epsilon}(z)\, dz,$$

where

$$l^{\mathsf{c}_\epsilon}(z) = \frac{\int_{-\infty}^{\infty} \frac{d\mathsf{c}_\epsilon(w)}{d\epsilon} \log(1 + e^{-z-w})\, dw}{\int_{-\infty}^{\infty} \frac{d\mathsf{c}_\epsilon(w)}{d\epsilon} \log(1 + e^{-w})\, dw}.$$

Note that the kernel is normalized not with respect to $d\epsilon$ *but with respect to* $d\mathtt{h}$, *i.e., with respect to changes in the entropy. The families are required to be smooth in the sense that* $\{H(\mathsf{c}_\epsilon), G(\mathsf{c}_\epsilon, \mathsf{a}_\epsilon)\}$ *forms a piecewise continuous curve.*

Lemma 1 (GEXIT and Dual GEXIT Function) *Consider a binary code* C *and transmission over a complete family of BMS channels characterized by their family of L-densities* $\{\mathsf{c}_\epsilon\}$. *Let* $\{\mathsf{a}_\epsilon\}$ *denote the corresponding family of (average) extrinsic MAP densities. Then the standard GEXIT curve is given in parametric form by* $\{H(\mathsf{c}_\epsilon), G(\mathsf{c}_\epsilon, \mathsf{a}_\epsilon)\}$. *The* dual *GEXIT curve is defined by* $\{G(\mathsf{a}_\epsilon, \mathsf{c}_\epsilon), H(\mathsf{a}_\epsilon)\}$. *Both, standard and dual GEXIT curve have an area equal to* $r(C)$, *the rate of the code.*

Discussion: Note that both curves are "comparable" in that they first component measures the channel c and the second argument measure the MAP density a. The difference between the two lies in the choice of measure which is applied to each component.

Proof 1 *Consider the entropy $H(\mathsf{c}_\epsilon \star \mathsf{a}_\epsilon)$. We have*

$$\begin{aligned} H(\mathsf{c}_\epsilon \star \mathsf{a}_\epsilon) &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} \mathsf{c}_\epsilon(w) \mathsf{a}_\epsilon(v-w)\, dw \Big) \log(1+e^{-v})\, dv \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathsf{c}_\epsilon(w) \mathsf{a}_\epsilon(z) \log(1+e^{-w-z})\, dw\, dz \end{aligned}$$

Consider now $\frac{dH(\mathsf{c}_\epsilon \star \mathsf{a}_\epsilon)}{d\epsilon}$. Using the previous representation we get

$$\begin{aligned} \frac{dH(\mathsf{c}_\epsilon \star \mathsf{a}_\epsilon)}{d\epsilon} = &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\mathsf{c}_\epsilon(w)}{d\epsilon} \mathsf{a}_\epsilon(z) \log(1+e^{-w-z})\, dw\, dz + \\ &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathsf{c}_\epsilon(w) \frac{d\mathsf{a}_\epsilon(z)}{d\epsilon} \log(1+e^{-w-z})\, dw\, dz. \end{aligned}$$

The first expression can be identified with the standard GEXIT curve except that it is parameterized by a generic parameter ϵ. The second expression is essentially the same, but the roles of the two densities are exchanged.

Integrate now this relationship over the whole range of ϵ and assume that this range goes from "perfect" (channel) to "useless". The integral on the left clearly equals 1. To perform the integrals over the right reparameterize the first expression with respect to $\mathtt{h} \triangleq \int_{\infty}^{\infty} \mathsf{c}_\epsilon(w) \log(1+e^{-w}) dw$ so that it becomes the standard GEXIT curve given by $\{H(\mathsf{c}_\epsilon), G(\mathsf{c}_\epsilon, \mathsf{a}_\epsilon)\}$. In the same manner reparameterize the second expression by $\mathtt{h} \triangleq \int_{\infty}^{\infty} \mathsf{a}_\epsilon(w) \log(1+e^{-w}) dw$ so that it becomes the curve given by $\{H(\mathsf{a}_\epsilon), G(\mathsf{a}_\epsilon, \mathsf{c}_\epsilon)\}$. Since the sum of the two areas equals one and the area under the standard GEXIT curve equals $r(C)$, it follows that the area under the second curve equals $1 - r(C)$. Finally, note that if we consider the inverse of the second curve by exchanging the two coordinates, i.e., if we consider the curve $\{G(\mathsf{a}_\epsilon, \mathsf{c}_\epsilon), H(\mathsf{a}_\epsilon)\}$, then the area under this curve is equal to $1 - (1 - r(C)) = r(C)$, as claimed.

