

Nonlinear Phenomena in the Turbo Decoding Algorithm

Z. Tasev[†], L. Kocarev[†], F. Lehmann[‡] and G. M. Maggio[‡]

[†] Institute for Nonlinear Science
University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093-0402
E-mail: ztasev@ucsd.edu, lkocarev@ucsd.edu

[‡] STMicroelectronics, Inc. &
Center for Wireless Communications
University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093-0407
E-mail: lehmann@ece.ucsd.edu, gmaggio@ucsd.edu

Abstract

In this paper, we treat the turbo decoding algorithm as a dynamical system parameterized by a single parameter that closely approximates the signal-to-noise ratio (SNR). A whole range of phenomena known to occur in nonlinear systems, like the existence of multiple fixed points, oscillatory behavior, bifurcations, chaos and transient chaos are found in the turbo decoding algorithm.

1. Introduction

In 1948, Shannon formulated the famous noisy channel coding theorem [1], which establishes the fundamental limits for digital communication in terms of channel capacity. Recently, it has been recognized that two classes of codes, namely *turbo codes* [2] and *low-density parity-check* (LDPC) codes [3, 4] perform at rates extremely close to the Shannon limit. Both codes are based on a similar philosophy: constrained random code ensembles, described by some fixed parameters plus randomness, decoded using iterative decoding algorithms (or message passing decoders). Turbo codes were discovered by Berrou *et al.* in 1993 [2]. On the other hand, LDPC codes were originally introduced by Gallager [3] in 1962. The crucial innovation of LDPC codes being the introduction of iterative decoding algorithms. LDPC codes were rediscovered by MacKay *et al.* [4] in 1996. Moreover, iterative decoding of turbo codes was recognized as instances of sum-product algorithms for codes defined on general graphs [5].

So far, most studies about turbo and LDPC codes have relied upon methods of information theory. However, iterative decoding algorithms can be viewed as complex nonlinear dynamical systems. Very recently, in a pioneering paper [6], Richardson has presented a geometrical interpretation of the turbo decoding algorithm and formalized it as a discrete-time dynamical system defined on a continuous set. This approach

clearly demonstrates the relationship between turbo decoding and maximum-likelihood decoding. The turbo decoding algorithm appears as an iterative algorithm aimed at solving a system of $2n$ equations in $2n$ unknowns, where n is the block-length size. If the turbo decoding algorithm converges to a certain codeword, then the latter constitutes a solution to this set of equations. Conversely, solutions to these equations provide fixed points of the turbo decoding algorithm, seen as a nonlinear mapping. In a follow-up of this work [7], Agrawal and Vardy parameterized the turbo decoding algorithm, as a dynamical system, in terms of the SNR.

The aim of the present work is to contribute to the in-depth understanding of the turbo decoding algorithm, based on the well developed theory of nonlinear dynamical systems [8]. A whole range of phenomena, known to occur in nonlinear systems, such as multiple fixed points, oscillatory behavior, bifurcations, chaos and transient chaos, are shown to occur in the turbo decoding algorithm.

2. Turbo Decoding as Nonlinear Dynamical System

The block diagram of a classical turbo code system is shown in Fig. 1. The turbo code consists of a parallel concatenation of two recursive systematic binary convolutional codes, CC1 and CC2, separated by a random interleaver [2]. Let $\mathbf{i}$ be the information bit sequence of length n at the input to the turbo encoder, and let $\mathbf{p}_1(\mathbf{i})$ (resp. $\mathbf{p}_2(\mathbf{i})$) be the parity bits produced by the first (resp. second) encoder. The input sequence of the second encoder has the same Hamming weight as $\mathbf{i}$, but the bits are rearranged due to a permutation defined by the choice of the random interleaver, π . The information bit sequence, $\mathbf{i}$, along with the parity bit sequences $\mathbf{p}_1(\mathbf{i})$ and $\mathbf{p}_2(\mathbf{i})$, forms a turbo codeword $(\mathbf{i}, \mathbf{p}_1(\mathbf{i}), \mathbf{p}_2(\mathbf{i}))$ of length $3n$.

