Magnetization enumerator of real-valued symmetric channels in Gallager error-correcting codes

N. S. Skantzos, J. van Mourik, and D. Saad

Neural Computing Research Group, Aston University, Birmingham, B4 7ET, United Kingdom
(Received 11 December 2002; published 21 March 2003)

Using the magnetization enumerator method, we evaluate the practical and theoretical limitations of symmetric channels with real outputs. Results are presented for several regular Gallager code constructions.

Error-correcting codes play a central role in modern communication, especially in noisy media such as in satellite and mobile communication. A broad range of error-correcting codes is on offer: they vary significantly in their practical and theoretical performance depending on the specific code chosen within a given code ensemble. Evaluating the limitations of specific code constructions is important for determining the code efficiency and for optimizing channel performance. One of the leading code ensembles to date is the family of low density parity check codes [1,2], which attracted significant interest both within and outside of the information theory community.

Methods of statistical physics have been recently employed to study the typical performance of various coding schemes, most notably of low density parity check codes [3–8]. Such studies have led to precise estimations of critical channel-noise levels (beyond which decoding is not possible) and also provided additional insight through the physical interpretation of various decoding schemes. The emerging picture for Gallager-type codes is that for sufficiently small noise levels, decoding is possible and the error-free (ferromagnetic) state is the only solution. For higher noise levels, one finds a transition to a regime where suboptimal solutions are created (spinodal or dynamical transition) and where existing practical decoding algorithms fail to find the most probable solution. For higher noise levels, a second transition occurs (thermodynamic transition) where the error-free solution ceases to be dominant. This marks the upper theoretical bound for error-free communication using the code specified. The number of equally plausible solutions to the decoding problem thereafter is exponential in the number of degrees of freedom. The thermodynamic transition approaches Shannon's limit with an increasing number of parity checks per bit.

One of the most important aspects in the decoding problem is the enumeration of possible solutions, as it provides a direct indication of the practical and theoretical performance of various decoding methods. A method for carrying out the analysis has been reported recently in Ref. [5]; the new approach generated interest in the application of the same method to other channel types characterized by real noise and real output values studied in this paper. In particular we study the cases of the Gaussian and Laplace channels which are of high practical relevance, and are used as standard benchmark channels for evaluating the performance of codes. We consider the channels to be symmetric, i.e., the probabilities $P(z_{\text{out}}|z_{\text{in}}) = P(-z_{\text{out}}|-z_{\text{in}})$, where z_{out} and z_{in} are the input and output channel values, respectively. We also

consider information vectors representing binary messages $t \in \{0,1\}^N$ to be encoded via Gallager's scheme.

A Gallager code is defined by the binary parity check matrix $A = [C_1|C_2]$ of dimensionality $(M-N) \times M$, which is a concatenation of two submatrices. A *regular* (K,C) Gallager code has a fixed number K of nonzero elements per row in A, and a fixed number C of nonzero elements per column. It follows that $C \equiv K(M-N)/M$.

The message vector $t \in \{0,1\}^N$ is encoded to the codeword $x = G^T t \in \{0,1\}^M$ prior to transmission, using the generator matrix $G = [I|(C_2^{-1}C_1)^T]$. This construction ensures that $AG^T = 0 \pmod 2$. Redundancy in the codeword, in the case of unbiased messages, is measured by the rate $R \equiv N/M = 1 - C/K$. After transmission of the codeword through the noisy channel, the following message is received:

$$y = x + n^r, \tag{1}$$

where $n^r \in \mathbb{R}^M$ represents the real channel noise which, in the case of Gaussian and Laplace channels, has the distribution

$$\text{Gaussian:} \quad P(n_i^r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(n_i^r)^2\right], \tag{2}$$

$$\text{Laplace:} \quad P(n_i^r) = \frac{1}{2\lambda} \exp\left[-\frac{|n_i^r|}{\lambda}\right], \tag{3}$$

per bit i. The channel-noise level is measured by the parameters σ and λ, respectively.

