# A Homological Approach to Belief Propagation and Bethe Approximations

Olivier Peltre(✉)

Université Paris 7 Denis Diderot, équipe Géométrie et Dynamique, Paris, France
olivier@peltre.xyz

**Abstract.** We introduce a differential complex of local observables given a set of random variables covered by subsets. Its boundary operator $\partial$ allows us to define a transport equation $\dot{u} = \partial \Phi(u)$ equivalent to Belief Propagation. This definition reveals a maximal set of conserved quantities under Belief Propagation and gives new geometric insight on the relationship of its equilibria with the critical points of Bethe free energy.

## 1 Introduction

A common feature of statistical physics and statistical learning is to consider a very large number of random variables, each of them mostly interacting with only a small subset of neighbours. Both lead to the effort of extracting relevant information about collective phenomena in spite of intractable global computations, hence motivating the development of local techniques where only small enough subsets of variables are simultaneously considered.

In the present note, we work with a collection of local algebras of observables, on which a boundary operator describes relations between intersecting subsystems. The construction of this differential complex is exposed in Sect. 2. It allows for a homological interpretation of the equivalence established by Yedidia *et al.* between critical points of the Bethe free energy approximation and fixed points of the Belief Propagation algorithm. We review this beautiful theorem bridging statistical learning and thermodynamics[1] in Sect. 3.

This is part of a PhD work done under the kind supervision of Daniel Bennequin. I wish to thank him as well as Grégoire Sergeant-Perthuis and Juan-Pablo Vigneaux, for their sustained collaboration and many fruitful discussions.

[1] See the longer version of this note for more context and historical background.


F. Nielsen and F. Barbaresco (Eds.): GSI 2019, LNCS 11712, pp. 218–227, 2019.
https://doi.org/10.1007/978-3-030-26980-7_23

# 2 Differential and Combinatorial Structures

## 2.1 Statistical System

**Regions.** We call system a finite set $\Omega$ equipped with a covering $X \subseteq \mathcal{P}(\Omega)$ by subsets such that:

- the empty set $\varnothing$ is in $X$,
- if $\alpha \in X$ and $\beta \in X$, then $\alpha \cap \beta$ is also in $X$.

We view $X$ as a subcategory of the partial order $\mathcal{P}(\Omega)$ having an arrow $\alpha \to \beta$ whenever $\beta \subseteq \alpha$. We call $\alpha \in X$ a region[2] of $\Omega$ and denote by $\Lambda^\alpha \subseteq X$ the subsystem of those regions contained in $\alpha$.

**Chains and Nerve.** A $p$-chain $\bar{\alpha}$ is a totally ordered sequence $\alpha_0 \to \ldots \to \alpha_p$ in $X$, it is said non-degenerate when all inclusions are strict. A $p$-chain $\bar{\alpha}$ may be viewed as a $p$-simplex, whose $p+1$ faces are the chains $\bar{\alpha}^{(k)}$ obtained by removing $\alpha_k$, for $0 \leq k \leq p$. The nerve of $X$ is the simplicial complex $NX = \bigsqcup_p N^p X$ formed by all non-degenerate chains.

**Microscopic States.** For each $i \in \Omega$, suppose given a finite set $E_i$. A microscopic state of a region $\alpha \subseteq \Omega$ is an element of the cartesian product[3]:

$$E_\alpha = \prod_{i \in \alpha} E_i$$

We denote by $\pi^{\beta\alpha} : E_\alpha \to E_\beta$ the canonical projection of $E_\alpha$ onto $E_\beta$ whenever $\beta$ is a subregion of $\alpha$.

## 2.2 Scalar Fields

**Differentials.** We call scalar field a collection $\lambda \in \mathbb{R}(X)$ of scalars indexed by the nerve of $X$. We denote by $\mathbb{R}_p(X) = \mathbb{R}^{N^p X}$ the space of $p$-fields, vanishing everywhere but on the $p$-simplices of $NX$.

