Planar and Surface Graphical Models which are EASY

Vladimir Chernyak[(1,2)] and **Michael Chertkov**[1]

[1] Center for Nonlinear Studies & Theory Division, LANL
[2] Chemistry Department, Wayne State, Detroit

"Physics of Algorithms" Workshop
Santa Fe, September 1, 2009

Outline

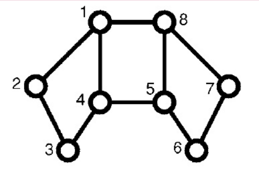
1. Introduction
   - Graphical Models
   - Easy and Difficult
   - Dimer and Ising Models on Planar Graphs
2. Planar is not necessarily easy ... but
   - Holographic Algorithms & Gauge Transformations
   - Edge-Binary models of degree ≤ 3
   - Edge-Binary Wick Models (of arbitrary degree)
3. Surface-Easy
   - Kasteleyn Conjecture for Dimer Model on Surface Graphs
   - Edge-Binary Graph-Model which are Surface-Easy
4. Conclusions & Path forward
   - Main "take home" message
   - Where do we go from here?

Binary Graphical Models

Forney style - variables on the edges

$$\mathcal{P}(\vec{\sigma}) = Z^{-1} \prod_a f_a(\vec{\sigma}_a)$$

$$\underbrace{Z = \sum_{\boldsymbol{\sigma}} \prod_a f_a(\vec{\sigma}_a)}_{\text{partition function}}$$



$f_a \geq 0$

$\sigma_{ab} = \sigma_{ba} = \pm 1$

$\vec{\sigma}_1 = (\sigma_{12}, \sigma_{14}, \sigma_{18})$

$\vec{\sigma}_2 = (\sigma_{12}, \sigma_{23})$

- Most Probable Configuration = Maximum Likelihood = Ground State: $\arg\max \mathcal{P}(\vec{\sigma})$
- Marginal Probability: e.g. $\mathcal{P}(\sigma_{ab}) \equiv \sum_{\vec{\sigma} \setminus \sigma_{ab}} \mathcal{P}(\vec{\sigma})$
- **Partition Function: Z** – Our main object of interest

Easy & Difficult Boolean Problems

EASY

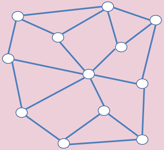
- Any graphical problems on a tree (Bethe-Peierls, Dynamical Programming, BP, TAP and other names)
- Ground State of a Rand. Field Ferrom. Ising model on any graph
- Partition function of planar Ising & Dimer models
- Finding if 2-SAT is satisfiable
- Decoding over Binary Erasure Channel = XOR-SAT
- Some network flow problems (max-flow, min-cut, shortest path, etc)
- Minimal Perfect Matching Problem
- Some special cases of Integer Programming (TUM)

Typical graphical problem, with loops and factor functions of a general position, is DIFFICULT

Glassy Ising & Dimer Models on a Planar Graph

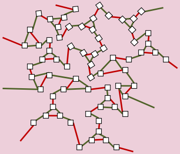
Partition Function of $J_{ij} \gtrless 0$ Ising Model, $\sigma_i = \pm 1$

$$Z = \sum_{\vec{\sigma}} \exp\left(\frac{\sum_{(i,j)\in\Gamma} J_{ij}\sigma_i\sigma_j}{T} \right)$$



Partition Function of Dimer Model, $\pi_{ij} = 0, 1$

perfect matching

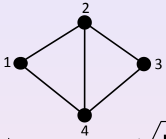


$$Z = \sum_{\vec{\pi}} \prod_{(i,j)\in\Gamma} (z_{ij})^{\pi_{ij}} \prod_{i\in\Gamma} \delta\left(\sum_{j\in i} \pi_{ij}, 1 \right)$$

