LETTER

A Phenomenological Study on Threshold Improvement via Spatial Coupling*

Keigo TAKEUCHI†(a), Toshiyuki TANAKA††(b), and Tsutomu KAWABATA††(c), Members

SUMMARY Kudekar et al. proved an interesting result in low-density parity-check (LDPC) convolutional codes: The belief-propagation (BP) threshold is boosted to the maximum-a-posteriori (MAP) threshold. Furthermore, the authors showed that the BP threshold for code-division multiple-access (CDMA) systems is improved up to a threshold below the optimal one via spatial coupling. In this letter, a phenomenological model for elucidating the essence of these phenomena, called threshold improvement, is proposed. The main result implies that threshold improvement occurs for spatially-coupled general graphical models.

key words: spatial coupling, threshold saturation, belief propagation (BP), dynamical systems, mean-field models.

1. Introduction

Low-density parity-check (LDPC) convolutional codes have been shown to outperform conventional LDPC block codes when iterative decoders based on belief propagation (BP) are used [1,2]. An LDPC convolutional code is constructed as a one-dimensionally coupled chain of LDPC block codes. As an explanation of this interesting result, it has been shown theoretically [3] and numerically [4] that the BP threshold of an LDPC convolutional code is boosted to the maximum-a-posteriori (MAP) threshold of the corresponding LDPC block code. This phenomenon is called “threshold saturation” via spatial coupling [3].

Recently, we showed that a phenomenon similar to the threshold saturation occurs in spatially-coupled code-division multiple-access (CDMA) systems: The BP threshold for randomly-spread CDMA systems is boosted to a threshold below the optimal one via spatial coupling [5]. In this letter, this phenomenon is called “threshold improvement” via spatial coupling. It is believed that threshold improvement is universal, i.e., the performance of the BP algorithm can be improved by coupling so-called mean-field graphical models spatially. This conjecture has been verified for several graphical models [6,7]. The purpose of this letter is to present a phenomenological study that supports the universality of threshold improvement.

Threshold improvement is a static property of spatially-coupled graphical models, rather than dynamical properties. It is well-known that BP can calculate the MAP solution exactly if the factor graph defining the BP is a tree [8]. Furthermore, if the BP algorithm for a general factor graph converges, the BP fixed-points correspond to the stationary points of the so-called Bethe free energy for the factor graph [9], while the MAP solution corresponds to the global minimizer of the true free energy. Roughly speaking, the Bethe free energy is obtained by approximating the original factor graph by a tree. These results allow us to characterize the static properties of the BP algorithm by stationary solutions to a dynamical system that has a potential energy function whose fixed-points coincide with those for the Bethe free energy, which is analogous to the density evolution (DE) equation for LDPC codes. In this letter, we restrict graphical models to such a class of graphical models that BP algorithms converge asymptotically.

As a phenomenological model for elucidating threshold improvement, we propose a spatially-coupled dynamical system with multiple stable solutions. This letter is organized as follows: In Section 2, the BP is characterized via a dynamical system with multiple stable solutions after presenting a motivating example for regular LDPC codes. In Section 3, a spatially-coupled dynamical system with multiple stable solutions is defined. Furthermore, an intuitive understanding of threshold improvement is presented. Section 4 presents the main result of this letter.

2. Systems without Spatial Coupling

2.1 Density Evolution for regular LDPC Codes

The explicit formula of the Bethe free energy for LDPC codes is unknown. We shall construct potential energy associated with the Bethe free energy from DE. Let us consider the DE equation for a regular LDPC codes over binary erasure channel (BEC) with erasure probability $\epsilon$ [10, Theorem 3.50]

