Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

Research Article

Information Geometry for Decoding

A. Addadi[1], **K. Abdelmoumen**[2], **H. Ben-Azza**[3*], **I. Chana**[4]

[1] Mathematics and Computer Science Department, Ecole Nationale Superieure des Arts et Metiers, Moulay Ismail University, Meknes, Morocco
[2] Mathematics and Computer Science Department, Ecole Normale Superieure, Sidi Mohamed Ben Abdellah University, Fes, Morocco
[3] Industrial Engineering Department, Ecole Nationale des Arts et Metiers, Moulay Ismail University, Meknes, Morocco
[4] Computer Science Department, Ecole Superieure de Technologie, Moulay Ismail University, Meknes, Morocco
E-mail: h.benazza@umi.ac.ma

Received: 29 July 2024; **Revised:** 6 September 2024; **Accepted:** 28 October 2024

Abstract: We revisit the idea of using elements of information geometry for decoding low-density parity-check (LDPC) codes, as introduced by Ikeda et al. In this work, we explicitly compute the m-projection to an e-flat submanifold, in the case of a binary symmetric channel and the Gaussian channel. We exemplify the algorithm by testing moderate size Gallager codes. To approach decoding problems, we show general theorems based on alternating projections in the framework of information geometry, inspired by von Neumann's theorem for the convergence of alternating projections in Hilbert spaces. More precisely, consider the manifold S of the probability distributions on the n-dimensional hypercube (i.e., the set of binary sequences of length n). Let p be in S. In the case of two intersecting m-flat or e-flat submanifolds, the method of alternating projections on the two submanifolds converges to the projection of p on their intersection. This result is also generalized to a finite family of submanifolds of S.

***Keywords*:** decoding, information geometry, LDPC codes, alternating projections

MSC: 62B11, 94B27, 94B35

1. Introduction

The LDPC codes have generated a rich and varied type of research (construction, performance analysis, etc. see, for example, the references [1–4]). Here, we restrict ourselves to decoding binary codes. Iterative decoding algorithms play a central role in coding theory [5, 6]. Some theoretical works such as [7, 8] investigates the excellent performances of LDPC codes [5, 9] and turbo codes [6], in terms of geometrical concepts. We are especially interested in the fundamental and unifying work of [10] which interprets iterative decoding or belief propagation (BP) algorithms [11] using information geometry [12, 13].

We describe the contributions of this paper. Consider the manifold S of the probability distributions on the n-dimensional hypercube (i.e., the set of binary sequences of length n). The BP algorithm of [10] uses m-projections onto e-flat submanifolds of S (see Section 2.2 for formal definitions). In the first part of our work, we explicitly compute the m-projection to an e-flat submanifold, in the case of a binary symmetric channel and the Gaussian channel. Furthermore, we exemplify the algorithm by testing moderate size Gallager codes. In the second part of this paper, to approach decoding

DOI: https://doi.org/10.37256/cm.6120255392

problems for block codes, we show general theorems based on alternating projections in the framework of information geometry, inspired by von Neumann's theorem for the convergence of alternating projections in Hilbert spaces. More precisely, let p be in S. In the case of two intersecting m-flat or e-flat submanifolds, the method of alternating projections on the two submanifolds converges to the projection of p on their intersection. This result is also generalized to a finite family of submanifolds of S. The usefulness of these results is to allow the design of iterative decoding algorithms taking advantage of information geometry methods. For example, we propose a general methodology in Section 5 to approach the decoding problem, where we include projections to the product submanifold M_0, because the m-projection to M_0 does not change the expectation of a probability distribution.

The method of alternating projections has been widely studied in the framework of Hilbert spaces. It consists of finding a point at the intersection of several closed subspaces by projecting sequentially onto each of the sets (see [14–19]). The first major result relating to the method of alternating projections is due to von Neumann in 1949 [18]. Early works of [20–22] show existence and uniqueness of m or e-projections in a more general setting. It must be noted that the work of [7] relates cross-entropy minimization and iterative decoding for product codes and turbo decoding using mainly e-projections. The works of [23–25] use Dykstra's algorithm and (symmetric) Bregman's projections to capture modified BP algorithms which use extrinsic information.

The rest of the paper is organized as follows. In Section 2, we recall MPM (abbreviation of maximization of the posterior marginals) decoding problem, followed by information geometry concepts. Section 3 presents MPM decoding of LDPC codes, an exact expression for m-projection, and examples to test the BP algorithm. In Section 4, we show some general theorems concerning the method of alternating projections. Section 5 presents a methodology to approach the decoding of linear block codes. Section 6 concludes the paper with brief comments.

2. Related theorems

We formulate in this section the decoding problem of interest, followed by a short background on information geometry; we adopt the notations and conventions of [10] with minor modifications.

2.1 *MPM decoding problem*

For a vector $x = (x_1, \ldots, x_N)^T \in \{-1, +1\}^N$, we consider probability distributions given by:

$$q(x) = C_{\exp}(c_0(x) + \cdots + c_K(x)). \tag{1}$$

The function $c_0(x)$ consists of linear terms, and $c_r(x),\ r = 1, \ldots, K$, consists of higher order terms of the variables $\{x_i\}$. The constant C is the normalization constant. MPM decoding is to estimate the information bits, x, based on $q(x)$. Let $\eta = (\eta_1, \ldots, \eta_N)^T$ be the expectation of x, and $\widetilde{x}$ be the decoded MPM estimator. Then

$$\eta = \sum_x q(x)x, \quad \eta = (\eta_1, \ldots, \eta_N).$$

The sign of each η_i is the decoding result $\widetilde{x_i}$. Let $q(x_i)$ be the marginal distribution of one component x_i in $q(x)$, and let Π denote the operator of marginalization, which maps $q(x)$ to an independent distribution having the same marginal distribution:

$$\Pi \circ q(x) = \prod_{i=1}^{N} q(x_i).$$

The soft bit η_i depends only on the marginal distribution $q(x_i)$. Since $q(x_i)$ is a binary distribution, η_i has one-to-one correspondence to $q(x_i)$. Therefore, soft decoding is equivalent to the marginalization of $q(x)$. Computation of the expectation η is equivalent to the computation of the marginalization operator [10].