Note that since the channel noise $\{n_i^r\}$ consists of real-valued variables (unlike the binary symmetric channel studied in Ref. [5]), the channel output y is also real valued and the evaluation of the syndrome vector cannot be based on simple modulo 2 operations. Decoding is carried out by applying the Bayes rule and the corresponding noise model to calculate $P(x = \pm 1|y)$ for each bit. To bring the problem back to a binary setting, we follow the procedure of Ref. [1], and consider a fictitious channel where the sent message x is corrupted by binary noise $n^f \in \{0,1\}^M$:

$$r = x + n^f \pmod 2. \tag{4}$$

Without loss of generality, one can take $n^f = x$ such that $r = 0$ and $P(n^r) = P(y|x) = P(y|n^f)$.

We denote the set of (fictitious) noise vectors n that satisfy the parity check equations $An = Ar = 0$ by $\mathcal{I}_{pc} = \{n|An = 0\}$, the parity check set. To infer the original message one needs to find the original fictitious noise n^f from the parity

check set on the basis of its statistics. The conditional probabilities of the fictitious noise variables (that satisfy the parity checks) follow from applying the Bayes rule (per bit i):

$$P(n_i|y_i) = \frac{P(y_i|n_i)P(n_i)}{\sum_{n_i'} P(y_i|n_i')P(n_i')}. \qquad (5)$$

It was shown (e.g., in Refs. [3,4]) that this problem can be cast into a statistical mechanical formulation replacing the field $(\{0,1\}, +\bmod 2)$ by the field $(\{1,-1\}, \times)$ and by suitably adapting the parity checks. Using the fact that n_i are Ising variables with prior $P(n_i) = 1/2$, Eq. (5) can now be written as $P(\boldsymbol{n}|\boldsymbol{y}) = \exp[\beta H(\boldsymbol{n})]$ where the *energy* $H(\boldsymbol{n})$ is up to a constant given by

$$H(\boldsymbol{n}) = \sum_i \ln P(n_i|y_i) = \frac{1}{d} \sum_i n_i h_i,$$

$$h_i \equiv \frac{d}{2} \sum_{\tau \in \{-1,1\}} \tau \ln P(y_i|\tau) \qquad (6)$$

and $\beta = 1$ (which corresponds to Nishimori's condition [10]). To unify the notation for the Gaussian and the Laplace channel, we denote the channel degradation parameter variable by d, where $d = \sigma^2, \lambda$ for the Gaussian and Laplace channel, respectively. With these definitions, for any symmetric channel with real outputs, the local energies are staggered magnetizations along the fields h_i, the distribution of which follows from Eq. (6) in combination with Eqs. (2) and (3), respectively:

Gaussian: $\quad p(h_i) = \sqrt{\dfrac{\sigma^2}{2\pi}} \exp\left[-\dfrac{\sigma^2}{2}\left(h_i - \dfrac{1}{\sigma^2} \right)^2 \right]$, (7)

Laplace: $\quad p(h_i) = \dfrac{1}{2}\delta(h_i - \lambda^{-1}) + \dfrac{e^{-2/\lambda}}{2}\delta(h_i + \lambda^{-1})$

$$+ \Theta[\lambda^{-1} - |h_i|]\frac{1}{2}e^{h_i - \lambda^{-1}}, \qquad (8)$$

where $\Theta[x]$ is the Heavyside function returning 1 if $x \geq 0$ and 0 if $x < 0$.