$$y_{t+1} - y_t = -\frac{dU}{dy}(y_t; \epsilon), \quad y_0 = 1,$$

where the potential energy associated with the Bethe free energy is given by
In these expressions, $y_t$ denotes the bit error rate (BER) for the BP decoder in iteration $t$. Furthermore, $\lambda(y)$ and $\rho(y)$ are the variable and check degree distributions from an edge perspective, respectively, satisfying $\lambda(0) = 0$, $\rho(1) = 1$, and $\rho'(y) > 0$. Note that the BP fixed-point $y_\infty$ corresponds to a stationary point of the potential energy (2). When $\epsilon$ is smaller than the BP threshold $\epsilon_{\text{BP}}$, the DE equation (1) has the unique solution $y_t = 0$. This observation implies that the BER converges to zero in $t \to \infty$ for $\epsilon < \epsilon_{\text{BP}}$. When $\epsilon > \epsilon_{\text{BP}}$, on the other hand, the DE equation (1) has two stable solutions $y_- = 0$ and $y_+ > 0$, and one unstable solution $y_0 > 0$, satisfying $y_- < y_0 < y_+$. The BER in this case converges to the strictly positive value $y_+$ in $t \to \infty$. Threshold saturation [3] implies that $y_t$ can approach the stable solution $y_- = 0$ for $\epsilon \in [\epsilon_{\text{BP}}, \epsilon_{\text{MAP}}]$ by spatial coupling, with $\epsilon_{\text{MAP}}$ denoting the MAP threshold. Thus, spatial coupling can be regarded as a method for conveying a metastable solution to the (global) stable solution.}

\[ U(y, \epsilon) = \int_0^y \lambda(c(z)) [\tilde{c}(z) - \epsilon] \, dz, \]  

with \[ c(y) = 1 - \rho(1 - y), \quad \tilde{c}(y) = \frac{y}{\lambda(c(y))}. \] \hspace{1cm} (3)

In order to investigate the universality of threshold improvement, we consider a dynamical system with a general potential energy function $U(y)$

\[ \frac{dy}{dt} = -\frac{dU}{dy}(y), \quad y(0) = y_0 \in \mathbb{R}. \] \hspace{1cm} (4)

Figure 1 shows a typical landscape of the potential energy $U(y)$. $y_-, y_+$, and $y_0$ denote a metastable solution, an unstable solution, and the stable solution of $U(y)$, respectively. For convenience in analysis, we have selected $y_+$ as the global minimizer of $U(y)$. Without loss of generality, $y_- < y_+$ is assumed. If the initial value $y_0$ is larger than the unstable solution $y_+$, the state $y(t)$ converges to the stable solution $y_+$ in $t \to \infty$. Otherwise, $y(t) \to y_-$. The stable solution $y_+$, which is the minimizer of $U(y)$, corresponds to the optimal solution. The typical BP solution corresponds to the metastable solution $y_-$, because the initial values for the BP algorithm are commonly closer to a metastable solution than to the stable solution, e.g. see [5][11] for code-division multiple-access (CDMA) systems. Thus, there is a gap between the BP performance and the optimal performance when the potential energy $U(y)$ have multiple stable solutions. We hereafter refer to metastable solutions and the stable solution as BP solutions and the optimal solution, respectively.

\[ y = y_+, \quad y = y_-, \quad y = y_0 \]

\[ U(y) \]

\[ y = y_u, \quad y = y_l, \quad y = y_r \]

2.2 Dynamical System with Multiple Stable Solutions

In order to move the BP solution to the optimal solution $y_+$, we consider one-dimensional coupling with a positive coupling function $D(y) > 0$,

\[ \frac{dy}{dt} = -\frac{dU}{dy}(y) - \frac{1}{2} D'(y) \left( \frac{\partial y}{\partial x} \right)^2 + \frac{\partial}{\partial x} \left( D(y) \frac{\partial y}{\partial x} \right), \] \hspace{1cm} (5)

with the initial and boundary conditions

\[ y(x, 0) = y_0 \in \mathbb{R} \quad \text{for} \quad x \in (-x_{\max}, x_{\max}), \] \hspace{1cm} (6)

\[ y(\pm x_{\max}, t) = y_+, \] \hspace{1cm} (7)

for $x_{\max} > 0$. The variable $x$ denotes the spatial position of coupled systems. More precisely, the lattice points $[-x_{\max}, -x_{\max} + 1, \ldots, x_{\max}]$ should be regarded as positions. Thus, $2[x_{\max} + 1] \approx O(1)$ corresponds to the number of coupled systems. This interpretation is reasonable when the difference of the lattice points $x = \pm x_{\max}$ tends to infinity in $x_{\max} \to \infty$. In this limit, $x_{\max} \to \infty$, the difference $y(x + 1, t) - y(x, t)$ is expected to be sufficiently small. The limit $x_{\max} \to \infty$ is equivalent to the limit in which $D$ tends to zero while $x_{\max} = O(1)$. The system [5] can be regarded as a dynamical system with the so-called Ginzburg-Landau free energy $H(y)$ [12] as its potential energy

\[ \frac{\delta y}{\delta t} = -\frac{\delta H}{\delta y}(y), \] \hspace{1cm} (8)

with

\[ H(y) = \int_{-x_{\max}}^{x_{\max}} \left( U(y(x, t)) + \frac{D(y(x, t))}{2} \left( \frac{\partial y}{\partial x} \right)^2 \right) \, dx. \] \hspace{1cm} (9)