Through the canonical scalar product of $\mathbb{R}(X)$, scalar fields can be identified with chains or cochains with real coefficients in $NX$ eitherwise. We denote by $\partial$ the boundary operator of $\mathbb{R}(X)$ and by $d$ its adjoint differential:

$$\begin{aligned} \partial &: \mathbb{R}_0(X) \leftarrow \mathbb{R}_1(X) \leftarrow \ldots \\ d &: \mathbb{R}_0(X) \to \mathbb{R}_1(X) \to \ldots \end{aligned}$$

[2] The term refers to the notion of region-graphs in Yedidia et al.

[3] The configuration space $E_\varnothing$ is thus a point, unit for the cartesian product and terminal element in **Set**.

**Convolution.** Let $\tilde{\mathbb{R}}_1(X) = \mathbb{R}_0(X) \oplus \mathbb{R}_1(X)$. Identifying the degenerate 1-chain $\alpha \to \alpha$ with $\alpha$, an element of $\tilde{\mathbb{R}}_1(X)$ is indexed by general 1-chains in $X$. Equipped with Dirichlet convolution, $\tilde{\mathbb{R}}_1(X)$ is the incidence algebra[4] of $X$:

$$(\varphi * \psi)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} \varphi_{\alpha\beta'} \cdot \psi_{\beta'\gamma}$$

The unit of $*$ is $1 \in \mathbb{R}_0(X)$, sometimes viewed as a Kronecker symbol in $\tilde{\mathbb{R}}_1(X)$.

The space of 0-fields $\mathbb{R}_0(X)$ also has a $\tilde{\mathbb{R}}_1(X)$-bimodule structure, where the left action of $\varphi \in \tilde{\mathbb{R}}_1(X)$ on $\lambda \in \mathbb{R}_0(X)$ is given by:

$$(\varphi \cdot \lambda)_\alpha = \sum_{\alpha \to \beta'} \varphi_{\alpha\beta'} \lambda_{\beta'}$$

**Möbius Inversion.** The Dirichlet zeta function $\zeta \in \tilde{\mathbb{R}}_1(X)$ is defined by $\zeta_{\alpha\beta} = 1$ for every $\alpha \to \beta$ in $X$. When $X$ is locally finite[5], $\zeta$ is invertible. Its inverse $\mu$, known as the Möbius function, satisfies:

$$\mu_{\alpha\beta} = \sum_{k \geq 0} (-1)^k (\zeta - 1)^{*k}_{\alpha\beta}$$

where $(\zeta - 1)^{*k}_{\alpha\beta}$ counts the number of non-degenerate $k$-chains from $\alpha$ to $\beta$.

The coefficients $c = (1 \cdot \mu) \in \mathbb{R}_0(X)$ contain all the combinatorics of Bethe approximations. They satisfy the following «inclusion-exclusion» formula:

$$(c \cdot \zeta)_\beta = \sum_{\alpha' \to \beta} c_{\alpha'} = 1$$

### 2.3 Observables, Densities and Statistical States

**Observable Fields.** Denote by $\mathfrak{a}_\alpha = \mathbb{R}^{E_\alpha}$ the commutative algebra of observables on $\alpha \subseteq \Omega$. For every subregion $\beta \subseteq \alpha$, an observable $u_\beta \in \mathfrak{a}_\beta$, as a real function on $E_\beta$, admits a cylindrical extension $j_{\alpha\beta}(u_\beta)$ on $E_\alpha$.

For every $\bar{\alpha} \in N_p X$, let $\mathfrak{a}_{\bar{\alpha}}$ denote a copy of the algebra $\mathfrak{a}_{\alpha_p}$ of observables on its smallest region. There is an injection $j_{\bar{\beta}\bar{\alpha}} : \mathfrak{a}_{\bar{\alpha}} \to \mathfrak{a}_{\bar{\beta}}$ whenever $\bar{\beta}$ is a subchain of $\bar{\alpha}$.[6] We define the graded vector space $\mathfrak{a}(X)$ of observable fields by:

$$\mathfrak{a}_p(X) = \bigoplus_{\bar{\alpha} \in N^p X} \mathfrak{a}_{\bar{\alpha}}$$

[4] See Rota [12] for a deeper treatment of these combinatorial structures.