The entropy of the solutions to the decoding problem with a given magnetization $m(\boldsymbol{n};\boldsymbol{h}) = 1/M\sum_i n_i h_i = m$ is

$$\mathcal{M}(m) = \frac{1}{M}\left\langle \ln \sum_{\boldsymbol{n} \in \mathcal{I}_{pc}(\boldsymbol{n},\boldsymbol{n}^f;A)} \delta[m - m(\boldsymbol{n};\boldsymbol{h})] \right\rangle. \qquad (9)$$

Averages in Eq. (9) are taken over the fields $\{h_i\}$ [Eqs. (7) or (8)], the parity check constructions A, whereas the original fictitious noise $\boldsymbol{n}^f$ is gauged away using the transformations $n_i \to n_i n_i^f$ and $y_i \to y_i n_i^f$. In order to perform the averages, we employ the replica identity: $\langle \ln \mathcal{M}(m) \rangle = \lim_{n \to 0} 1/n \ln \langle \mathcal{M}^n(m) \rangle$. In the limit $n \to 0$ and within the replica symmetric assumption (shown to be exact in this case for obtaining the theoretical critical noise levels [5]; for technical details also see, e.g., Refs. [3,9]), we find that

$$\mathcal{M}(m) = \mathrm{Extr}_{\pi,\hat{\pi},\hat{m}} \left\{ -\hat{m}m - \frac{C}{K}\ln 2 + \frac{C}{K}I_1[\pi] - CI_2[\pi,\hat{\pi}] \right.$$

$$\left. + I_3[\hat{\pi};\hat{m}] \right\}, \qquad (10)$$

with

$$I_1[\pi] = \int \prod_{k=1}^{K} \{dx_k \pi(x_k)\} \ln\left[1 + \prod_{k=1}^{K} x_k \right], \qquad (11)$$

$$I_2[\pi,\hat{\pi}] = \int dx d\hat{x} \ \pi(x)\hat{\pi}(\hat{x})\ln(1 + x\hat{x}), \qquad (12)$$

$$I_3[\hat{\pi};\hat{m}] = \int \prod_{c=1}^{C} \{d\hat{x}_c \hat{\pi}(\hat{x}_c)\} \left\langle \ln \sum_{\tau = \pm} e^{\tau\hat{m}h}\prod_{c=1}^{C}(1 + \tau\hat{x}_c) \right\rangle_h. \qquad (13)$$

The functional extremization problem in Eq. (10) results in the following saddle point equations:

$$\hat{\pi}(\hat{x}) = \int \prod_{k=1}^{K-1} \{dx_k \pi(x_k)\}\delta\left[x - \prod_{k=1}^{K-1} x_k \right], \qquad (14)$$

$$\pi(x) = \int \prod_{c=1}^{C-1} \{d\hat{x}_c \hat{\pi}(\hat{x}_c)\} \left\langle \delta\left[x - \tanh\left[\hat{m}h \right.\right.\right.$$

$$\left.\left.\left. + \sum_{c=1}^{C-1} \mathrm{atanh}(\hat{x}_c) \right]\right] \right\rangle_h, \qquad (15)$$

$$m = \int \prod_{c=1}^{C} \{d\hat{x}_c \hat{\pi}(\hat{x}_c)\} \left\langle \frac{h\sum_{\tau = \pm} \tau e^{\hat{m}\tau h}\prod_c(1 + \tau\hat{x}_c)}{\sum_{\tau = \pm} e^{\tau\hat{m}h}\prod_c(1 + \tau\hat{x}_c)} \right\rangle_h. \qquad (16)$$

Equations (14) and (15) are the infinite system equivalent of the so-called density-evolution equations [2], and iteratively converge to the stationary distributions $\pi^*(x)$, $\hat{\pi}^*(x)$. Equations (10) and (16) are then evaluated for these stationary distributions. One should note that the solution $\hat{\pi}^*(\hat{x}) = \delta[\hat{x} - 1]$ and $\pi^*(x) = \delta[x - 1]$ always exists, and has a magnetization $m_0 = \langle h \rangle_h$ (16) and zero entropy. This (ferromagnetic) solution corresponds to perfect retrieval ($\boldsymbol{n}^f$ after the gauge), and should be compared to alternative solutions, if they exist, which correspond to the other (suboptimal) noise candidates in $\mathcal{I}_{pc}(\boldsymbol{n},A)$.