In [8], $\delta/\delta y$ denotes the functional derivative with respect to $y$. Expression (9) implies that we impose spatial coupling that smooths the state $y(x, t)$ spatially. The point of spatial coupling is that the boundary is fixed to the optimal

\[ \delta \] Threshold improvement for spatially-coupled CDMA systems can be understood from [5] with a constant coupling function [5]. For an ensemble of spatially-coupled LDPC codes [3] a higher-order approximation is needed at the boundaries, while the DE equation can be approximated by [5] in a region far from the boundary. This implies that spatially-coupled LDPC codes are a peculiar example.
solution \(y_{\pm}\). A “stretched rubber rope” whose both ends are fixed to the optimal solution \(y_{0}\) is utilized to “climbs” the potential barriers between the BP solution(s) and the optimal solution. The elastic force of the rubber rope lifts the state \(y(x,t)\) toward the optimal solution. In LDPC convolutional codes, such a boundary condition results from termination of convolutional codes [3].

4. Main Result

For simplicity, we hereafter assume that the coupling function does not depend on \(y\), i.e., \(D(y) = D > 0\). We believe that the main result holds for state-dependent coupling functions. We focus on stationary solutions \(y(x)\) to \((5)\), satisfying

\[
0 = - \frac{dU}{dy}(y) + D \frac{d^2y}{dx^2}, \quad y(\pm x_{\text{max}}) = y_{\pm}. \tag{10}
\]

Figure 2 shows examples of the stationary solution \(y(x)\) for a double-well potential \(U(y) = y^4/4 - y^2/2 - hy\) with a parameter \(h \in \mathbb{R}\). The double-well potential has a metastable solution \(y_{-} < 0\) \((y_{+} > 0)\) and a stable solution \(y_{+} > 0\) \((y_{-} < 0)\) for \(h > 0\) \((h < 0)\). When \(h = 0.01\), the state approaches the uniform solution \(y(x) = y_{+}\). In other words, the BP solution coincides with the optimal solution. When \(h = -0.01\), on the other hand, the stationary solution \(y(x)\) is a pot-shaped solution. Note that this solution is a natural solution for the case where the state \(y(x,t)\) cannot climb potential barriers.

The main result of this letter is that there are no pot-shaped stationary solutions if the boundary is fixed to the optimal solution.

**Definition 1:** A stationary solution \(y(x)\) is called a pot-shaped solution if the following conditions are satisfied:

- \(y(0) < y(\pm x_{\text{max}})\).
- \(dy/dx \geq 0\) for \(x > 0\).

Note that any solution to \((10)\) is an even function of \(x\), because the differential equation \((10)\) is invariant under the transformation \(x' = -x\). Thus, the second condition implies \(dy/dx \leq 0\) for \(x < 0\).

**Theorem 1:** Suppose that the coupling function \(D(y) > 0\) does not depend on \(y\). If the optimal solution \(y_{+}\) of \(U(y)\) is unique, then, there are no pot-shaped stationary solutions.

Before proving Theorem 1 we shall present the significance of Theorem 1. Obviously, the uniform solution \(y(x) = y_{0}\) is a stationary solution to the spatially-coupled dynamical system \((5)\). Intuitively, non-monotonic solutions cannot become stationary solutions to \((5)\) with the uniform initial condition \(y(x,0) = y_{0}\), because the state \(y(x,t)\) should move closer to \(y_{0}\) as the position \(x\) gets closer to the boundary. Thus, Theorem 1 implies that the state \(y(x,t)\) converges to the uniform solution \(y(x) = y_{0}\) in \(t \to \infty\) for any initial value, i.e., the BP solution coincides with the optimal solution for general potential energy \(U(y)\) having the unique optimal solution.