[5] $X$ is locally finite if for any $\alpha, \beta \in X$ there is only a finite number of non-degenerate chains from $\alpha$ to $\beta$.

[6] Observable fields form a simplicial algebra $\mathfrak{a}(X) : \mathbf{Ord}^{op} \to \mathbf{Alg}$. To relate to this more general theory, see Segal's note on classifying spaces [14] for instance.

It is equipped with a boundary[7] operator $\partial : \mathfrak{a}_{p+1}(X) \to \mathfrak{a}_p(X)$. When $p = 0$, we have for instance[8]:

$$\partial_\beta \varphi = \sum_{\alpha' \to \beta} \varphi_{\alpha'\beta} - \sum_{\beta \to \gamma'} \varphi_{\beta\gamma'}$$

Belief Propagation is essentially a dynamic up to a boundary term $\partial\varphi$ in $\mathfrak{a}_0(X)$, although it is usually viewed in the multiplicative group $G_0(X) = \prod_{\alpha \in X} G_\alpha$ with $G_\alpha = (\mathbb{R}_+^*)^{E_\alpha}$.

**Density Fields.** We call density on $\alpha \subseteq \Omega$ a linear form on observables $\omega_\alpha \in \mathfrak{a}_\alpha^*$. Denote by $\Sigma^{\beta\alpha}(\omega_\alpha) \in \mathfrak{a}_\beta^*$ the partial integration of $\omega_\alpha$ along the fibers of $\pi^{\beta\alpha}$:

$$\Sigma^{\beta\alpha}(\omega_\alpha)(x_\beta) = \sum_{x' \in E_{\alpha \setminus \beta}} \omega_\alpha(x_\beta, x')$$

It satisfies $\langle\, \Sigma^{\beta\alpha}(\omega_\alpha) \,|\, u_\beta \,\rangle = \langle\, \omega_\alpha \,|\, j_{\alpha\beta}(u_\beta) \,\rangle$ for every $u_\beta \in \mathfrak{a}_\beta$.

The complex $\mathfrak{a}^*(X)$ is equipped with a differential $d : \mathfrak{a}_p^*(X) \to \mathfrak{a}_{p+1}^*(X)$, adjoint of $\partial$. For $p = 0$, we have for instance:

$$(d\omega)_{\alpha\beta} = \omega_\beta - \Sigma^{\beta\alpha}(\omega_\alpha)$$

A field $\omega \in \mathfrak{a}_0^*(X)$ is said consistent if $d\omega = 0$. The notion of consistent densities will replace that of a global measure on $E_\Omega$.

**Statistical Fields.** Denote by $\Delta_\alpha \subseteq \mathfrak{a}_\alpha^*$ the convex subset of probability measures. It consists of all the positive densities $\omega_\alpha$ satisfying $\omega_\alpha(1_\alpha) = 1$, and any non-trivial positive density $\omega_\alpha \in \mathfrak{a}_\alpha^*$ defines a normalised density $[\omega_\alpha] \in \Delta_\alpha$.

Its interior $\mathring{\Delta}_\alpha$ admits a natural Lie group structure, as it is diffeomorphic to the quotient of $G_\alpha$ by scalings of $\mathbb{R}_+^*$, itself isomorphic to the quotient of $\mathfrak{a}_\alpha$ by the action of additive constants. We denote by $[\mathrm{e}^{-U_\alpha}] \in \mathring{\Delta}_\alpha$ the Gibbs state associated to $U_\alpha \in \mathfrak{a}_\alpha$ and by $[\mathrm{e}^{-}]_\alpha : \mathfrak{a}_\alpha \to \mathring{\Delta}_\alpha$ this surjective group morphism.