We assume that the turbo code is transmitted over an AWGN (additive white Gaussian noise) binary-input memoryless channel, using the BPSK (binary phase shift keying) modulation. We denote the transmitted turbo codeword by $(\mathbf{s}^0, \mathbf{s}^1, \mathbf{s}^2) \equiv (\mathbf{i}, \mathbf{p}_1(\mathbf{i}), \mathbf{p}_2(\mathbf{i}))$.

This work was supported in part by STMicroelectronics Inc., University of California (DiMi program), and the ARO (Army Research Office).

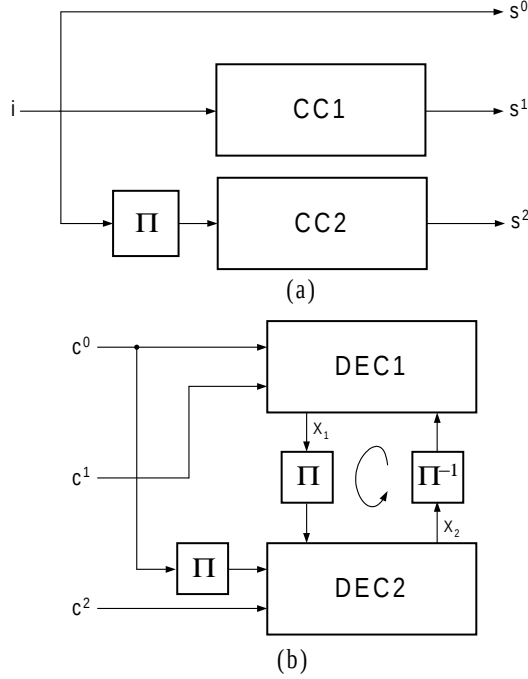


Figure 1: Classical turbo code: a) CC1 and CC2 are recursive convolutional codes; π indicates a random permutation. b) DEC1 and DEC2 are the decoders for CC1 and CC2, respectively; X_i ($i = 1, 2$) denotes the extrinsic information.

Let c^0 , c^1 and c^2 be the channel outputs corresponding to the input sequences s^0 , s^1 and s^2 , respectively. The turbo decoder consists of two components: a decoder DEC1 for the convolutional code CC1, and a decoder DEC2 for the code CC2. These decoders use iteratively the BCJR [9] algorithm in order to compute the extrinsic information, X_i , of the information bits. The input for the decoder DEC1 (resp. DEC2) is the extrinsic information X_2 (resp. X_1) from the decoder DEC2 (resp. DEC1), which is used as a priori information after interleaving (resp. deinterleaving).

The turbo decoding algorithm may be described by a discrete-time dynamical system of the form:

$$X_1(l+1) = F_1[X_2(l); c^0, c^1] \quad (1)$$

$$X_2(l) = F_2[X_1(l); c^0, c^2] \quad (2)$$

where $F_i = (F_{i1}, \dots, F_{in})$ are nonlinear functions which depend on the constituent codes.¹ Equations (1)-(2) define an n -dimensional mapping, because there are only n independent variables, namely X_1 (or alternatively X_2). Assuming

¹Namely, F_1 has the following form (similar expression applies for F_2):

$$F_{1k} = \log \frac{\sum_{i=1}^{2^n} f_{ik}}{\sum_{i=1}^{2^n} g_{ik}} - \frac{4}{\sigma^2} c_k^0 - X_{2k}, \quad k = 1, \dots, n$$

where f_{ik} and g_{ik} are some exponential functions of X_2 , c^0 , c^1 , and σ^2 is the noise variance.

that a uniform probability distribution of the a priori information is induced, the initial condition of the algorithm should be set at the origin.² The quantities c^0 , c^1 and c^2 are completely characterized by the channel likelihood ratios. Consequently, the mapping (1)-(2) depends on $3n$ parameters. As shown in [6], this mapping depends smoothly on its variables and parameters.

After l iterations, a hard decision on the k -th bit can be made according to the sign of the log-likelihood ratio:

$$\Lambda_k(l) = \log \frac{p_i^1(l)}{p_i^0(l)} = X_{1k}(l) + X_{2k}(l) + \frac{4}{\sigma^2} c_k^0 \quad (3)$$

where σ^2 represents the noise variance.