Ising & Dimer Classics

- L. Onsager, *Crystal Statistics*, Phys.Rev. **65**, 117 (1944)
- M. Kac, J.C. Ward, *A combinatorial solution of the Two-dimensional Ising Model*, Phys. Rev. **88**, 1332 (1952)
- C.A. Hurst and H.S. Green, *New Solution of the Ising Problem for a Rectangular Lattice*, J.of Chem.Phys. **33**, 1059 (1960)
- M.E. Fisher, *Statistical Mechanics on a Plane Lattice*, Phys.Rev **124**, 1664 (1961)
- P.W. Kasteleyn, *The statistics of dimers on a lattice*, Physics **27**, 1209 (1961)
- P.W. Kasteleyn, *Dimer Statistics and Phase Transitions*, J. Math. Phys. **4**, 287 (1963)
- M.E. Fisher, *On the dimer solution of planar Ising models*, J. Math. Phys. **7**, 1776 (1966)
- F. Barahona, *On the computational complexity of Ising spin glass models*, J.Phys. A **15**, 3241 (1982)

Pfaffian solution of the Matching problem

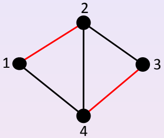


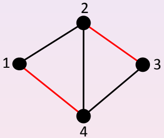
$$Z = z_{12}z_{34} + z_{14}z_{23} = \sqrt{\mathrm{Det}\hat{A}} = \mathrm{Pf}[\hat{A}]$$

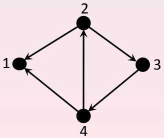
$$\hat{A} = \begin{pmatrix} 0 & -z_{12} & 0 & -z_{14} \\ +z_{12} & 0 & +z_{23} & -z_{24} \\ 0 & -z_{23} & 0 & +z_{34} \\ +z_{14} & +z_{24} & -z_{34} & 0 \end{pmatrix}$$

Odd-face [Kasteleyn] rule (for signs)

Direct edges of the graph such that for every internal face the number of edges oriented clockwise is odd







▸ Fermion/Grassman Representation

Planar Spin Glass and Dimer Matching Problems

The Pfaffian formula with the "odd-face" orientation rule extends to any planar graph thus proving constructively that

- Counting weighted number of dimer matchings on a planar graph is easy
- Calculating partition function of the spin glass Ising model on a planar graph is easy

Planar is generally difficult [Barahona '82]

- Planar spin-glass problem with magnetic field is difficult
- Dimer-monomer matching is difficult even in the planar case

Outline

Are there other graphical models which are easy?

Holographic Algorithms [Valiant '02-'08]

- reduction to dimers via
- "classical" one-to-one gadgets
  (e.g. Ising model to dimer model)
- "holographic" gadgets (e.g. Ice model to Dimer model)
- resulted in discovery of variety of new easy planar models

Gauge Transformations [Chertkov, Chernyak '06-'09]

- Equivalent to the holographic gadgets Gauge Transformations
  (different gauges = different transformations)
- Belief Propagation (BP) Loop Calculus/Series
  is one special choice of the gauge freedom

Are there other graphical models which are easy?

Holographic Algorithms [Valiant '02-'08]

- reduction to dimers via
- "classical" one-to-one gadgets (e.g. Ising model to dimer model)
- "holographic" gadgets (e.g. ▸ Ice model to Dimer model)
- resulted in discovery of variety of new easy planar models

Gauge Transformations [Chertkov, Chernyak '06-'09]

- Equivalent to the holographic gadgets ▸ Gauge Transformations (different gauges = different transformations)
- Belief Propagation (BP) ▸ Loop Calculus/Series is one special choice of the gauge freedom

Are there other graphical models which are easy?

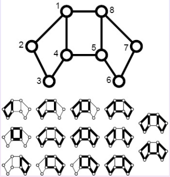
Holographic Algorithms [Valiant '02-'08]

- reduction to dimers via
- "classical" one-to-one gadgets
  (e.g. Ising model to dimer model)
- "holographic" gadgets (e.g. Ice model to Dimer model)
- resulted in discovery of variety of new easy planar models

Gauge Transformations [Chertkov, Chernyak '06-'09]

- Equivalent to the holographic gadgets Gauge Transformations
  (different gauges = different transformations)
- Belief Propagation (BP) Loop Calculus/Series
  is one special choice of the gauge freedom

BP+ for Planar [degree ≤ 3]



Loop Series (general)
[MC,Chernyak '06]

$$Z = Z_0 \cdot z, \;\; z \equiv 1 + \sum_C r_C$$

Summing 2-regular (closed curve) partition is easy!!
[MC,Chernyak,Teodorescu '08]