Example 2 (EXIT Versus GEXIT) *Fig. 4 compares the EXIT function to the GEXIT function for the $[3, 1, 3]$ repetition code and the $[6, 5, 2]$ single parity-check code when transmission takes place over the BSC. As we can see, the two curves are similar but distinct. In particular note that the areas under the GEXIT curves are equal to the rate of the codes but that this is not true for the EXIT functions.*

Example 3 (GEXIT Versus Dual GEXIT) *Fig. 5 shows the standard GEXIT function and the dual GEXIT function for the $[5, 4, 2]$ code and transmission over the BSC. Although the two curves have quite distinct shapes, the area under the two curves is the same.*

Lemma 2 *Consider a degree distribution pair (λ, ρ) and transmission over an BMS channel characterized by its L-density c so that density evolution converges to Δ_∞. Let $\{\mathsf{a}_\alpha\}_{\alpha=-1}^{\infty}$ and $\{\mathsf{b}_\alpha\}_{\alpha=0}^{\infty}$ denote the interpolated families as defined in Definition 1.*

Then the two GEXIT curves parameterized by

$$\{H(\mathsf{a}_\alpha), G(\mathsf{a}_\alpha, \mathsf{b}_{\alpha+1})\}, \qquad \text{GEXIT of check nodes}$$

$$\{H(\mathsf{a}_\alpha), G(\mathsf{a}_\alpha, \mathsf{b}_\alpha)\}, \qquad \text{inverse of dual GEXIT of variable nodes}$$

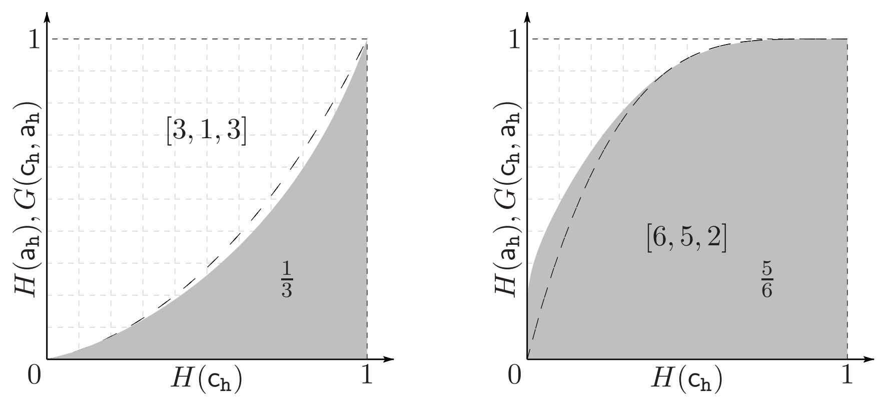

Figure 4: A comparison of the EXIT with the GEXIT function for the $[3, 1, 3]$ and the $[6, 5, 2]$ code.