In the limit $K, C \to \infty$ (while keeping the rate R finite) it is possible to derive these alternative solutions for all values of m, analytically. This case is, however, of little practical interest, and will not be discussed in this paper.

For finite K, C, alternative analytic solutions can no longer be obtained for both Gaussian and Laplace channels, and one has to solve the saddle point equations numerically to obtain $\mathcal{M}(m)$.
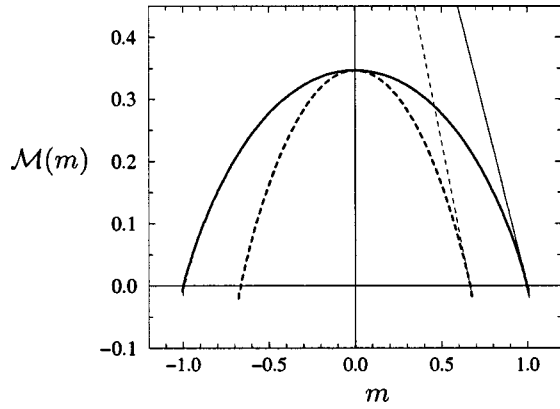
FIG. 1. The magnetization enumerator for the Gaussian (solid curve) and the Laplace (dashed curve) channels for a $(K,C)=(6,3)$ code at the thermodynamic transition noise levels ($\sigma_c^2 =0.899$, $\lambda_c=0.712$). For both channels the maximum number of solutions is $\mathcal{M}(0)=R \ln 2$. The energy of solutions is given by $E(m)=-(1/d)m$, while their free energy at the Nishimori temperature [10] is up to a constant given by the orthogonal distance to the straight lines. At the thermodynamic transition point, these are tangents to $\mathcal{M}(m)$ at $m_\star=m_0$ (=1 and 0.665 for the two channels, respectively).

The critical noise values of the thermodynamic and spinodal transitions now follow directly from the graphs of $\mathcal{M}(m)$, for different values of d.

As explained in full detail in Ref. [5], the theoretical critical degradation value $d=d_c$ is reached, for *maximum a posteriori* and typical set decoding methods, when the magnetization at which the entropy of the suboptimal solutions vanishes, coincides with that of the ferromagnetic solution. For finite temperature decoding at the Nishimori temperature [10], the critical noise value coincides with the thermodynamic transition at which the ferromagnetic solution ceases to be dominant, and $\mathcal{M}(m)$ has a slope $\partial\mathcal{M}(m)/\partial m|_{m=\langle h \rangle_h} =-1/d$ at $m=m_0=\langle h \rangle_h$ (see Fig. 1).

Of more practical interest is the limiting practical noise level $d=d_d$, above which practical algorithms such as density evolution [2] break down. This transition is signaled by the emergence of suboptimal solutions for Eqs. (14) and (15). These correspond to local minima of the free energy in which the algorithm gets trapped; this is known as a spinodal point or dynamical transition. The noise level $d=d_d$ can be

obtained from $\mathcal{M}(m)$; it is the smallest positive value d for which there exists a value m^* such that $\partial\mathcal{M}(m)/\partial m|_{m*}= -1/d$. This typically happens at a value of m^* for which $\mathcal{M}(m^*)<0$, such that the practical transition point is upper bounded by the thermodynamic transition point, which is, in turn, upper bounded by Shannon's information-theoretic limit, and the following inequalities hold: $d_d \leq d_c \leq d_S$.

In Fig. 1 we present the magnetization enumerator for the Gaussian and Laplace channels for a $(K,C)=(6,3)$ code at the thermodynamic transition. The maximum number of solutions to both channels depends only on the code rate R as $\mathcal{M}(0)=R \ln 2$.