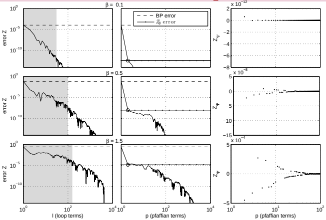
$$Z_s = Z_0 \cdot z_s, \quad z_s = 1 + \sum_{C \in \mathcal{G}}^{\forall a \in C,\ |\delta(a)|_C = 2} r_C$$

[JSTAT '08]

Efficient Approximate Scheme
[Gomez,MC,Kappen '09]

http://arXiv.org/abs/0901.0786

UAI, 2009 + submitted to JML



Easy Models of degree ≤ 3 [MC,Chernyak,Teodorescu '08]

Generic planar problem is difficult

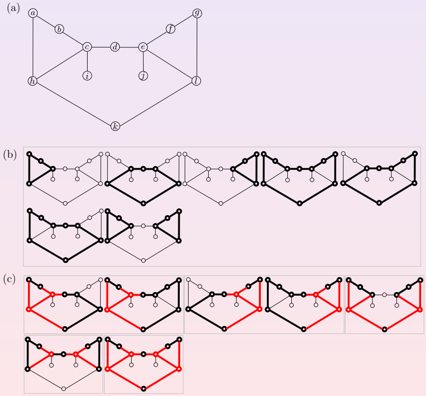
A planar problem is easy if

- the factor functions satisfy

$$\forall\, a \in \mathcal{G}: \quad \sum_{\vec{\sigma}_a} f_a(\vec{\sigma}_a) \times \prod_{b}^{(a,b)\in\mathcal{E}} \exp\left(\eta_{ab}\sigma_{ab}\right) \times \left(\sigma_{ab} - \tanh\left(\eta_{ab} + \eta_{ba}\right)\right) = 0$$

where η are messages from a BP solution for the model

- i.e. when all (!!) "three-colorings" are zero after a BP-transformation [BP gauge= all (!!) "one-colorings" are zero]

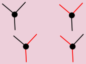


"three-colorings" are shown in red

Easy Models of degree ≤ 3 (II)

To describe the family of easy edge-binary models of degree not larger than three (partition function is reducible to Pfaffian of a $|\mathcal{G}_1| \times |\mathcal{G}_1|$-dimensional skew-symmetric matrix) one needs to:

Item #1: Generate an arbitrary factor-function set which satisfies: $\forall a: \quad W^{(a)}(\vec{\sigma}_a) = 0 \quad \text{if} \quad \sum_{b\sim a} \sigma_{ab} \neq 0 (\text{mod } 2)$



Item #2: Apply an arbitrary skew-orthogonal Gauge-transformation:

$$W^{(a)}(\boldsymbol{\pi}_a) \to f_a(\boldsymbol{\pi}_a) = \sum_{\boldsymbol{\pi}'_a} \left(\prod_{b\sim a} G_{ab}(\pi_{ab}, \pi'_{ab}) \right) W^{(a)}(\boldsymbol{\pi}'_a)$$

$$\forall \{a, b\} \in \mathcal{G}_1: \quad \sum_{\pi} G_{ab}(\pi, \pi') G_{ba}(\pi, \pi'') = \delta(\pi', \pi'')$$

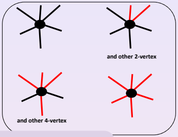
$$Z = \sum_{\boldsymbol{\pi}} \prod_{a\in\mathcal{G}_0} f_a(\boldsymbol{\pi}_a) = \sum_{\boldsymbol{\pi}} \prod_{a\in\mathcal{G}_0} \left(\sum_{\boldsymbol{\pi}'_a} \left(\prod_{b\sim a} G_{ab}(\pi_{ab}, \pi'_{ab}) \right) W^{(a)}(\boldsymbol{\pi}_a) \right)$$

Next Step:

Generalize construction (Item #1) to degree> 3 [Item #2 is already generic]