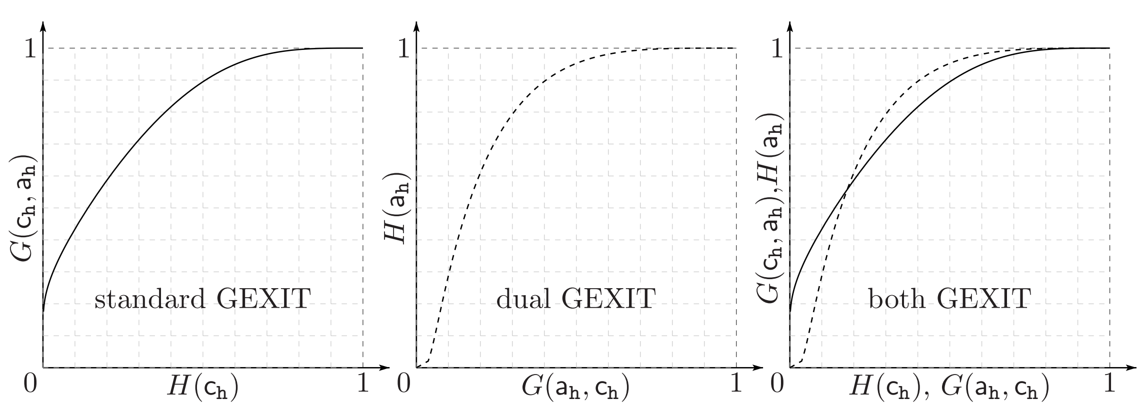

Figure 5: Standard and dual GEXIT function of $[5, 4, 2]$ code and transmission over the BSC.

do not overlap and faithfully represent density evolution. Further, the area under the "check-node" GEXIT function is equal to $1 - \int \rho$ and the area to the left of the "inverse dual variable node" GEXIT function is equal to $H(\mathsf{c}) \int \lambda$. It follows that $r(\lambda, \rho) \leq 1 - H(\mathsf{c})$, i.e., the transmission rate can not exceed the Shannon limit.

This implies that transmission approaching capacity requires a perfect matching of the two curves.

Proof 2 *First note that $\{H(\mathsf{a}_\alpha), G(\mathsf{a}_\alpha, \mathsf{b}_{\alpha+1})\}$ is the standard GEXIT curve representing the action of the check nodes: a_α corresponds to the density of the messages* entering *the check nodes and $\mathsf{b}_{\alpha+1}$ represents the density of the corresponding output messages. On the other hand, $\{H(\mathsf{a}_\alpha), G(\mathsf{a}_\alpha, \mathsf{b}_\alpha)\}$ is the inverse of the dual GEXIT curve corresponding to the action at the variable nodes: now the input density to the check nodes is b_α and a_α denotes the corresponding output density.*

The fact that the two curves do not overlap can be seen as follows. Fix an entropy value. This entropy value corresponds to a density a_α for a unique value of α. The fact that $G(\mathsf{a}_\alpha, \mathsf{b}_\alpha) \geq G(\mathsf{a}_\alpha, \mathsf{b}_{\alpha+1})$ now follows from the fact that $\mathsf{b}_{\alpha+1} \prec \mathsf{b}_\alpha$ and that for any symmetric a_α this relationship stays preserved by applying the GEXIT functional.

The statements regarding the areas of the two curves follow in a straightforward manner from the GAT and Lemma 1. The bound on the achievable rate follows in the same manner as for the BEC: the total area of the GEXIT box equals one and the two curves do not overlap and have areas $1 - \int \rho$ and $H(\mathsf{c})$. It follows that $1 - \int \rho + H(\mathsf{c}) \int \lambda \leq 1$, which is equivalent to the claim $r(\lambda, \rho) \leq 1 - H(\mathsf{c})$.

We see that the matching condition still holds even for general channels. There are a few important differences between the general case and the simple case of transmission over the BEC. For the BEC, the intermediate densities are always the BEC densities independent of the degree distribution. This of course enormously simplifies the task. Further, for the BEC, given the two EXIT curves, the progress of density evolution is simply given by a staircase function bounded by the two EXIT curves. For the general case, this staircase function still has vertical pieces but the "horizontal" pieces are in general at an angle. This is true since the y-axis for the "check node" step measures $G(\mathsf{a}_\alpha, \mathsf{b}_{\alpha+1})$, but in the subsequent "inverse variable node" step it measures $G(\mathsf{a}_{\alpha+1}, \mathsf{b}_{\alpha+1})$. Therefore, one should think of two sets of labels on the y-axis, one measuring $G(\mathsf{a}_\alpha, \mathsf{b}_{\alpha+1})$, and the second one measuring $G(\mathsf{a}_{\alpha+1}, \mathsf{b}_{\alpha+1})$. The "horizontal" step then consists of first switching from the first y-axis to the second, so that the labels correspond to the same density b and then drawing a horizontal line until it crosses the "inverse variable node" GEXIT curve. The "vertical" step stays as before, i.e., it really corresponds to drawing a vertical line. All this is certainly best clarified by a simple example.