We denote by $\Delta(X) \subseteq \mathfrak{a}^*(X)$ the convex subset of statistical fields, by $\mathring{\Delta}(X)$ its interior, and by $\mathring{\Delta}^d \subseteq \mathring{\Delta}_0(X)$ the subset of consistent ones.

## 2.4 Homology

**Gauss Formulas.** For every region $\alpha \in X$, let us define the coboundary of the subsystem $\Lambda^\alpha$ as the subset of arrows $\delta\Lambda^\alpha \subseteq N^1 X$ that meet $\Lambda^\alpha$ but are not contained in $\Lambda^\alpha$:

$$\delta\Lambda^\alpha = \{\alpha' \to \beta' \mid \alpha' \notin \Lambda^\alpha \quad \text{and} \quad \beta' \in \Lambda^\alpha\}$$

[7] A boundary $\partial$ satisfies $\partial^2 = \partial \circ \partial = 0$.

[8] We will generally drop the injection in our notation.

The following proposition may then be thought of as a Gauss formula on $\Lambda^\alpha$:

**Proposition 1.** *For every $\varphi \in \mathfrak{a}_1(X)$ and $\alpha \in X$ we have:*

$$\sum_{\beta' \in \Lambda^\alpha} \partial_{\beta'} \varphi = \sum_{\alpha'\beta' \in \delta\Lambda^\alpha} \varphi_{\alpha'\beta'}$$

*In particular, the above vanishes if $\varphi$ is supported in $\Lambda^\alpha$.*

A similar formula holds on the cone $V_\beta$ over $\beta$ in $X$, formed by all the regions containing $\beta$ with coboundary the set $\delta V_\beta$ of arrows leaving $V_\beta$. The sums however need to be embedded in the space of global observables.

**Proposition 2.** *For every $\varphi \in \mathfrak{a}_1(X)$ and $\beta \in X$ we have:*

$$\sum_{\alpha' \in V_\beta} \partial_{\alpha'} \varphi = - \sum_{\alpha'\beta' \in \delta V_\beta} \varphi_{\alpha'\beta'}$$

*as global observables of $\mathfrak{a}_\Omega$.*

**Interaction Decomposition.** We call boundary observable on a region $\alpha \in X$ any observable generated by observables on strict subregions of $\alpha$ in $X$. Suppose chosen for every $\alpha$ a supplement $\mathfrak{z}_\alpha$ of boundary observables, so that:

$$\mathfrak{a}_\alpha = \mathfrak{z}_\alpha \oplus \Big( \sum_{\alpha > \beta'} \mathfrak{a}_{\beta'} \Big)$$

We may inductively continue this procedure, as illustrated by the following well known[9] theorem.

**Theorem 1** (Interaction Decomposition)**.** *Given supplements $(\mathfrak{z}_\alpha)$ of boundary observables for every $\alpha \in X$, we have the decompositions:*

$$\mathfrak{a}_\alpha = \bigoplus_{\alpha \to \beta'} \mathfrak{z}_{\beta'}$$

*They induce a projection $P$ of $\mathfrak{a}_0(X)$ onto $\mathfrak{z}_0(X) = \bigoplus_\alpha \mathfrak{z}_\alpha$ defined by:*

$$P^\beta(u) = \sum_{\alpha' \to \beta} P^{\beta\alpha'}(u_{\alpha'})$$

*where $P^{\beta\alpha}$ denotes the projection of $\mathfrak{a}_\alpha$ onto $\mathfrak{z}_\beta$ for all $\alpha \to \beta$ in $X$.*

Given a field $u \in \mathfrak{a}_0(X)$, define the global observable $\zeta_\Omega(u) \in \mathfrak{a}_\Omega$ by:

$$\zeta_\Omega(u) = \sum_{\alpha \in X} u_\alpha$$

**Corollary 1.** *For any $u \in \mathfrak{a}_0(X)$, we have the equivalence:*

$$P(u) = 0 \quad \Leftrightarrow \quad \zeta_\Omega(u) = 0$$

*In particular, $\mathfrak{z}_0(X)$ is isomorphic to the image of $\zeta_\Omega$ in $\mathfrak{a}_\Omega$*[10].