It should be noted that values of $\mathcal{M}(m)<0$ are unphysical, and are an artifact of the replica symmetric assumption. Nevertheless, this region turns out to be irrelevant for the determination of the dynamical (spinodal) transition. This can be understood by the fact that the replica symmetric fixed point equations are the exact infinite size equivalent of the practical density evolution equations [2]. Therefore, although quantities related to the $\mathcal{M}(m)<0$ region are unphysical and should be corrected by a refined replica symmetry ansatz, the region $\mathcal{M}(m)<0$ and the fixed point equations associated with it still allow us to determine the dynamical transition point.

The calculated critical noise levels are presented in the Table I for several code constructions and for the Gaussian and Laplace channels. These values are in excellent agreement with those obtained independently in Ref. [8] using a different method, as well as with the corresponding practical upper bounds of Ref. [2] obtained using density evolution. Note that we have also presented the critical degradation values for different code constructions of identical rates to illustrate the opposite tendencies for the theoretical and practical critical values with increasing K and C. Increasing K and C (keeping K/C fixed) pushes the thermodynamic critical value closer to Shannon's information-theoretic limit, but adversely affects the practically admissible degradation value. This is in agreement with the common belief that code constructions with higher connectivity are less practical.

In this paper we have shown how the magnetization enumerator formalism [5] can be easily extended to real-valued channels, in order to obtain both theoretical and practical critical values for the degradation parameter. Following the method presented in Ref. [1] we have mapped the real-

TABLE I. Values of the critical noise levels (spinodal and thermodynamic transitions) of the Gaussian and Laplace channels for various regular (K,C) Gallager codes. For comparison, in the last column we present Shannon's information-theoretic bound.

(K,C)	R	σ_d^2	σ_c^2	σ_S^2	(K,C)	R	λ_d	λ_c	λ_S
(6,3)	0.5	0.775	0.899	0.958	(6,3)	0.5	0.651	0.712	0.752
(5,3)	0.4	1.017	1.253	1.321	(5,3)	0.4	0.773	0.875	0.914
(6,4)	0.333	1.020	1.666	1.681	(6,4)	0.333	0.782	1.045	1.055
(9,6)	0.333	0.379	1.679	1.681	(9,6)	0.333	0.661	1.048	1.055
(4,3)	0.25	1.598	2.325	2.401	(4,3)	0.25	1.018	1.260	1.298
(8,6)	0.25	0.880	2.396	2.401	(8,6)	0.25	0.619	1.271	1.298

valued channel onto an equivalent fictitious binary channel, and employed methods of statistical physics to calculate the magnetization enumerator for the Gallager code ensemble. The magnetization enumerator is instructive in the way it nicely links the various decoding schemes [5] and facilitates the derivation of both practical and theoretical critical noise levels.

Using Nishimori's gauge theory, the theoretical critical noise levels can be shown to be exact, while our practical critical degradation parameters are in excellent agreement with those obtained known in the literature [8,2], when available.

Studying the magnetization enumerator further, beyond the practical limiting noise level, may provide additional insight into the decoding complexity the performance of Gallager-type codes.

---

[1] D.J.C. MacKay, IEEE Trans. Inf. Theory **45**, 399 (1999).

[2] T.J. Richardson and R.L. Urbanke, IEEE Trans. Inf. Theory **47**, 599 (2001).

[3] Y. Kabashima and D. Saad, Europhys. Lett. **45**, 97 (1999).

[4] Y. Kabashima, T. Murayama, and D. Saad, Phys. Rev. Lett. **84**, 1355 (2000).

[5] J. van Mourik, Y. Kabashima, and D. Saad, Phys. Rev. E **66**, 026705 (2002).

[6] A. Montanari, Eur. Phys. J. B **23**, 121 (2001).

[7] S. Franz, M. Leone, A. Montanari, and F. Ricci-Tersenghi, Phys. Rev. E **66**, 046120 (2002).

[8] T. Tanaka and D. Saad, e-print http://xxx.soton.ac.uk/abs/cond-mat/0204554.

[9] K. Wong and D. Sherrington, J. Phys. A **20**, L785 (1987).

[10] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, Oxford, UK, 2001).