Example 4 ($(3,6)$-Regular Ensemble and Transmission over BSC) *Consider the $(3,6)$-regular ensemble and transmission over the BSC(0.07). The corresponding illustrations are shown in Fig. 6. The top-left figure shows the standard GEXIT curve for the check node side. The top-right figure shows the dual GEXIT curve corresponding to the variable node side. In order to use these two curves in the same figure, it is convenient to consider the inverse function for the variable node side. This is shown in the bottom-left figure. In the bottom-right figure both curves are shown together with the "staircase" like function which represents density evolution. As we see, the two curves to not overlap and have both the correct areas.*

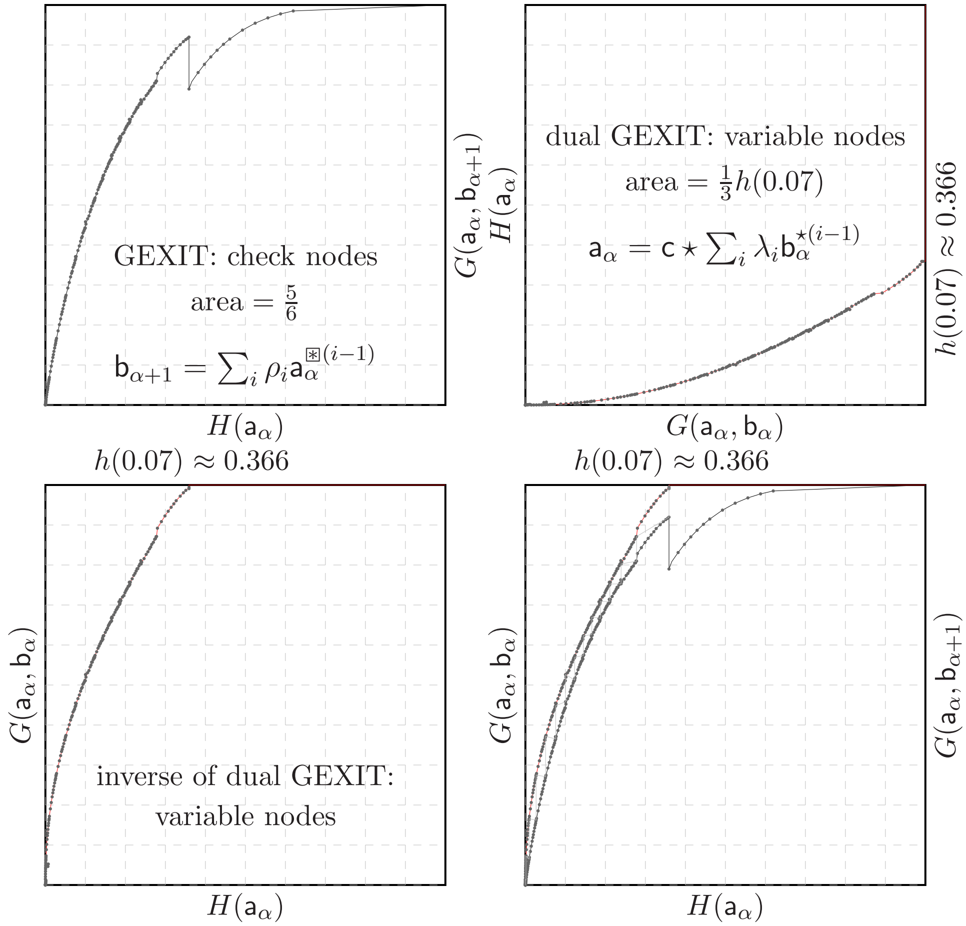

Figure 6: Faithful representation of density evolution by two non-overlapping component-wise GEXIT functions which represent the "actions" of the check nodes and variable nodes, respectively. The area between the two curves equals is equal to the additive gap to capacity.