2.1. A posteriori average entropy

Since the turbo decoding algorithm is a high dimensional dynamical system, we suggest the following representation of its trajectories in the state space. At each iteration l , the turbo decoding algorithm computes $2n$ values of the extrinsic information, X_i . From Eq. (3) one can compute, for each iteration, the probabilities $p_k^0(l)$ and $p_k^1(l)$ that the k -th bit is either 0 or 1. Let us define

$$E(l) = -\frac{1}{n} \sum_{i=1}^n p_i^l(0) \ln p_i^l(0) + p_i^l(1) \ln p_i^l(1)$$

Then, E represents the *a posteriori average entropy*, which is a measure of the reliability of bit decisions, for an information block of size n . The advantage being that $E \in [0, 1]$. Namely, when all bits are detected correctly, p_i is either 0 or 1 for all i and, therefore, $E = 0$. On the other hand, when all bits are equally probable, i.e. ambiguous, $E = 1$.

3. Bifurcation Analysis of the Turbo Decoding Algorithm

As a complex dynamical system with a large number of parameters, the turbo decoding algorithm is not readily amenable for analysis. Namely, the mapping is parameterized by the noise values $c_1^0, c_2^0, \dots, c_n^0$, $c_1^1, c_2^1, \dots, c_n^1$, and $c_1^2, c_2^2, \dots, c_n^2$. To analyze the trajectories of the turbo decoding algorithm we parameterize it by the SNR $1/\sigma^2$ [7]. For large n , typical for turbo codes, the noise variance σ^2 is approximately equal to $\tilde{\sigma}^2 = \sum [(c_k^0)^2 + (c_k^1)^2 + (c_k^2)^2] / 3n$. We fix $(3n-1)$ noise ratios $c_1^0/c_2^0, c_2^0/c_3^0, \dots, c_{n-1}^0/c_n^0$ and treat the turbo decoding algorithm as it was depending on a single parameter $1/\tilde{\sigma}^2$, which is closely related to the SNR.

²However, one may introduce different initial conditions (needed for instance to compute Lyapunov exponents and average chaotic transient lifetime), without assuming that the information is known.

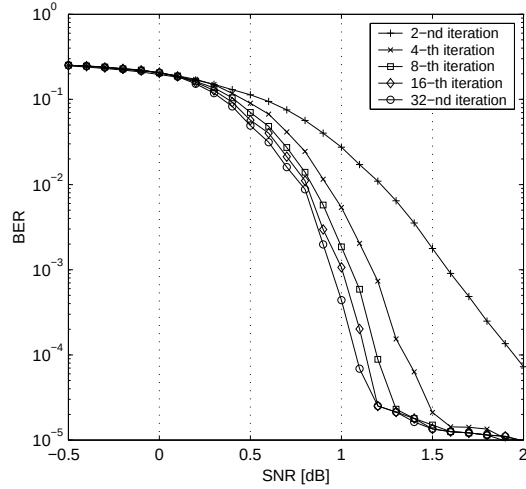


Figure 2: BER performance of the classical turbo code with interleaver length 1024 and rate 1/3, for increasing number of iterations. The so-called *waterfall* region of the turbo code corresponds approximately to the SNR region 0.25—1.25 dB.

3.1. Bifurcation diagram

In our analysis we consider the classical turbo code [2] with identical constituent recursive convolutional codes generated by the polynomials $\{D^4 + D^3 + D^2 + D^1 + 1, D^4 + 1\}$, producing a rate-1/3 turbo code. The codewords are transmitted over an AWGN channel using BPSK modulation. The length of the interleaver is $n = 1024$. Figure 2 shows the BER (bit-error-rate) performance of the turbo code versus the SNR, expressed in dB.

We have performed many simulations changing the parameter SNR from $-\infty$ to $+\infty$, with different realizations of the noise (different noise ratios $c_1^0/c_2^0, c_2^0/c_3^0, \dots, c_{n-1}^0/c_n^2$). The results of our analysis are summarized in Fig. 3.