**Remark 1:** Suppose that the potential energy \(U(y)\) contains a parameter \(\epsilon\), which corresponds to the erasure probability for LDPC codes over BEC or to the system load for CDMA systems. We rewrite the potential energy as \(U(y; \epsilon)\). The shape of \(U(y, \epsilon)\) continuously changes with the parameter \(\epsilon\). Theorem 1 allows us to find the BP threshold for a spatially-coupled graphical model as such a point \(\epsilon_{\text{BP}}^{(SC)}\) that there are two distinct global stable solutions \(y_{-}\) and \(y_{+}\), satisfying \(U(y_{-}; \epsilon_{\text{BP}}^{(SC)}) = U(y_{+}; \epsilon_{\text{BP}}^{(SC)})\). The BP threshold \(\epsilon_{\text{BP}}^{(SC)}\) does not coincide with the optimal one for spatially-coupled CDMA systems [5], while \(\epsilon_{\text{MAP}}^{(SC)}\) is equal to the MAP threshold \(\epsilon_{\text{MAP}}\) for spatially-coupled LDPC codes over BEC [3]. These observations imply that it depends on graphical models whether \(\epsilon_{\text{BP}}^{(SC)}\) coincides with the optimal one.

**Proof of Theorem 1** We shall prove Theorem 1 by contradiction. Suppose that there is a pot-shaped stationary solution \(y(x)\). Integrating \((11)\) after multiplying both sides by \(dy/dx\), we obtain

\[
\frac{D}{2} \left( \frac{dy}{dx} \right)^2 = U(y) + C, \tag{11}
\]

with a constant \(C\). We use the boundary condition \(y(\pm x_{\text{max}}) = y_{\pm}\) and the positivity of the left-hand side on \((11)\) to find \(C \geq -U(y_{+})\), where we have used the assumption that \(y_{+}\) is the unique minimizer of \(U(y)\). If \(y_{+}\) was a local minimizer, the inequality would have to be replaced by \(C \geq -U(y_{+}) + \Delta U\), with \(\Delta U > 0\) denoting the energy gap between \(y_{+}\) and the global minimizer. Since we have assumed that there is a stationary solution satisfying \(dy/dx \geq 0\) \((dy/dx \leq 0\) for \(x > 0\) \((x < 0)\), integrating \((11)\) after taking the square root of both sides yields

\[
F(y) = x - \bar{x} \quad \text{for} \quad x > 0, \tag{12}
\]

with

\[
F(y) = \frac{\sqrt{D}}{2} \int_{y_{0}}^{y} \frac{dy'}{\sqrt{U(y') + C}}, \tag{13}
\]
where $\bar{y}$ denotes a value between $y(0)$ and $y_*$ that does not minimize the potential energy, i.e., $U(\bar{y}) > U(y_*)$. In (12), we have selected a constant of integration such that $y(\bar{x}) = \bar{y}$. Repeating the same argument for $x < 0$, we obtain

$$y = F^{-1}(|x| - \bar{x}) \quad \text{for } x \in (-x_{\text{max}}, x_{\text{max}}),$$

(14)

where $F^{-1}$ denotes the inverse function of (13).

Any stationary solution must be differentiable since it is a solution to the second-order differential equation (10). However, the solution (14) is indifferenciable at the origin unless $dF^{-1}/dy|_{y=\bar{y}}(0) = 0$ or $dF/dy|_{y=\bar{y}}(0) = \infty$, in which the value $y(0)$ at the origin is given by $y(0) = F^{-1}(\bar{x})$. The uniqueness of the optimal solution implies that the integrand in (13) can diverge only at $y = y_*$. Thus, a necessary condition for $dF/dy|_{y=\bar{y}}(0) = \infty$ is $y(0) = y_*$, which contradicts $y(0) < y(\pm x_{\text{max}})$.

What occurs when the boundary is fixed to a metastable solution? Figure 3 shows a bifurcation diagram of stationary solutions for a double-well potential $U(y) = y^4/4 - y^2/2 - hy$ with a parameter $h$. In this case, the boundary is fixed to the metastable solution $y_*$ for $h < 0$. The state converges to a pot-shaped stationary solution in $t \to \infty$ when $(D, -h)$ locates above the line shown in Fig. 3. Otherwise, the state converges to the uniform solution $y = y_*$. As long as $D$ is finite, the state is conveyed to the metastable solution $y = y_*$ for small $h$. However, such a solution seems to disappear in $D \to 0$, which corresponds to the limit in which the number of coupled systems tends to infinity [3].

References