2.2 *Information geometry background*

We consider the family of all probability distributions over the variable x. We denote it by:

$$S = \left\{ q / q(x) > 0,\ x \in \{-1,\ +1\}^N,\ \sum_x q(x) = 1 \right\}.$$

We define a submanifold

$$M_0 = \left\{ p_0(x;\ \theta) = \exp\left(c_0(x) + \theta \cdot x - \varphi_0(\theta)\right);\ \theta \in \mathbb{R}^N \right\},$$

where $\varphi_0(\theta)$ is a normalization factor known as free energy. Each component is independent for the distributions of M_0 and $\Pi \circ q(x) = \prod_{i=1}^N q(x_i) \in M_0$. We define e-flat and m-flat submanifold of S. The submanifold $M \subset S$ is said to be e-flat when the following $r(x,\ t)$ belongs to M for all $t \in [0,\ 1]$, $q(x),\ p(x) \in M$ where

$$\ln r(x,\ t) = (1-t)\ln q(x) + t\ln p(x) + c(t),$$

with $c(t)$ a normalization term. The submanifold $M \subset S$ is said to be m-flat when the following $r(x,\ t)$ belongs to M for all $t \in [0,\ 1]$, $q(x),\ p(x) \in M$ where

$$r(x,\ t) = (1-t)q(x) + tp(x).$$

From its definition, we see that M_0 is e-flat. Next, we define m-projection to an e-flat submanifold, after defining the divergence between probability distributions. Let $D[q(x),\ p(x)]$ be the Kullback-Leibleir (KL) divergence for $p,\ q \in S$ defined as

$$D[q(x),\ p(x)] = \sum_x q(x)\ln\frac{q(x)}{p(x)}.$$

The KL divergence is nonnegative and verifies $D[q,\ p] = 0$ if and only if $q = p$. Note that, in general, the KL divergence is not symmetric.

Definition 1 Let M be an e-flat (resp. m-flat) submanifold in S, and let $q(x) \in S$. The distribution in M that minimizes the Kl-divergence from $q(x)$ on M is denoted by

$$\Pi_M^m oq(x) = \underset{p(x)\in M}{argmin} D[q(x),\ p(x)],$$

$$(resp.\ \Pi_M^e oq(x) = \underset{p(x)\in M}{argmin} D[p(x),\ q(x)])$$

and is called the m-projection (resp. e-projection) of $q(x)$ to M.

Theorem 1 [10] Let M be an e-flat (resp. m-flat) submanifold in S and let $q(x) \in S$. The m-projection (resp. e-projection) of $q(x)$ to M is unique.

It is a fundamental fact [10] that the computation of the marginalization operator is equivalent to the m-projection of $q \in S$ to M_0, $\Pi_{M_0} \circ q(x) = \Pi \circ q(x) = \prod_{i=1}^{N} q(x_i)$.

Theorem 2 (Pythagorean theorem [10, 12]). Let $p(x)$, $q(x)$ and $r(x)$ be three distributions in S. Suppose that the m-geodesic connecting $r(x)$ and $q(x)$ is orthogonal at $q(x)$ to the e-geodesic connecting $q(x)$ and $p(x)$. Then we have:

$$D[p(x),\ r(x)] = D[p(x),\ q(x)] + D[q(x),\ r(x)].$$

3. MPM decoding for LDPC codes

In this section we present the relevant information geometry (IG) formulation of [10] for belief propagation algorithm. In the second subsection, we give an explicit expression for computing projections on relevant submanifolds.

3.1 *LDPC codes*

The structure of LDPC codes is shown in Figure 1. Let $s = (s_1,\ s_2,\ \ldots,\ s_M)^T$, $s_i \in \{0,\ 1\}$, be the information bits. The parity check aligned is $H = (h_{ij}) \in \{0,\ 1\}^{K\times N}$. The code $u = (u_1,\ \ldots,\ u_N)^T$ is generated with $G^T s$ mod 2, and u is sent through a channel. We assume a binary symmetric channel BSC with bit error σ, code word u is disturbed and received as $\widetilde{u} = u + x$ mod 2, $x = (x_1,\ \ldots,\ x_N)^T$, $x_i \in \{-1,\ +1\}$, be the noise vector.

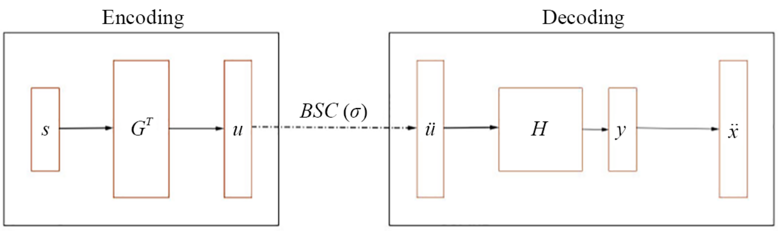

Figure 1. Structure of LDPC code: encoding and decoding

We consider an LDPC code given by its binary control aligned $H = (h_{jr}) \in \{0,\ 1\}^{K\times N}$, over $BSC(\sigma)$, where the bit error rate is $\sigma = (1 - tanh\beta)/2$.

Let $x = (x_1,\ \ldots,\ x_N)^T \in \{-1,\ +1\}^N$ be the noise vector and $\widetilde{y} = (\widetilde{y}_1,\ \ldots,\ \widetilde{y}_K)$ the observed syndrome vector.

The decoding is to infer x such that $\widetilde{y} = Hx = y(x)$ mod $2 = (y_1(x),\ \ldots,\ y_K(x))$, where $y_r(x) = \prod_{j\in L_r} x_j$ and $L_r = \{j : h_{jr} = 1\}$.

Let $c_0(x) = \alpha \cdot x$, where $\alpha = (\beta, \ldots, \beta)$. Then the posterior distribution $p(x \mid \widetilde{y})$ is of the form given in Equation (1). LDPC decoding approximates the computation of $\Pi \circ p(x \mid \widetilde{y})$ by introducing parameterized distributions $p_r(x;\ \varsigma_r)$ for each parity check equation of the code, $\varsigma_r \in \mathbb{R}^N$, $r = 1, \ldots, K$:

$$p_r(x;\ \varsigma_r) = p(x \mid \widetilde{y}_r,\ \varsigma_r) = \exp(c_0(x) + c_r(x) + \varsigma_r \cdot x - \varphi_r(\varsigma_r)) \tag{2}$$

with $c_r(x) = \rho \widetilde{y}_r y_r(x)$, and $\varphi_r(\varsigma_r) = \ln \sum_x \exp(c_0(x) + c_r(x) + \varsigma_r \cdot x)$. The parameter ρ governs the quality of decoding: it is a positive real number which may be taken in the case of regular LDPC to be at least $N/(4KQ)$ where Q is the number per row of 1 s in H (see Appendix II of [10]).

In the language of information geometry, we consider the e-flat submanifold $M_r = \{p_r(x;\ \varsigma_r)\}$, for $r = 1, \ldots, K$ and the product e-flat submanifold:

$$M_0 = \left\{p_0(x;\ \theta) = \exp(c_0(x) + \theta \cdot x - \varphi_0(\theta));\ \theta \in \mathbb{R}^N\right\} \tag{3}$$

to which the M_r will project.