[9] The first appearance of this now very common result in statistics seems to be in Kellerer [4]. See also [7] for a proof via harmonic analysis.

[10] They both represent the inductive limit of $\mathfrak{a}$ over $X$.

**Homology Groups.** The complex of observable fields $\mathfrak{a}(X)$ is acyclic[11] and we only focus on the first homology group.

**Theorem 2.** *The interaction decomposition $P$ induces an isomorphism on the first homology group of observable fields:*

$$\mathfrak{a}_0(X)/\partial\mathfrak{a}_1(X) \sim \mathfrak{z}_0(X)$$

*Proof.* The Gauss formula on the cone $V_\beta$ above $\beta$ in $X$ first ensures that $P$ vanishes on boundaries:

$$P^\beta(\partial\varphi) = \sum_{\alpha'\to\beta} P^\beta(\partial_{\alpha'}\varphi) = \sum_{\alpha'\to\beta}\sum_{\beta'\not\to\beta} P^\beta(\varphi_{\alpha'\beta'}) = 0$$

as $P^\beta(\mathfrak{a}_{\beta'})$ is non-zero if and only if $\beta'$ contains $\beta$. Let us denote by $[P]$ the quotient map induced by $P$. Given $u \in \mathfrak{a}_0(X)$, consider the flux $\varphi$ defined by $\varphi_{\alpha\beta} = P^{\beta\alpha}(u_\alpha)$:

$$\partial_\beta\varphi = \sum_{\alpha'\to\beta} \varphi_{\alpha'\beta} - \sum_{\beta\to\gamma'} \varphi_{\beta\gamma'} = P^\beta(u) - u_\beta$$

When $P(u) = 0$ this gives $u = -\partial\varphi$, hence $[P]$ is injective.

**Corollary 2.** *Let $V = \zeta \cdot v$ in $\mathfrak{a}_0(X)$. We have the equivalence:*

$$cV \in \mathrm{Im}(\partial) \quad \Leftrightarrow \quad v \in \mathrm{Im}(\partial)$$

*Proof.* According to the theorem, it suffices to show that $P(v) = P(cV)$ and:

$$P^\gamma(v) = P^\gamma(\mu\cdot V) = \sum_{\alpha'\to\beta'\to\gamma} P^{\gamma\beta'}(\mu_{\alpha'\beta'}V_{\beta'}) = \sum_{\beta'\to\gamma} P^{\gamma\beta'}(c_{\beta'}V_{\beta'}) = P^\gamma(cV)$$

# 3 First Applications

## 3.1 Critical Points of Bethe Free Energy

**Gibbs Free Energy.** For every $\alpha \subseteq \Omega$, denote by $\mathcal{F}_\alpha$ its local Gibbs free energy, viewed as the functional on $\Delta_\alpha \times \mathfrak{a}_\alpha$ defined by:

$$\mathcal{F}_\alpha(p_\alpha, H_\alpha) = \mathbb{E}_{p_\alpha}[H_\alpha] - S(p_\alpha)$$

where $S(p_\alpha) = -\sum p_\alpha \ln(p_\alpha)$ denotes Shannon entropy.

Given a global hamiltonian $H_\Omega \in \mathfrak{a}_\Omega$, the global Gibbs state $[\mathrm{e}^{-H_\Omega}] \in \Delta_\Omega$ is the global minimum of $\mathcal{F}_\Omega(\,\cdot\,, H_\Omega)$. This definition being hardly computable in practice, we shall seek to estimate its marginals $\Sigma^{\alpha\Omega}(p_\Omega)$ by an approximation on the global Gibbs free energy $\mathcal{F}_\Omega$.

[11] We do not provide a proof here, a treatment of higher degrees shall be given in later work.