Edge Binary Wick (EBW) Models [Chernyak, MC '09]

$$Z_{EBW}(W)=\sum_{\gamma=\{\gamma_{ab}\}\in\mathcal{Z}_1(\mathcal{G};\mathbb{Z}_2)}\ \prod_{b\in\mathcal{G}_0}^{\sum_{a\sim b}\gamma_{ab}\neq 0} W^{(b)}_{\{a_1,\cdots,a_{2k}\}\equiv\{a|a\sim b;\gamma_{ab}=1\}}$$



- All odd weights are zero
- Even ($d>2$) weights are expressed via pair-wise weights

$$W^{(b)}_{\{a_1,\cdots,a_{2k}\}}\equiv\sum_{\xi\in P([2k-1])}W^{(b)}_{\xi,a_1\cdots a_{2k}},\quad W^{(b)}_{\xi,a_1\cdots a_{2k}}\equiv(-1)^{\overbrace{\sum_{p,p'\in\xi}^{p<p'} C_{\alpha(p)}\cdot C_{\alpha(p)'}}^{\text{number of crossings (mod 2)}}}\cdot\prod_{p\in\xi}W^{(b)}_{\alpha(p)}$$

Examples of 6-colorings and extensions of a EBW-model 6 vertex

$W_{16}W_{25}W_{34}$ [zero crossing] $\quad -W_{12}W_{35}W_{46}$ [one crossing] $\quad W_{13}W_{25}W_{46}$ [two crossings] $\quad -W_{14}W_{25}W_{36}$ [three crossings]

Edge Binary Wick Models (II)

Known Easy Planar Graphical Models & EBW

$\exists$ a gauge transformation reducing any easy planar model to a EBW

- Dimer Model
- Ising Model
- Ice Model
- Possibly all models discussed in the "holographic" papers

Any EBW model on a planar graph is EASY

- Equivalent to Gaussian Grassman Models on the same graph
- Partition function is Pfaffian of a $|\mathcal{G}_1| \times |\mathcal{G}_1|$ matrix

Related Grassmann/Fermion Models

Vertex Gaussian Grassmann Graphical (VG^3) Models

$$Z_{VG^3}(\varsigma,\boldsymbol{\sigma};\mathbf{W}) = \frac{\int \exp\left(\frac{1}{2}\sum_{(b\to a\to c)\in\mathcal{G}_1} \varphi_{ab}\varsigma^{(a)}_{bc}W^{(a)}_{bc}\varphi_{ac}\right)\exp\left(\frac{1}{2}\sum_{(a,b)\in\mathcal{G}_1}\varphi_{ab}\sigma_{ab}\varphi_{ba}\right)\prod_{(a,b)} d\varphi_{ab}}{\int \exp\left(\frac{1}{2}\sum_{(a,b)\in\mathcal{G}_1}\varphi_{ab}\sigma_{ab}\varphi_{ba}\right)\prod_{(a,b)} d\varphi_{ab}}$$

$$= \frac{\text{Pf}(H(\varsigma,\boldsymbol{\sigma};\mathbf{W}))}{\text{Pf}(H(\varsigma,\boldsymbol{\sigma};\mathbf{0}))}, \quad H_{ij} = \begin{cases} \varsigma^{(a)}_{bc}W^{(a)}_{bc}, & i=(a,b)\ \&\ j=(a,c),\ \text{where } b\neq c\sim a, \\ \sigma_{ab}, & i=(a,b),\ \&\ j=(b,a). \end{cases}$$

Grassmann (anti-commuting) variables: $\forall (a,b),(c,d)\in\mathcal{G}_1 \quad \varphi_{ab}\varphi_{cd} = -\varphi_{cd}\varphi_{ab}$

Berezin (formal) integration rules: $\forall (a,b)\in\mathcal{G}_1: \quad \int d\varphi_{ab} = 0, \quad \int \varphi_{ab}d\varphi_{ab} = 1$

Main Theorem of [Chernyak,MC '09/planar]

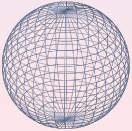
- $\exists \sigma,\varsigma = \pm 1: \quad$ s.t. $\quad Z_{VG^3}(\varsigma,\boldsymbol{\sigma};\mathbf{W}) = Z_{EBW}(\mathbf{W})$
- The special configuration of σ,ς corresponds to Kastelyan (spinor) orientation on the extended planar graph