3.2 *Explicit m-projection computation*

For the purpose of implementing belief propagation algorithm, information geometry version, we explicitly compute the m-projection of $p_r(x,\ \varsigma_r)$, $r = 1, \ldots, K$, to the e-flat submanifold M_0 in the case of a binary symmetric channel $BSC(\sigma)$.

Theorem 3 On a $BSC(\sigma)$, with $\beta = arctanh(1 - 2\sigma)$, if $\theta = \theta(r) = \Pi^m_{M_0} \circ p_r(x;\ \varsigma_r)$, then

$$\theta_i = arctanh\left(\sum_x p_r(x,\ \varsigma_r) x_i\right) - \beta,\ i = 1, \ldots, N \tag{4}$$

Proof. We have

$$\begin{aligned}
\theta &= \Pi^m_{M_0} o p_r(x;\ \varsigma_r) \\
&= arg \min_{\theta \in \mathbb{R}^N} D[p_r(x;\ \varsigma_r),\ p_0(x;\ \theta)] \\
&= arg \min_{\theta \in \mathbb{R}^N} \sum_x \left(p_r(x,\ \varsigma_r) \ln \frac{p_r(x,\ \varsigma_r)}{p_0(x;\ \theta)}\right) \\
&= arg \min_{\theta \in \mathbb{R}^N} \sum_x (p_r(x,\ \varsigma_r) \ln p_r(x,\ \varsigma_r)) - arg \min_{\theta \in \mathbb{R}^N} \sum_x (p_r(x,\ \varsigma_r) \ln p_0(x;\ \theta)) \\
&= arg \min_{\theta \in \mathbb{R}^N} \left(-\sum_x (p_r(x,\ \varsigma_r) \ln p_0(x;\ \theta))\right)
\end{aligned}$$

$$= arg \min_{\theta\in\mathbb{R}^N} \left(-\sum_x \left(p_r\left(x,\ \varsigma_r\right)\left(c_0(x)+\theta\cdot x-\varphi_0(\theta)\right)\right)\right)$$

$$= arg \min_{\theta\in\mathbb{R}^N} \left(\sum_x \left(p_r\left(x,\ \varsigma_r\right)\left(\varphi_0(\theta)-\theta\cdot x\right)\right)\right).$$

Let $g_x(\theta)=\varphi_0(\theta)-\theta\cdot x$ and $G=G(\theta)=\sum_x\left(p_r\left(x,\ \varsigma_r\right)g_x(\theta)\right)$. Since the free energy function φ_0 is convex, then g_x is convex (it is the difference of two convex functions). Then we have the following:

$$\frac{\partial G(\theta)}{\partial\theta}=\sum_x p_r\left(x,\ \varsigma_r\right)\frac{\partial}{\partial\theta}\left(g_x(\theta)\right)$$

$$\frac{\partial G(\theta)}{\partial\theta_i}=0\Leftrightarrow\sum_x p_r\left(x,\ \varsigma_r\right)\frac{\partial}{\partial\theta_i}\left(\varphi_0(\theta)\right)=\sum_x p_r\left(x,\ \varsigma_r\right)\frac{\partial}{\partial\theta_i}(\theta\cdot x)$$

$$\Leftrightarrow\frac{\partial}{\partial\theta_i}\left(\varphi_0(\theta)\right)=\sum_x p_r\left(x,\ \varsigma_r\right)x_i.$$

Let $\alpha=(\beta,\ \ldots,\ \beta)\in\mathbb{R}^N$, and $c_0(x)=\alpha\cdot x$. We make the variable change:

$$\theta_{\text{old}}=\theta+\alpha \text{ and } \varphi_0(\theta)=\varphi(\theta+\alpha).$$

We have

$$\varphi\left(\theta_{\text{old}}\right)=\sum_{i=1}^N\varphi\left(\theta_i\right)=\sum_{i=1}^N ln\left(e^{-\theta i}+e^{\theta i}\right),\ \theta_{old}\in\mathbb{R}^N$$

and

$$\partial\theta_i\varphi\left(\theta_{\text{old}}\right)=\frac{\left(e^{\theta i}-e^{-\theta i}\right)}{\left(e^{\theta i}+e^{-\theta i}\right)}=tanh\left(\left(\theta_{\text{old}}\right)_i\right),$$

$$tanh\left(\left(\theta_{\text{old}}\right)_i\right)=\sum_x p_r\left(x,\ \varsigma_r\right)x_i.$$

Thus

$$\left(\theta_{old}\right)_i=arctanh\left(\sum_x p_r\left(x,\ \varsigma_r\right)x_i\right).$$

Finally,

$$\theta_i = arctanh\left(\sum_x p_r(x,\ \varsigma_r)x_i\right) - \beta,\ i = 1,\ \ldots,\ N.$$

$\square$

With the same argument, we have:

Theorem 4 On a binary white Gaussian additive channel, with $\sigma^2 = \dfrac{N_0}{2}$ representing the variance of noise and $\beta = \dfrac{1}{\sigma^2}$, if $\theta = \Pi^m_{M_0} op_r(x;\ \varsigma_r)$, then

$$\theta_i = arctanh\left(\sum\nolimits_x p_r(x,\ \varsigma_r)x_i\right) - \beta,\ i = 1,\ \ldots,\ N \tag{5}$$

Now, the interpretation of LDPC decoding by information geometry [10] is: Initialization: For $t = 0$; set $\xi_r^t = 0,\ \varsigma_r^t = 0,\ (r = 1,\ \ldots,\ K)$.

For $t = 0,\ 1,\ 2,\ \ldots$ do compose $p_r(x,\ \varsigma_r^t) \in M_r$,

Horizontal step: Compute the m-projection θ^t of $p_r(x,\ \varsigma_r^t)$ to M_0 with

$$\theta_i^t = arctanh\left(\sum_x p_r\left(x,\ \varsigma_r^t\right)x_i\right) - \beta\ i = 1,\ \ldots,\ N$$

and define ξ_r^{t+1} by $\xi_r^{t+1} = \theta^t - \varsigma_r^t,\ r = 1,\ \ldots,\ K$.

Vertical step: Update $\left\{\varsigma_r^{t+1}\right\},\ \varsigma_r^{t+1} = \theta^{t+1} - \xi_r^{t+1};\ r = 1,\ \ldots,\ K$.

Convergence: if θ^t does not converge (that is $\theta^t \neq \theta^{t+1}$), repeat the steps by incrementing t by 1.

3.3 *Examples*

We give examples of LDPC decoding via IG to test its quality and efficiency based on the computation of Theorem 3 for small dimensions of LDPC matrices H. The algorithms are implemented in C language.