**Bethe Approximation.** The Bethe-Peierls approach and its refinements[12] essentially consist in writing an approximate decomposition of $\mathcal{F}_\Omega$ as a sum of local free energy summands $f_\beta$, for $\beta \in X$. This localisation procedure can be made exact on any $\alpha \in X$ by Möbius inversion:

$$\mathcal{F}_\alpha = \sum_{\alpha \to \beta'} f_{\beta'} \quad \Leftrightarrow \quad f_\beta = \sum_{\beta \to \gamma'} \mu_{\beta\gamma'} \, \mathcal{F}_{\gamma'}$$

The approximation only comes when $\Omega$ is not in $X$ and we may then write the error $\mathcal{F}_\Omega - \check{\mathcal{F}}$ as a global free energy summand $f_\Omega$. One should expect $f_\Omega$ to be small when sufficiently large regions are taken in $X$, by extensivity of the global Gibbs free energy[13].

The Bethe free energy $\check{\mathcal{F}}$ is thus defined for $p \in \Delta_0(X)$ and $H \in \mathfrak{a}_0(X)$ by:

$$\check{\mathcal{F}}(p, H) = \sum_{\beta \in X} c_\beta \cdot \mathcal{F}_\beta(p_\beta, H_\beta)$$

Given $H \in \mathfrak{a}_0(X)$, we denote by $\check{\mathcal{F}}^H$ the induced functional on $\Delta_0(X)$.

**Critical Points.** Because of the Möbius numbers $c_\beta$ appearing in its definition, the Bethe free energy $\check{\mathcal{F}}$ is no longer convex in general, and $\check{\mathcal{F}}^H$ might have a great multiplicity[14] of consistent critical points in $\mathring{\Delta}^d$.

**Theorem 3.** *A non-vanishing consistent statistical field $p \in \mathring{\Delta}^d$ is a critical point of the Bethe free energy $\check{\mathcal{F}}^H$ constrained to $\mathring{\Delta}^d$ if and only if there exists a flux $\varphi \in \mathfrak{a}_1(X)$ such that:*

$$-\ln(p) \simeq H + \zeta \cdot \partial\varphi \mod \mathbb{R}_0(X)$$

*Proof.* To describe the normalisation constraints, we may look at the quotient $\mathfrak{a}_0(X)/\mathbb{R}_0(X)$ as the cotangent space of $\mathring{\Delta}_0(X)$ at $p$, and write the differential of $\check{\mathcal{F}}^H$ as:

$$\frac{\partial \check{\mathcal{F}}}{\partial p} \simeq \sum_{\beta \in X} c_\beta \big( H_\beta + \ln(p_\beta) \big) \mod \mathbb{R}_0(X)$$

The flux term comes as a collection of Lagrange multipliers for the consistency constraints. Whenever $p$ is a critical point, the differential of $\check{\mathcal{F}}^H$ vanishes on $\mathrm{Ker}(d) = \mathrm{Im}(\partial)^\perp$ and we have:

$$c\big(H + \ln(p)\big) \in \mathrm{Im}(\partial) + \mathbb{R}_0(X)$$

The corollary of Theorem 2 is crucial[15] to state that this implies:

$$H + \ln(p) \in \zeta \cdot \mathrm{Im}(\partial) + \mathbb{R}_0(X)$$

---

[12] For reference see [2,5,8,11].

[13] Schlijper [13] proved this procedure convergent to the true free energy per lattice point for the infinite Ising 2D-model.

[14] For numerical studies see [6,9,15], A first mathematical proof of multiplicity is given by Bennequin in [1].

[15] The proof given in [17] is problematic when there exists $\beta$ such that $c_\beta = 0$.