We found that the turbo decoding algorithm admits two types of fixed points:

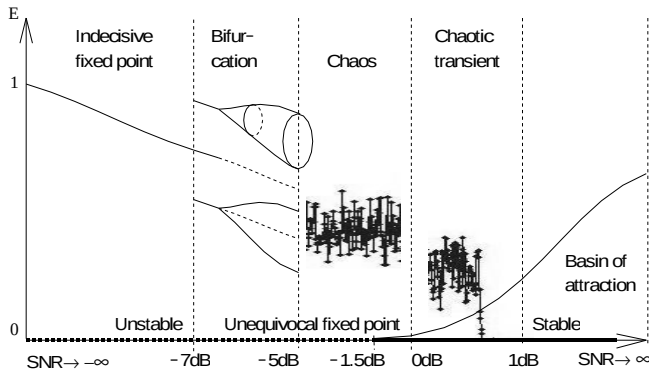


Figure 3: Bifurcation diagram for the classical turbo code.

- *Unequivocal fixed point*: In this case, p_k^0 is close to 0 or 1 for all k , so the log-likelihoods Λ_k (and X_{ik}) assume large values, and consequently $E = 0$. Hard decisions corresponding to unequivocal fixed point will typically form a valid codeword, usually coinciding with the transmitted codeword.
- *Indecisive fixed point*: This corresponds to $\Lambda_k \rightarrow 0$, i.e. equal probability that the k -th bit is 0 or 1, hence $E \rightarrow 1$. At this fixed point the turbo decoding algorithm is relatively ambiguous regarding the values of the information bits. Hard decisions corresponding to this fixed point typically will *not* form a valid codeword.

In each instance of the turbo decoding algorithm that we analyzed, an unequivocal fixed point existed for all values of SNR, from $-\infty$ to $+\infty$. This point is always represented by the average entropy $E = 0$. The unequivocal fixed point becomes stable at around -1.5 dB. However, the algorithm “cannot see” this point until 0–0.5 dB when, in some cases, the initial point of the algorithm (corresponding to $E = 1$) is within the basin of attraction of the fixed point.

For low values of SNR, when the SNR goes to $-\infty$, the indecisive fixed point is represented by $E = 1$. In our simulations we found that the indecisive fixed point moves toward smaller values of E , with increasing SNR. Then, it loses its stability (or disappears) at low SNRs, typically in the range from -7 dB to -5 dB (see Fig. 3). There are three ways in which a fixed point of a discrete-time dynamical system may fail to be hyperbolic [8]: when the Jacobian matrix evaluated at the fixed point admits complex conjugate eigenvalues on the unit circle, an eigenvalue +1, or an eigenvalue -1. We have observed all three corresponding bifurcations (discrete-time Hopf, tangent and flip bifurcation) in the turbo decoding algorithm.

For example, Figure 4 a) and b) illustrate a discrete-time Hopf bifurcation, also referred to as Neimark-Sacker bifurcation [8]. After the bifurcation, the resulting unstable fixed point is surrounded by an isolated, stable, close invariant limit cycle, with largest Lyapunov exponent [10] equal to zero. Correspondingly, a transition from indecisive to unequivocal fixed point occurs in the turbo decoding algorithm. In Fig. 3 this transition spans a large region of SNR: from -7 dB to 1 dB. In particular, the region -5 dB–0 dB is characterized by chaos: the turbo decoding algorithm as a dynamical system has a chaotic attractor. In the region 0 dB–1 dB chaotic transients occur.

We note here, by comparison with Fig. 2, that the waterfall region of the turbo code corresponds to the transient chaos behavior. Also, in this region the unequivocal fixed point is stable and the size of its basin of attraction gradually grows, for increasing SNR. In the waterfall region, the turbo decoding algorithm converges either to the chaotic invariant set or to the unequivocal fixed point, after a long transient behavior