Example 1 We consider the regular parity check aligned with small number of one's $K = 15,\ N = 20$, and a maximum of 100 iterations; the number of frames is 10 (Gallager's LDPC code construction):

$$
\begin{pmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

Let $s = (1;\ 1;\ 1;\ 1;\ 1)^T$ be the information vector, note that 0 in the binary form correspond to +1 in the bipolar form and 1 correspond to -1, and vice versa.

Let $u = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)^T$ be the encoded message, and $\widetilde{u} = u + x \bmod 2$. Since $sign(\eta_i) = sign(\theta_i)$ we have:

$$
\widetilde{x_i} = \begin{cases} 1 & \text{if } \theta_i > 0 \\ \\ -1 & \text{if } \theta_i < 0. \end{cases}
$$

For example, if $\rho = 3$ we construct Table 1, where $\widetilde{u}$ is the received word and $\widetilde{x}$ is the decoded word:

Table 1 shows the word $\widetilde{x}$ for different σ if $\rho = 3$ and $u = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$.

For example if $\sigma = 0.3$ we have $\widetilde{u} = (1;\ 1;\ 1;\ -1;\ 1;\ 1;\ 1;\ -1;\ -1;\ -1;\ 1;\ -1;\ -1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$ and $\widetilde{x} = (-1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1;\ -1)$, if $\sigma = 0.2$ we have $\widetilde{u} = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ -1;\ 1;\ -1;\ 1;\ 1;\ 1;\ -1;\ 1;\ -1;\ 1;\ 1;\ 1;\ 1)$ and $\widetilde{x} = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$.

The efficiency of the decoder $(15,\ 20)$ LDPC depends on the parameter ρ.

If we change ρ, for example if $\sigma = 0.3$, $\rho = 3$ and $u = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$ we have $\widetilde{u} = (1;\ 1;\ 1;\ -1;\ 1;\ 1;\ 1;\ -1;\ -1;\ -1;\ 1;\ -1;\ -1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$ and $\widetilde{x} = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1)$.

Table 1. Decoder $(15,\ 20)$ LDPC code to recover the information word fixing $\rho = 3$

σ	$\widetilde{u}$	$\widetilde{x}$	Number of iterations
0.3	(1; 1; 1; −1; 1; 1; 1; −1; −1; −1; 1; −1; −1; 1; 1; 1; 1; 1; 1; 1)	(−1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1)	1
0.29	(−1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; −1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; −1)	1
0.28	(−1; −1; −1; −1; 1; 1; 1; 1; −1; 1; −1; −1; −1; −1; 1; 1; 1; 1; −1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	53
0.27	(1; 1; 1; 1; 1; −1; 1; 1; 1; 1; −1; 1; 1; 1; −1; 1; 1; −1; −1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1)	30
0.26	(1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1; −1; −1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	55
0.25	(1; −1; 1; −1; 1; 1; 1; 1; −1; 1; 1; −1; −1; 1; 1; 1; −1; 1; 1; 1)	(−1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1)	1
0.24	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; −1; −1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	58
0.23	(1; 1; −1; 1; 1; 1; −1; 1; 1; 1; 1; 1; −1; 1; 1; −1; 1; −1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	100
0.22	(1; 1; −1; 1; 1; 1; −1; 1; 1; 1; 1; −1; 1; 1; 1; 1; −1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	100
0.21	(1; 1; 1; 1; −1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1)	100
0.2	(1; 1; 1; 1; 1; 1; 1; −1; 1; −1; 1; 1; 1; −1; 1; −1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	16
0.19	(−1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; −1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	21
0.18	(1; 1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	100
0.17	(−1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; −1; −1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	3
0.16	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; −1; 1; 1; −1; 1; 1; 1; 1)	(−1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1)	23
0.15	(1; −1; 1; −1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	6
0.14	(1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	39
0.13	(1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	24
0.12	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	7
0.11	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	(−1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1; −1)	1
0.1	(1; −1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; −1; 1; 1; 1; 1; −1; 1; 1)	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	33

Table 2 shows the word $\widetilde{x}$ if we change ρ.

Table 2. The word $\widetilde{x}$ of a $(15,\ 20)$ LDPC decoder if we change ρ

σ	$\widetilde{x}$	ρ	Number of iterations
0.3	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	1
0.28	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	1
0.27	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	23
0.25	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	72
0.21	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	0.5	54
0.16	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	48
0.11	(1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1)	1	6

Often, the quality of decoding is measured by the bit error rate (BER) defined as:

$$\text{BER} = \frac{\text{Number of bit errors}}{(N-K) \times \text{ Number of frames}} \tag{6}$$

For 10 frames we have Table 3:

Table 3. BER for 10 frames and 100 iterations

ρ	σ			
	0.5	1	2	3
0.3	1	0	0	0.5
0.29	0.7	0	0.8	0
0.28	0	0	0	0
0.27	0	0	0	0
0.26	0.1	0	0.1	0
0.25	0	0.1	0	0.3
0.24	0	0	0	0
0.23	0	0	0	0
0.22	0	0	0	0
0.21	0	0	0	0.3

Table 3 shows the convergence and efficiency of LDPC decoding by information geometry for different values of σ depending on the aligned H and the value of ρ. Recall that the parameter ρ intervenes in the local distributions as:

$$p_r(x;\ \varsigma_r) = p(x \mid \widetilde{y}_r,\ \varsigma_r) = \exp(c_0(x) + c_r(x) + \varsigma_r \cdot x - \varphi_r(\varsigma_r)),\ r = 1,\ \ldots,\ K$$

with, $c_r(x) = \rho \widetilde{y}_r \cdot y_r(x),\ \rho \in \mathbb{R}$, and $\rho > 0$. For example, for $\sigma = 0.26$ if $\rho = 0.5$, the BER for 10 frames of one decoder is 0.1 and if $\rho = 1$ the BER is 0.

Example 2 We consider the regular parity check aligned has small number of one's, $K = 10,\ N = 15.$ (Gallager's LDPC code construction):

$$\begin{pmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

$s = (1;\ 1;\ 1;\ 1;\ 1)^T$ is the message sent, $u = (1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;)^T$ is the encoded message, for 1,000 frames we have Table 4.

Table 4. BER for 1,000 frames and 100 iterations

ρ	σ			
	0.5	1	2	3
0.3	1	0.736	0.518	1
0.29	0	0.003	0.87	0.243
0.28	1	0.281	1	0.908
0.27	0.437	0.71	0.108	0
0.26	0.299	1	1	1
0.25	0	0	0	0.7
0.24	1	0.31	1	0.923
0.23	0	0.63	0	0.63
0.22	0.001	0.064	0.001	0.178
0.21	0.014	0	0.32	0.014

Table 4 shows the convergence and efficient of LDPC decoding by information geometry if we change the parameter ρ for different σ. For example, for $\sigma = 0.29$ if $\rho = 1$, the BER of one decoder is 0.003, and if $\rho = 0.5$, the BER is 0.