Outline

Dimer Model on Surface Graphs (I)

Partition function of dimer model on a surface graph of genus g is expressed in terms of a (± 1)-weighted sum over 2^{2g} determinants $=$ surface-easy

- Kasteleyn '63;'67 - non-constructive (??) conjecture
- Gallucio, Loebl '99 - first [combinatorial] proof
- Cimasoni, Reshetikhin '07 - topological proof and relation to gauge fermion models



genus $g = 0$

genus $g = 1$

genus $g = 2$

Dimer Model on Surface Graphs (II)

Partition Function of Dimer Model, $\pi_{ij} = 0, 1$, on a surface graph $\mathcal{G}$

$$Z(\mathcal{G}; \mathbf{z}) = \sum_{\vec{\pi}}^{\text{dimers}} \prod_{(i,j)\in\Gamma} (z_{ij})^{\pi_{ij}}$$

Theorem: (formulation of Cimasoni, Reshetikhin)

$$Z(\mathcal{G}; \mathbf{z}) = \frac{1}{2^g} \sum_{[\mathbf{s}]} \underbrace{\text{Arf}(q^{\mathbf{s}}_{\boldsymbol{\pi}_0})\varepsilon^{\mathbf{s}}(\boldsymbol{\pi}_0)}_{=\pm 1;\ \boldsymbol{\pi}_0-\text{independent};\ \text{depends only on } [\mathbf{s}]} \text{Pf}(A^{\mathbf{s}}(z))$$

- $\boldsymbol{\pi}_0$ is a reference dimer configuration
- $\mathbf{s}$ is a Kasteleyn orientation; $[\mathbf{s}]$ equivalence classes of the Kasteleyn orientations, 2^{2g} of them
- $\varepsilon^{\mathbf{s}}(\boldsymbol{\pi}) = \pm 1$ defines total signature of the dimer configuration π wrt the Kasteleyn orientation $\mathbf{s}$
- $q^{\mathbf{s}}_{\boldsymbol{\pi}_0}(\alpha)$ is a well-defined quadratic form associated with $\mathbf{s}$, $\boldsymbol{\pi}_0$ and α is a closed curve on $\mathcal{G}$; $\text{Arf}(q^{\mathbf{s}}_{\boldsymbol{\pi}_0})$ is the Arf-invariant of the quadratic form.

Dimer Model on Surface Graphs (III)

[Cimasoni, Reshetikhin]

$$Z(\mathcal{G};\mathbf{z}) = \frac{1}{2^g}\sum_{[\mathbf{s}]} \mathrm{Arf}(q^{\mathbf{s}}_{\boldsymbol{\pi}_0})\varepsilon^{\mathbf{s}}(\boldsymbol{\pi}_0)\mathrm{Pf}(A^{\mathbf{s}}(z))$$

- the sum over determinants can be transformed into the sum over partition functions of Kasteleyn-fermion models
- Kasteleyn orientation is a discrete version of spin(or) structures [from topological field theories]
- Powerful derivation techniques from topology [homology and immersion theories]

Generic graphical model on a surface graph is SURFACE-DIFFICULT

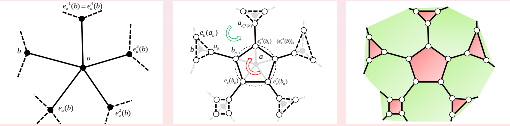
Our next task is:
To classify graphical models which are SURFACE-EASY

Edge-Binary-Wick (EBW) Models and Vertex Gaussian Grassman Graphical (VG^3) models on Surface Graphs

Main Theorem of [Chernyak,MC '09/surface]

$Z_{EBW}(\mathbf{W})Z_{EBW}(\mathbf{1}) = \sum_{[\mathbf{s}]} Z_{VG^3}([\mathbf{s}];\mathbf{1})Z_{VG^3}([\mathbf{s}];\mathbf{W})$ where

- $\mathbf{s} = (\sigma;\varsigma)$ corresponds to a Kastelyan/spinor orientation defined on extended graph
- $[\mathbf{s}]$ are equivalence classes (2^{2g} of them) of the Kastelyan/spinor $\mathbf{s}$ orientations