### 3.2 Belief Propagation as a Transport Equation

**Effective Energy.** For every $\alpha \to \beta$ in $X$, call effective energy the smooth submersion $\mathbb{F}^{\beta\alpha}$ of $\mathfrak{a}_\alpha$ onto $\mathfrak{a}_\beta$ defined by:

$$\mathbb{F}^{\beta\alpha}(U_\alpha) = -\ln\left(\Sigma^{\beta\alpha}(\mathrm{e}^{-U_\alpha})\right)$$

Physically $\mathbb{F}^{\beta\alpha}(U_\alpha)(x_\beta)$ can be thought of as the conditional free energy of $\Lambda^\alpha$ given $x_\beta$. It is functorial in the category of smooth manifolds and we have the commutative diagram: $\Sigma^{\beta\alpha} \circ [\mathrm{e}^{-}]_\alpha = [\mathrm{e}^{-}]_\beta \circ \mathbb{F}^{\beta\alpha}$.

Let us call effective gradient the smooth functional $\nabla^{\mathbb{F}}$ from $\mathfrak{a}_0(X)$ to $\mathfrak{a}_1(X)$ defined by:

$$\nabla^{\mathbb{F}}(H)_{\alpha\beta} = H_\beta - \mathbb{F}^{\beta\alpha}(H_\alpha)$$

The hamiltonian $H$ is related to a field of local potentials $h$ by $H = \zeta \cdot h$. Letting $\Phi = -\nabla^{\mathbb{F}} \circ \zeta$, we have:

$$\Phi_{\alpha\beta}(h) = \mathbb{F}^{\beta\alpha}\Bigg(\sum_{\beta' \in \Lambda^\alpha \setminus \Lambda^\beta} h_{\beta'}\Bigg)$$

which is the effective contribution of $\Lambda^\alpha \setminus \Lambda^\beta$ to the energy of $\Lambda^\beta$.

**Belief Propagation.** Consider the following transport equation:

$$\dot{u} = \partial\Phi(u)$$

and denote by $\Xi = \partial\Phi$ the induced vector field on $\mathfrak{a}_0(X)$. In absence of normalisation, Belief Propagation[16] is equivalent to the naive Euler scheme[17] approximating the flow of $\Xi$ by:

$$\mathrm{e}^{n\tau\Xi} \simeq (1 + \tau\Xi)^n$$

The beliefs are given by $q = \mathrm{e}^{-U}$ with $U = \zeta \cdot u$.

This new perspective reveals the strong homological character of Belief Propagation. Denote by $T_h(\varphi)$ the transport of $h$ by a flux $\varphi \in \mathfrak{a}_1(X)$:

$$T_h(\varphi) = h + \partial\varphi$$

With initial condition $h$, the potentials $u$ remain in the image of $T_h$. This yields a maximal set of conserved quantities in light of Theorem 2.

**Theorem 4.** *Let $q \in G_0(X)^{\mathbb{N}}$ denote a sequence of belief fields obtained by iterating BP. The following quantity remains constant:*

$$q_\Omega = \prod_{\alpha \in X} (q_\alpha)^{c_\alpha}$$

[16] For reference and the algorithm formula see [3,9,10,15–17].

[17] BP is actually for $\tau = 1$, a different time scale would appear as exponent in the multiplicative formulation.

*Proof.* The fact that $u \in \mathrm{Im}(T_h)$ is equivalent to $P(u) = P(h)$. According to Corollary 1, this is also equivalent to $\zeta_\Omega(u) = \zeta_\Omega(h)$ where:

$$\zeta_\Omega(u) = \sum_{\alpha \in X} u_\alpha = \sum_{\beta \in X} c_\beta U_\beta$$

Letting $u = h + \partial\varphi$, BP can also be viewed as a dynamic over messages:

$$\dot{\varphi} = \Phi\big(T_h(\varphi)\big)$$

Although it converges on trees, this algorithm is generally divergent in presence of loops, and beliefs need to be normalised in order to attain projective equilibria.

**Normalisation.** Because the effective gradient $\nabla^{\mathbb{F}}$ is additive along constants and both $\zeta$ and $\partial$ preserve scalar fields, $\Xi$ induces a vector field on the quotient $\mathfrak{a}_0(X)/\mathbb{R}_0(X)$. Normalised belief are given by $q = [\mathrm{e}^{-U}]$ with $U = \zeta \cdot u$.