Example 3 For a Gallager's LDPC code of size $(20,\ 25)$ with $\rho = 1$, let $s = (1;\ 1;\ 1;\ 1;\ 1)^T$ be the information vector; we obtain Table 5.

Table 5. Number of bit errors with $(20,\ 25)$ LDPC decoder for $\rho = 1$

σ	Number of bit errors
0.2	0
0.19	0
0.18	0
0.17	0
0.16	0
0.15	0
0.14	0
0.13	0
0.12	0
0.11	0

Table 5 shows the number of bit errors of Gallager's LDPC code $(20,\ 25)$ for different σ and for $\rho = 1$.

Remark 1 The same implementation is valid for turbo decoding with $K = 2$. As shown in [10], we can also compute the error correction for LDPC aligned, which vanishes for large girths. That is, any two columns of the parity check aligned have at most one overlapping positions of 1.

4. Alternating projections: Information geometry view

This section concerns the derivation of general results in the framework of the information geometry of the manifold S, where Theorems 1 and 2 are extensively used.

The following proposition shows that projections on the intersection are the same (recall that the KL divergence is asymmetric).

Proposition 1 Suppose that M_1, M_2 are two submanifolds such that M_1 is e-flat and M_2 is m-flat in S with $M_1 \cap M_2 \neq \varnothing$. Let $p_0 \in S$.

Then we have $\Pi^m_{M_1 \cap M_2} o p_0 = \Pi^e_{M_1 \cap M_2} o p_0$, where $\Pi^m_{M_1 \cap M_2} o p_0$ is the m-projection of p_0 to $M_1 \cap M_2$ and $\Pi^e_{M_1 \cap M_2} o p_0$ is the e-projection of p_0 to $M_1 \cap M_2$.

Proof. Let $p^* = \Pi^m_{M_1 \cap M_2} o p_0$ and $q^* = \Pi^e_{M_1 \cap M_2} o p_0$. Then, for all $p \in M_1 \cap M_2$, by the Pythagorean theorem we have:

$$D\left[p_0,\ q^*\right] + D\left[q^*,\ p\right] = D\left[p_0,\ p\right].$$

By taking $p = p^*$, we have:

$$D\left[p_0,\ q^*\right] + D\left[q^*,\ p^*\right] = D\left[p_0,\ p^*\right] \tag{7}$$

□

Similarly, for all $q \in M_1 \cap M_2$, by the Pythagorean theorem we have:

$$D\left[p_0,\ p^*\right] + D\left[p^*,\ q\right] = D\left[p_0,\ q\right] \tag{8}$$

If $q = q^*$, we have:

$$D\left[p_0,\ p^*\right] + D\left[p^*,\ q^*\right] = D\left[p_0,\ q^*\right] \tag{9}$$

By using Equation (7) + Equation (8), we have:

$$D\left[p_0,\ q^*\right] + D\left[q^*,\ p^*\right] + D\left[p_0,\ p^*\right] + D\left[p^*,\ q^*\right] = D\left[p_0,\ p^*\right] + D\left[p_0,\ q^*\right].$$

Then,

$$D\left[q^*,\ p^*\right] + D\left[p^*,\ q^*\right] = 0.$$

But $D\left[q^*,\ p^*\right] \geq 0$ and $D\left[p^*,\ q^*\right] \geq 0$. Thus, we deduce $p^* = q^*$.

Theorem 5 Suppose that M_1, M_2 are two submanifolds such that M_1 is e-flat and M_2 is m-flat in S where M_1 corresponds to m-projections Π^m and M_2 corresponds to e-projections Π^e. Suppose that $M_1 \cap M_2 \neq \varnothing$ and $p_0 \in S$. Then the sequence of alternating projections generated by:

$$p_1(x) = \Pi^m_{M_1} o p_0(x);\ q_1(x) = \Pi^e_{M_2} o p_1(x);\ p_2(x) = \Pi^m_{M_1} o q_1(x),\ \ldots$$

converges to a point p^* equal to p^{**} the m-projection of p_0 onto $M_1 \cap M_2$ (see Figure 2).

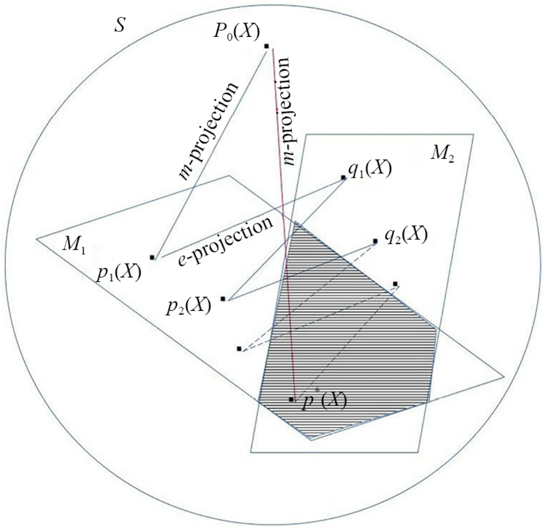

Figure 2. Von Neumann's theorem in S for an e-flat submanifolds M_1 and an m-flat submanifold M_2 (Information geometry view)

Proof. Let $p_0 \in S$ and p_1 the m-projection of p_0 to M_1. For $t = 1, 2, \ldots, p_t \in M_1$, search for $q_t \in M_2$ that minimizes $D[q_t, p_t]$, this is given by the e-projection of p_t to M_2, let it be $q_t \in M_2$; then search for the point p_{t+1} in M_1 that minimizes $D[q_t, p_{t+1}]$, let it be p_{t+1}; this is given by the m-projection of q_t to M_1. Thus, we have

$$q_t = \Pi^e_{M_2} o p_t(x) = \underset{q \in M_2}{argmin} D[q(x), p_t(x)],$$

$$p_{t+1} = \Pi^m_{M_1} o q_t(x) = \underset{p \in M_1}{argmin} D[q_t(x), p(x)].$$

Let p^* and q^* be the pair of minimizers of $D[q_t, p_{t+1}]$, then the m-projection of q^* to M_1 is p^* and the e-projection of p^* to M_2 is q^*. Since we have

$$D[q_t, p_t] \geq D[q_t, p_{t+1}] \geq D[q_{t+1}, p_{t+1}] \geq \ldots \geq D[p^*, q^*],$$

$$D[p^*, q^*] = \underset{p \in M_1}{argmin} \underset{q \in M_2}{argmin} D[p, q],$$

and $p^* \in M_1 \cap M_2$, $q^* \in M_1 \cap M_2$, we deduce that $D[p^*, q^*] = 0$, and consequently $p^* = q^*$. As a conclusion of the first part of the proposition, the sequence of alternating projections converges to p^*.