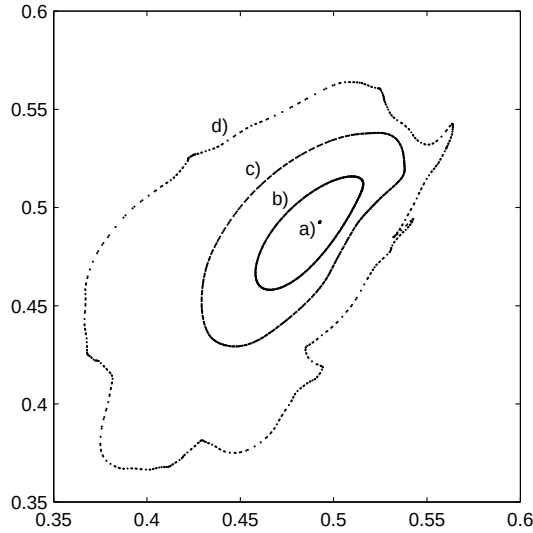


Figure 4: $E(l+1)$ vs. $E(l)$: a) and b) Discrete-time Hopf bifurcation; c) and d) Torus-breakdown route to chaos. Values of SNR: a) -6.7 dB; b) -6.5 dB; c) -6.3 dB; d) -6.2 dB.

indicating the existence of a chaotic non-attracting invariant set in the vicinity of the unequivocal fixed point. In some cases, the algorithm spends a few thousand iterations before reaching the fixed point solution. As SNR increases, the average chaotic transient lifetime [10] decreases.

Finally, for high SNR, the turbo decoding algorithm tends to converge to the unequivocal fixed point.

3.2. Chaotic behavior

The turbo decoding algorithm exhibits chaotic behavior, characterized by positive Lyapunov exponent [10], for a relatively large range of SNR values, from -6dB to 0dB (see Fig. 3). Figure 5 shows the largest Lyapunov exponent, λ , of the decoding algorithm, for one particular frame, as a function of the SNR.

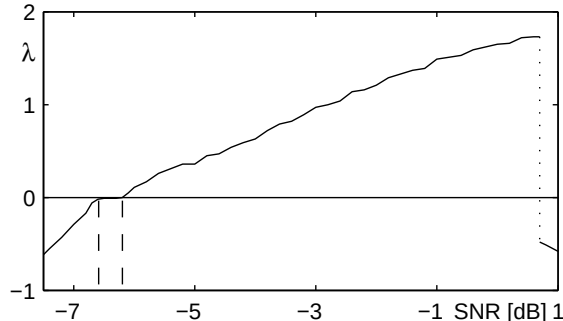


Figure 5: The largest Lyapunov exponent of the turbo decoding algorithm versus SNR, for one particular frame.

There are several ways (routes) by which chaotic attractors may arise. These include infinite period-doubling cascades, intermittency, crises and torus breakdown. A typical torus breakdown route to chaos is shown in Fig. 4 c) and d).

4. Conclusions

The turbo decoding algorithm can be viewed as a high-dimensional dynamical system parameterized by a large number of parameters. In this work, we have shown that the turbo decoding algorithm exhibits a whole range of phenomena known to occur in nonlinear systems. These include the existence of multiple fixed points, oscillatory behavior, bifurcations, chaos and transient chaos.

References

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, pp. 623–56, (1948).
- [2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-Codes," *Proc. IEEE International Communications Conference*, pp. 1064–70, (1993).
- [3] R.G. Gallager, "Low Density Parity Check Codes," *IRE Trans. Inf. Theory*, vol. 8, pp. 21–28, (1962).
- [4] D.J.C. MacKay and R.M. Neal, "Near Shannon limit performance of Low Density Parity Check Codes," *Electronics Letters*, vol. 32, pp. 1645–46, (1996).
- [5] N. Wiberg, "Codes and decoding on general graphs," *Dissertation thesis no. 440*, Dept. of Electrical Engineering, Linköping University, Sweden (1996).
- [6] T. Richardson, "The geometry of turbo-decoding dynamics," *IEEE Trans. Information Theory*, vol. 46, pp. 9–23 (2000).
- [7] D. Agrawal and A. Vardy, "The turbo decoding algorithm and its phase trajectories," *IEEE Trans. Information Theory*, vol. 47, No. 2, pp. 699–722 (2001).
- [8] S. Wiggins and F. John, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 3rd edition, Springer-Verlag, New York, (1990).
- [9] L.R. Bahl, J. Cocke, F. Jelinek, J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Information Theory*, vol. IT-20, no. 2, pp. 284–287 (1974).
- [10] E. Ott, *Chaos in dynamical systems*, Cambridge University Press, New York, (1993).