Given an initial hamiltonian $H = \zeta \cdot h$, a belief field $q \in \mathring{\Delta}_0(X)$ obtained by iterating BP satisfies:

$$-\ln(q) \simeq H + \zeta \cdot \partial\varphi \mod \mathbb{R}_0(X)$$

In virtue of Theorem 3, this implies that $q$ is a critical point of the Bethe free energy $\check{\mathcal{F}}^H$ constrained to $\mathring{\Delta}^d$ if and only if $q$ is a consistent statistical field. Considering all beliefs that may be obtained by such a choice of messages, let:

$$\mathring{\Delta}_H = \big\{[\mathrm{e}^{-U}] \mid U \in H + \zeta \cdot \mathrm{Im}(\partial)\big\} \subseteq \mathring{\Delta}_0(X)$$

Following Yedidia *et al.*, call any consistent $q \in \mathring{\Delta}_H \cap \mathring{\Delta}^d$ a fixed point of Belief Propagation[18]. With this terminology, we can rephrase their initial claim [17]:

**Theorem 5.** *Fixing a reference hamiltonian field $H$, fixed points of Belief Propagation are in one to one correspondence with critical points of the Bethe free energy.*

## References


1. Bennequin, D., Peltre, O., Sergeant-Perthuis, G., Vigneaux, J.P.: Informations, Energies and Messages. Preprint (2019)
2. Bethe, H.A., Bragg, W.L.: Statistical theory of superlattices. Proc. R. Soc. Lond. Ser. Math. Phys. Sci. **150**(871), 552–575 (1935)
3. Gallager, R.G.: Low-Density Parity-Check Codes. MIT Press, New York (1963)
4. Kellerer, H.G.: Maßtheoretische Marginalprobleme. Mathematische Annalen **153**(3), 168–198 (1964)
5. Kikuchi, R.: A theory of cooperative phenomena. Phys. Rev. **81**, 988–1003 (1951)


[18] This terminology is somewhat ambiguous as it does not mean that $q$ may be obtained by iterating BP from $[\mathrm{e}^{-H}]$.

6. Knoll, C., Pernkopf, F.: On loopy belief propagation - local stability analysis for non-vanishing fields. In: Uncertainty in Artificial Intelligence (2017)
7. Matúš, F.: Discrete marginal problem for complex measures. Kybernetika **24**, 36–46 (1988)
8. Morita, T.: Cluster variation method of cooperative phenomena and its generalization I. J. Phys. Soc. Jpn. **12**(7), 753–755 (1957)
9. Murphy, K.P., Weiss, Y., Jordan, M.I.: Loopy belief propagation for approximate inference: an empirical study. In: UAI (1999)
10. Pearl, J.: Networks of plausible inference. In: Probabilistic Reasoning in Intelligent Systems (1988)
11. Peierls, R.: On Ising's model of ferromagnetism. Math. Proc. Camb. Philos. Soc. **32**(3), 477–481 (1936)
12. Rota, G.-C.: On the foundations of combinatorial theory - I. Theory of Möbius functions. Z. Warscheinlichkeitstheorie **2**, 340–368 (1964)
13. Schlijper, A.G.: Convergence of the cluster-variation method in the thermodynamic limit. Phys. Rev. B **27**, 6841–6848 (1983)
14. Segal, G.: Classifying spaces and spectral sequences. Publications Mathématiques de l'IHÉS **34**, 105–112 (1968)
15. Weiss, Y.: Belief propagation and revision in networks with loops. Technical report, MIT (1997)
16. Yedidia, J.S., Freeman, W.T., Weiss, Y.: Bethe free energy, Kikuchi approximations, and belief propagation algorithms. Technical Report TR2001-16, MERL - Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, May 2001
17. Yedidia, J.S., Freeman, W.T., Weiss, Y.: Constructing free energy approximations and generalized belief propagation algorithms. IEEE Trans. Inf. Theory **51**(7), 2282–2312 (2005)