Let p^{**} be the m-projection of p_0 to $M_1 \cap M_2$ and $p^{**} = \Pi^m_{M_1 \cap M_2} o p_0$. For all $p \in M_1 \cap M_2$, by the Pythagorean theorem we have:

$$D[p_0, p^{**}] + D[p^{**}, p] = D[p_0, p] \tag{10}$$

If we set $p = p^*$, in Equation (10), we have:

$$D[p_0,\ p^{**}]+D[p^{**},\ p^*]=D[p_0,\ p^*] \tag{11}$$

where $p^* \in M_1 \cap M_2$ and $p^{**} \in M_1 \cap M_2$. Then $D[p^{**},\ p^*]=0$, and $D[p_0,\ p^{**}]=D[p_0,\ p^*]$. We deduce that $p^*=p^{**}=\Pi^m_{M_1\cap M_2} o p_0$. □

The three points $p_0,\ p^*$ and p form an orthogonal triangle, because the m-geodesic connecting p^* and p_0 is orthogonal to the e-geodesic connecting p and p^*. Hence, the Pythagorean theorem shows $D[p_0,\ p^*]+D[p^*,\ p]=D[p_0,\ p]$. We conclude that for $p_0 \in S$, the point p^* is the m-projection of p_0 to $M_1 \cap M_2$.

With the same argument, we have:

Theorem 6 Suppose that $M_1,\ M_2$ are two submanifolds such that M_1 is m-flat and M_2 is e-flat in S where M_1 corresponds to e-projections Π^m and M_2 corresponds to m-projections Π^e. Suppose that $M_1 \cap M_2 \neq \varnothing$ and $p_0 \in S$. Then the sequence of alternating projections generated by:

$$p_1(x)=\Pi^e_{M_1} o p_0(x);\ q_1(x)=\Pi^m_{M_2} o p_1(x);\ p_2(x)=\Pi^e_{M_1} o q_1(x),\ \ldots$$

converges to a point p^* equal to p^{**} the e-projection of p_0 to $M_1 \cap M_2$.

Theorem 7 [7] Suppose that $M_1,\ M_2$ are two m-flat submanifolds in S corresponding to e-projections Π^e. Suppose that $M_1 \cap M_2 \neq \varnothing$ and $p_0 \in S$. Then the sequence of alternating e-projections converges to p^* the eprojection of p_0 on $M_1 \cap M_2$.

Proof. Let $p_0 \in S$ and p_{t+1} the e-projection of p_t to $M_i,\ i=1,\ 2$:

$$\Pi^e_{M_i} o p_t(x)=\underset{p_{t+1}(x)\in M_i}{argmin}\ D[p_{t+1}(x),\ p_t(x)]$$

We have:

$$D[p_1,\ p_0]\geq D[p_2,\ p_1]\geq \ldots \geq D[p_{t+1},\ p_t]$$

This implies:

$$D[p^*,\ p_t]\geq D[p^*,\ p_{t+1}],\ p^* \in M_1 \cap M_2.$$

When $t\to\infty$, $D[p^*,\ p_{t+1}]=0$. □

Since the e-projection p^* of p_0 is unique and satisfies the Pythagorean theorem, p^* is the e-projection of p_0 to $M_1 \cap M_2$.

Remark 2 The m-flat submanifolds are convex sets.

Theorem 8 Suppose that $M_1,\ M_2$ are two e-flat submanifolds in S corresponding to m-projections Π^m. Suppose that $M_1 \cap M_2 \neq \varnothing$ and $p_0 \in S$. Then the sequence of alternating m-projections converges to p^*, the m-projection of p_0 to $M_1 \cap M_2$.

Proof. Let $p_0 \in S$ and let p_{t+1} be the m-projection of p_t to $M_i,\ i=1,\ 2$:

$$\Pi_{M_i}^m o p_t(x) = \underset{p_{t+1}(x) \in M_i}{argmin}\ D[p_t(x),\ p_{t+1}(x)]$$

We have

$$D[p_0,\ p_1] \geq D[p_1,\ p_2] \geq \ldots \geq D[p_t,\ p_{t+1}]$$

This implies

$$D[p_t,\ p^*] \geq D[p_{t+1},\ p^*],\ p^* \in M_1 \cap M_2.$$

When $t \to \infty$, $D[p_{t+1},\ p^*] = 0$. Since the m-projection p^* of p_0 is unique and satisfies the Pythagorean theorem, p^* is the m-projection of p_0 to $M_1 \cap M_2$. □

Theorem 9 Let $M_1,\ M_2,\ \ldots,\ M_r$ be m-flat submanifolds of the manifold S with $M = \cap_{i=1}^r M_i \neq \varnothing$.

Let $q_0 \in S$ and consider the sequence $q_1,\ q_2,\ \ldots$ defined by

$$q_n = \Pi_{M_n}^e o q_{n-1},\ \text{ for } n = 1,\ \ldots,\ r-1\ ;$$

$$q_n = \Pi_{M_n mod r}^e o q_{n-1},\ \text{ for } n \geq r\ .$$

Then q_n converges to the e-projection q^* of q_0 to M.

Proof. Under the hypotheses of the theorem, it follows that the e-projections $q_1,\ q_2,\ \ldots$ and q^* exist and $D[p,\ q_{n-1}] = D[p,\ q_n] + D[q_n,\ q_{n-1}]$ for any $p \in M_n,\ n = 1,\ 2,\ \ldots$. In particular, setting $p = q^*$, we obtain by induction:

$$D[q^*,\ q_0] = D[q^*,\ q_n] + \sum_{i=1}^{n} D[q_i,\ q_{i-1}],\ n = 1,\ 2,\ \ldots$$

□

By the same argument, we obtain:

Theorem 10 Let $M_1,\ M_2,\ \ldots,\ M_r$ be e-flat submanifolds of the manifold S with $M = \cap_{i=1}^r M_i \neq \varnothing$. Let $q_0 \in S$ and consider the sequence $q_1,\ q_2,\ \ldots$ defined by

$$q_n = \Pi_{M_n}^m o q_{n-1},\ \text{ for } n = 1,\ \ldots,\ r-1\ ;$$

$$q_n = \Pi_{M_n mod r}^m \text{ o } q_{n-1},\ \text{ for } n \geq r.$$

Then q_n converges to the m-projection q^* of q_0 to M.