EBW and VG^3 models on Surface Graphs (II)

$$Z_{EBW}(\mathbf{W})Z_{EBW}(\mathbf{1}) = \sum_{[\mathbf{s}]} Z_{VG^3}([\mathbf{s}];\mathbf{1})Z_{VG^3}([\mathbf{s}];\mathbf{W})$$

The multi-step proof of the main surface theorem includes

- Extended/fat graph construction and partitioning $\boldsymbol{\xi}$ of the even generalized loop γ configurations into closed curves [Wick structure]
- Analysis and relation between invariant objects (quadratic forms) for the generalized loops, $[\gamma]$, and spinors, $[s]$, defined on fat graphs and respective Riemann surfaces.
- Term by term comparison of the relation between the partial $\tilde{Z}_{EBW}([\gamma];\mathbf{W})$ and $\tilde{Z}_{VG^3}([\gamma],[\mathbf{s}];\mathbf{W})$, where $Z_{EBW}(\mathbf{W}) = \sum_{[\gamma]} \tilde{Z}_{EBW}([\gamma];\mathbf{W})$ and $Z_{VG^3}([\mathbf{s}];\mathbf{W}) = \sum_{[\gamma]} \tilde{Z}_{VG^3}([\gamma],[\mathbf{s}];\mathbf{W})$. This results in the system of 2^{2g} linear equations for 2^{2g} unknowns $\tilde{Z}_{EBW}([\gamma];\mathbf{W})$.
- Solving the linear equations we recover the main statement of the theorem.
- $2^g Z_{VG^3}([\mathbf{s}];\mathbf{1}) = \mathrm{Arf}(q([s]))Z_{EBW}(\mathbf{1})$, where $q(s)(\gamma) = q([s])([\gamma])$ is a well-defined quadratic form.

Q:

Describe the family of surface-easy edge-binary models on an arbitrary surface graph $\mathcal{G}$ (partition function is reducible to a sum of 2^{2g} Pfaffians)

A: [constructive]

- Generate an arbitrary Vertex Gaussian Grassmann binary-Gauge (VG^3) Model on the graph
- Fix the binary-gauge according to the Kasteleyn (spinor) rule on the extended graph
- Construct respective Edge-Binary Wick model on the original graph
- Apply an arbitrary skew-orthogonal (holographic) gauge/transformation

The partition function of the resulting model is the sum of 2^{2g} $\pm$-weighted Pfaffians. [All terms in the sum are explicitly known.]

Future work

- Use the described hierarchy of easy planar models as a basis for efficient variational approximation of generic (difficult) planar problems. (The approach may also be useful for building efficient variational matrix-product state wave functions for quantum models. Dynamical Bayesian Networks: 1+1, tree+1, ...)
- Study Wick Gaussian models on non-planar but Pfaffian orientable or k-Pfaffian orientable graphs (where any dimer model on surface graph of genus g is 2^{2g}-Pfaffian orientable).
- Almost Planar = Geographical Graphical Models, Renormalization Group, Generalized BP
- Analogs of all of the above for Surface-Difficult Problems

Example (1): Statistical Physics

Ising model $\sigma_i = \pm 1$

$$\mathcal{P}(\vec{\sigma}) = Z^{-1} \exp\left(\sum_{(i,j)} J_{ij}\sigma_i\sigma_j\right)$$

J_{ij} defines the graph (lattice)

Graphical Representation

Variables are usually associated with vertexes ... but transformation to the Forney graph (variables on the edges) is straightforward

- Ferromagnetic ($J_{ij} < 0$), Anti-ferromagnetic ($J_{ij} > 0$) and Frustrated/Glassy
- Magnetization (order parameter) and Ground State
- Thermodynamic Limit, $N \to \infty$
- Phase Transitions

Binary Graphical Models

Example (2): Information Theory, Machine Learning, etc

Probabilistic Reconstruction (Statistical Inference)

$\vec{\sigma}_{\text{orig}}$	$\Rightarrow$	$\vec{x}$	$\Rightarrow$	$\vec{\sigma}$
original *data* $\vec{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\vec{x}\|\vec{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\vec{\sigma} \in \mathcal{C}$