As remarked earlier, one potential use of the matching condition is to find capacity approaching degree distribution pairs. Let us quickly outline a further such potential application. Assuming that we have found a sequence of capacity-achieving degree distributions, how does the number of required iterations scale as we approach capacity. It has been conjectured that the the number of required iterations scales like $1/\delta$, where δ is the gap to capacity. This conjecture is based on the geometric picture which the matching condition implies. To make things simple, imagine the two GEXIT curves as two parallel lines, lets say both at a 45 degree angle, a certain distance apart, and think of density evolution as a staircase function. From the previous results, the area between the lines is proportional to δ. Therefore, if we half δ the distance between the lines has to be halved and one would expect that we need twice as many steps. Obviously, the above discussion was based on a number of simplifying assumptions. It remains to be seen if this conjecture can be proven rigorously.

Acknowledgments

The work of A. Montanari was partially supported by the European Union under the project EVERGROW.

References

[1] A. Shokrollahi, “Capacity-achieving sequences,” in *Codes, Systems, and Graphical Models* (B. Marcus and J. Rosenthal, eds.), vol. 123 of *IMA Volumes in Mathematics and its Applications*, pp. 153–166, Springer-Verlag, 2000.

[2] P. Oswald and A. Shokrollahi, “Capacity achieving sequences for the erasure channel,” in *Proceedings of the International Symposium on Information Theory, Washington DC*, p. 48, 2001.

[3] S. ten Brink, “Convergence of iterative decoding,” *Electron. Lett.*, vol. 35, pp. 806–808, May 1999.

[4] S. ten Brink, “Iterative decoding for multicode CDMA,” in *Proc. IEEE VTC*, vol. 3, pp. 1876–1880, May 1999.

[5] S. ten Brink, “Iterative decoding trajectories of parallel concatenated codes,” in *Proc. 3rd IEEE/ITG Conf. Source Channel Coding*, pp. 75–80, Jan. 2000.

[6] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 1727–1737, Oct. 2001.

[7] A. Ashikhmin, G. Kramer, and S. ten Brink, “Extrinsic information transfer functions: a model and two properties,” in *Proc. of Conference on Information Sciences and Systems (CISS)*, (Princeton University), Mar. 2002.

[8] A. Ashikhmin, G. Kramer, and S. ten Brink, “Code rate and the area under extrinsic information transfer curves,” in *Proc. of the IEEE Int. Symposium on Inform. Theory*, (Lausanne, Switzerland), p. 115, June 30–July 5 2002.

[9] A. Ashikhmin, G. Kramer, and S. ten Brink, “Extrinsic information transfer functions: model and erasure channel property,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 2657–2673, Nov. 2004.

[10] S. Huettinger, J. Huber, R. Johannesson, and R. Fischer, “Information processing in Soft-Ouput decoding,” in *Proc. of the Allerton Conf. on Commun., Control and Computing*, (Monticello, IL, USA), Oct. 2001.

[11] S. Huettinger and J. Huber, “Extrinsic and intrinsic information in systematic coding,” in *Proc. of the IEEE Int. Symposium on Inform. Theory*, (Lausanne, Switzerland), June 30–July 5 2002.

[12] C. Méasson and R. Urbanke, “Asymptotic analysis of turbo codes over the binary erasure channel,” in *Proc. of the 12th Joint Conference on Communications and Coding*, (Saas Fee, Switzerland), March 2002.

[13] C. Méasson and R. Urbanke, “Further analytic properties of EXIT-like curves and applications,” in *Proc. of the IEEE Int. Symposium on Inform. Theory*, (Yokohama, Japan), p. 266, June 29–July 4 2003.

[14] K. Bhattad and K. R. Narayanan, “An MSE based transfer chart to analyze iterative decoding schemes,” in *Proc. of the Allerton Conf. on Commun., Control and Computing*, (Monticello, IL, USA), Oct. 2004.

[15] D. Guo, S. Shamai, and S. Verdú, "Mutual information and MMSE in gaussian channels," in *IEEE International Symposium on Information Theory*, (Chicago, USA), p. 349, June 27 - July 2 2004.

[16] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1261–1882, Apr. 2005.

[17] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke, "Life above threshold: From list decoding to area theorem and MSE," in *Proc. of the IEEE Inform. Theory Workshop*, (San Antonio, Texas), October 24–29 2004.

[18] G. D. Forney, "Lecture notes." MIT, 2005.

[19] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke, "Maximum a posteriori decoding and turbo codes for general memoryless channels," in *Proc. of the IEEE Int. Symposium on Inform. Theory*, (Adelaide), 2005.