Theorem 11 Let $M_1,\ \ldots,\ M_k$ be a set of submanifolds, where M_i is e-flat or m-flat, with $M = \cap_{i=1}^k M_i \neq \varnothing$. Let σ be a permutation on $\{1,\ \ldots,\ k\}$,

$$\sigma = \begin{pmatrix} 1 & \dots & k \\ \\ i_1 & \dots & i_k \end{pmatrix}$$

Set

$$f_i = \begin{cases} m \text{ if } M_i \text{ is } e- \text{ flat} \\ & \text{for } i = 1, \dots, k \\ e \text{ if } M_i \text{ is } m- \text{ flat} \end{cases}$$

Then, if $p_0 \in S$, we have

$$\lim_{n \to \infty} \left(\Pi_{M_{i_k}}^{f_{i_k}} o \dots o \Pi_{M_{i_1}}^{f_{i_1}} \right)^n o p_0 = \Pi_M^m o p_0$$

Proof. Let $D_{\min} = \min_{i \neq j} D[M_i, M_j]$, then

$$\lim_{n \to \infty} \left(\Pi_{M_{i_k}}^{f_{i_k}} o \dots o \Pi_{M_{i_1}}^{f_{i_1}} \right)^n o p_0 = D_{\min}(M_i)$$

Since $M \neq \varnothing$, we get $D_{\min} = 0$. Then, there exists $p^* \in M$, such that, for all $p \in M_i$, $D[p, p^*] = 0$. Set $\Pi_M^m o p_0 = p^{**}$. For all $p \in M$, by Pythagorean theorem, we have:

$$D[p_0, p^{**}] + D[p^{**}, p] = D[p_0, p].$$

If $p = p^*$ we have:

$$D[p_0, p^{**}] + D[p^{**}, p^*] = D[p_0, p^*] \tag{12}$$

Then $D[p^{**}, p^*] = 0$ and $D[p_0, p^{**}] = D[p_0, p^*]$.
We conclude that $p^* = p^{**} = \Pi_M^m o p_0$. □

5. Interpretation of iterative decoding

The aim of this section is to show the possibility to devise decoding techniques for block codes, based on the general results of the preceding section. Recall from Section 2.2 that the computation of the marginalization operator is equivalent to the m-projection of $q \in S$ to a product submanifold M_0.

We consider two submanifolds in S, M_1 and M_2, where M_1 is m-flat and M_2 is e-flat of the form:

$$M_1 = \{p/\sum_x p(x)f_i(x) = \gamma_i,\ i = 1,\ \ldots,\ K\} \tag{13}$$

and

$$M_2 = \left\{p/p(x) = Cq(x)\exp\left(\sum_{i=1}^K \theta_i f_i(x)\right)\right\} \tag{14}$$

In M_2, q is a given distribution and C is a normalization factor. The m-flat submanifold M_1 is completely defined by the functions $f_1,\ \ldots,\ f_K$ and the scalars $\gamma_1,\ \ldots,\ \gamma_K$. Similarly, the e-flat submanifold M_2 is completely defined by the distribution q, the functions $f_1,\ \ldots,\ f_K$ and the parameters θ_i.

The m-projection p^* of $q \in S$ to M_2 is unique and satisfies the Pythagorean identity:

$$D[p(x),\ q(x)] = D\left[p(x),\ p^*(x)\right] + D\left[p^*(x),\ q(x)\right].$$

The e-projection q^* of p onto M_1 is unique and satisfies the Pythagorean identity:

$$D[p(x),\ q(x)] = D\left[p(x),\ q^*(x)\right] + D\left[q^*(x),\ q(x)\right].$$

The e-projection q^* of q onto the m-flat submanifold M_1 is given by:

$$q^*(x) = q(x)\exp\left(-\sum_{i=1}^K \mu_i\left(f_i(x) - \gamma_i\right)\right)$$

where the $\{\mu_i\}$ are Lagrange multipliers determined from the constraints.

In the decoding problem the functions $\{f_i\}$ correspond to the parity-check equations of the code and if $\{\gamma_i = 0\}$, then the e-projection is given by [7]:

$$q^*(x) = q(x)\exp\left(-\mu_0\right) I_1(x)\ldots I_K(x)$$

where $I_i(x)$ is the indicator function and $\exp\left(-\mu_0\right)$ is a normalization constant:

$$I_i(x) = \begin{cases} 1 \text{ if } f_i(x) = 0 \\ \\ 0 \text{ if } f_i(x) \neq 0. \end{cases}$$

The m-projection p^* of p onto the e-flat submanifold M_2 is given by:

$$p^*(x) = \prod_i p_i(x_i)$$

where $p_i(x_i)$ is the marginal distribution on x_i. In the following theorem, the notation $p(x,\ \theta)$ signify a probability distribution with e-affine coordinates θ.

Theorem 12 We consider two m-flat submanifolds M_1, M_2, and an e-flat product manifold M_0 in S given by:

$$M_1 = \left\{p/\sum_x p(x) f_i(x) = 0,\ i = 1,\ \ldots,\ r\right\},$$

$$M_2 = \left\{p/\sum_x p(x) f_i(x) = 0,\ i = r+1,\ \ldots,\ K\right\},$$

$$M_0 = \left\{p(x,\ \theta)/p(x,\ \theta) = \prod_{i=1}^{K} p(x_i,\ \theta_i)\right\}.$$

Let $p_0 \in S$ and p_1 the m-projection of p_0 to M_1. Suppose that $M_1 \cap M_2 \neq \varnothing$. Then the sequence of alternating projections generated by:

$$p_1(x) = \Pi^e_{M_1} o p_0(x);\ p_1^{'}(x) = \Pi^m_{M_0} o p_1(x);\ q_1(x) = \Pi^e_{M_2} o p_1^{'}(x)\ ;$$

$$q_1^{'}(x) = \Pi^m_{M_0} o q_1(x);\ p_2(x) = \Pi^e_{M_1} o q_1^{'}(x)\ldots$$

converges to $p^* \in M_1 \cap M_2$.

Proof. Let $p_0 \in S$ and p_1 the e-projection of p_0 to M_1. We have:

$$D\left[p_1,\ p_1^{'}\right] + D\left[q_1,\ p_1^{'}\right] \geq D\left[p_2,\ p_2^{'}\right] + D\left[q_2,\ p_2^{'}\right] \geq \ldots \geq D\left[p_t,\ p_t^{'}\right] + D\left[q_t,\ p_t^{'}\right]$$

when $t \to \infty$, $D\left[p_t,\ p_t^{'}\right] + D\left[q_t,\ p_t^{'}\right] = D[p_t,\ p^*] + D\left[p^*,\ p_t^{'}\right]$, with $p^* \in M_1 \cap M_2$.

Then,

$$\Pi^e_{M_1 \cap M_2} o p_t^{'}(x) = \Pi^e_{M_1 \cap M_2} o p_t(x) = p^*.$$

□

Corollary 1 With the same notations as the above theorem, let $p^{**} = \Pi^m_{M_0} op$. Then, we have $p^{**} = \Pi^m_{M_0} op^*$.

Indeed, we have $D[p^*,\ p^{**}] + D[p^{**},\ p_0] = D[p^*,\ p_0]$.

Let us consider an error-correcting code C given by its $K \times N$ parity-check aligned $H = H(C)$. The rows of H correspond to linear forms. We may associate to the code C a submanifold in S, by the operator $\mathscr{T}(C) = \mathscr{T}(H)$ of the form (13) defined by

$$\mathscr{T}(C) = \left\{ p / \sum_x p(x) f_i(x) = 0,\ i = 1,\ \ldots,\ K \right\}$$

where the f_i are the linear forms of H and the γ_i are null. Let $1 \leq r \leq K$. We partition the rows as

$$H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}$$

where H_1 is an $r \times N$ subaligned and H_2 is an $(K - r) \times N$ subaligned of H. Then $C = C_1 \cap C_2$ is non empty, by construction, and the codes C_1, C_2 correspond to the submatrices H_1, H_2 respectively. Thus, since $\mathscr{T}(C_1) \cap \mathscr{T}(C_2) \neq \varnothing$, it is possible to apply Theorem 12.

6. Conclusions

In this paper, our main sources of inspiration are (1) the von Neumann's theorem [19] on the convergence of alternating projection method in the case of Hilbert spaces, and (2) the work of Ikeda et al. [10] interpreting belief propagation algorithm in the frame of information geometry (IG). After a presentation of LDPC decoding, version IG, we explicitly compute the m-projection to an e-flat submanifold, in the case of a binary symmetric channel and the Gaussian channel. Moreover, we give moderate tests by implementing the algorithm. In a second part, we give general results on the convergence of IG alternating projections on e-flat and m-flat submanifolds. Towards decoding problems [7], in Section 5, we see how to transform the decoding problem by introducing convenient submanifolds.

Further work include (1) the study of the rate of convergence of our proposals by taking account of the angle between submanifolds (see [14, 15, 26]) in our case; (2) the quality of decoding linear block codes and performance analysis; (3) the search for the types of error-correcting codes adapted to IG alternating projections by using Theorem 11 or 12 for example.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Andreev K, Frolov A, Svistunov G, Wu K, Liang J. Deep neural network based decoding of short 5G LDPC codes. In: *2021 XVII International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY)*. Moscow, Russian: IEEE; 2021.

[2] Kruglik S, Potapova V, Frolov A. On performance of multilevel coding schemes based on non-binary LDPC codes. *European Wireless 2018; 24th European Wireless Conference*. Catania, Italy: IEEE; 2018.

[3] Tang H, Xu J, Kou Y, Lin S, Abdel-Ghaffar K. On algebraic construction of Gallager and circulant low-density parity-check codes. *IEEE Transactions on Information Theory*. 2004; 50(6): 1269-1279.

[4] Uglovskii AY, Melnikov IA, Alexeev IA, Kureev AA. Effective error floor estimation based on importance sampling with the uniform distribution. *Problems of Information Transmission*. 2023; 59(4): 217-224.

[5] Gallager RG. Low-density parity-check codes. *IRE Transactions on Information Theory*. 1962; 8(1): 21-28.

[6] Berrou C, Glavieux A, Thitimajshima P. Near Shannon limit error-correcting coding and decoding: Turbo-codes. *Proceedings of the IEEE International Conference on Communications*. Geneva, Switzerland: IEEE; 1993. p.1064-1070.

[7] Moher M, Gulliver TA. Cross-entropy and iterative decoding. *IEEE Transactions on Information Theory*. 1998; 44(7): 3097-3104.

[8] Richardson T. The geometry of turbo-decoding dynamics. *IEEE Transactions on Information Theory*. 2000; 46(1): 9-23.

[9] Mackay DJC. Good error-correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory*. 1999; 45(2): 399-431.

[10] Ikeda S, Tanaka T, Amari S. Information geometry of turbo and low-density parity-check codes. *IEEE Transactions on Information Theory*. 2004; 50(6): 1097-1114.

[11] McEliece RJ, Mackay DJC, Cheng JF. Turbo decoding as an instance of Pearl's belief propagation algorithm. *IEEE Journal on Selected Areas in Communications*. 1998; 16(2): 140-152.

[12] Amari S, Nagaoka H. *Methods of Information Geometry*. Oxford, UK: AMS and Oxford University Press; 2000.

[13] Nielson F. An elementary introduction to information geometry. *Entropy*. 2020; 22(10): 1100.

[14] Deutsch F. *Best Approximation in Inner Product Spaces*. New York: Springer-Verlag; 2001.

[15] Escalante R, Raydan M. *Alternating Projection Methods*. Philadelphia, USA: SIAM; 2011.

[16] Ginat O. *The Method of Alternating Projections*. Oxford, UK: University of Oxford; 2018.

[17] Halperin I. The product of projection operators. *Acta Scientiarum Mathematicarum (Szeged)*. 1962; 23: 96-99.

[18] Von Neumann J. On rings of operators reduction theory. *Annals of Mathematics*. 1949; 50(2): 401-485.

[19] Von Neumann J. *The Geometry of Orthogonal Spaces*. USA: Princeton University Press; 1950.

[20] Csiszar I. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*. 1967; 7(3): 200-217.

[21] Csiszar I, Matus F. Information projections revisited. *IEEE Transactions on Information Theory*. 2003; 49(6): 1474-1490.

[22] Csiszar I, Tusnady G. Information geometry and alternating minimization procedure. *Statistics and Decisions*. 1984; 1: 205-237.

[23] Alberge F. Iterative decoding as Dykstra's algorithm with alternate I-projection and reverse I-projection. *2008 16th European Signal Processing Conference*. Lausanne, Switzerland: IEEE; 2008.

[24] Walsh JM, Regalia PA. Belief propagation, dykstra's algorithm, and iterated information projections. *IEEE Transactions on Information Theory*. 2010; 56(8): 4114-4128.

[25] Walsh JM, Johnson CR, Regalia PA. A refined information geometric interpretation of turbo decoding. In: *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005*. Philadelphia, PA, USA: IEEE; 2005.

[26] Lewis AS, Malick J. Alternating projections on manifolds. *Mathematics of Operations Research*. 2008; 33(1): 216-234.