Maximum Likelihood [ground state]

Marginalization

$$\text{ML}(\vec{x}) = \arg\max_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma}) \qquad \sigma_i^*(\vec{x}) = \arg\max_{\sigma_i} \sum_{\vec{\sigma}\setminus\sigma_i} \mathcal{P}(\vec{x}|\vec{\sigma})$$

Counting (Partition Function): $Z(\vec{x}) = \sum_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$

Example (2): Information Theory, Machine Learning, etc

Probabilistic Reconstruction (Statistical Inference)

$\vec{\sigma}_{\text{orig}}$	$\Rightarrow$	$\vec{x}$	$\Rightarrow$	$\vec{\sigma}$
original *data* $\vec{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\vec{x}\vert\vec{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\vec{\sigma} \in \mathcal{C}$

Maximum Likelihood [ground state]

Marginalization

$$\text{ML}(\vec{x}) = \arg\max_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$$

$$\sigma_i^*(\vec{x}) = \arg\max_{\sigma_i} \sum_{\vec{\sigma}\setminus\sigma_i} \mathcal{P}(\vec{x}|\vec{\sigma})$$

Counting (Partition Function): $Z(\vec{x}) = \sum_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$

Example (2): Information Theory, Machine Learning, etc

Probabilistic Reconstruction (Statistical Inference)

$\vec{\sigma}_{\text{orig}}$	$\Rightarrow$	$\vec{x}$	$\Rightarrow$	$\vec{\sigma}$
original *data* $\vec{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\vec{x}\|\vec{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\vec{\sigma} \in \mathcal{C}$

Maximum Likelihood [ground state]

Marginalization

$$\text{ML}(\vec{x}) = \arg\max_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$$

$$\sigma_i^*(\vec{x}) = \arg\max_{\sigma_i} \sum_{\vec{\sigma}\setminus\sigma_i} \mathcal{P}(\vec{x}|\vec{\sigma})$$

Counting (Partition Function): $Z(\vec{x}) = \sum_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$

Binary Graphical Models

Example (2): Information Theory, Machine Learning, etc

Probabilistic Reconstruction (Statistical Inference)

$\vec{\sigma}_{\rm orig}$	$\Rightarrow$	$\vec{x}$	$\Rightarrow$	$\vec{\sigma}$
original *data* $\vec{\sigma}_{\rm orig}\in\mathcal{C}$ codeword	noisy channel $\mathcal{P}(\vec{x}\vert\vec{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\vec{\sigma}\in\mathcal{C}$

Maximum Likelihood [ground state]

Marginalization

$$\mathrm{ML}(\vec{x}) = \arg\max_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$$

$$\sigma_i^*(\vec{x}) = \arg\max_{\sigma_i} \sum_{\vec{\sigma}\setminus\sigma_i} \mathcal{P}(\vec{x}|\vec{\sigma})$$

Counting (Partition Function): $Z(\vec{x}) = \sum_{\vec{\sigma}} \mathcal{P}(\vec{x}|\vec{\sigma})$

Grassmann (fermion, nilpotent) Calculus for Pfaffians

Grassman (nilpotent) Variables on Vertexes

$$\forall (a,b) \in \mathcal{G}_e: \quad \theta_a\theta_b + \theta_b\theta_a = 0 \quad \int d\theta = 0, \quad \int \theta d\theta = 1$$

Pfaffian as a Gaussian Berezin Integral over the Fermions

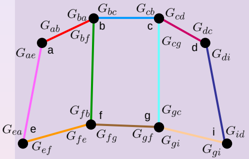
$$\int \exp\left(-\frac{1}{2}\vec{\theta}^t \hat{A}\vec{\theta}\right) d\vec{\theta} = \mathrm{Pf}(\hat{A}) = \sqrt{\det(\hat{A})}$$

Pfaffian Formula

Gauge Transformations

Chertkov, Chernyak '06

Local Gauge, G, Transformations



$$Z = \sum_{\vec{\sigma}} \prod_a f_a(\vec{\sigma}_a), \;\; \vec{\sigma}_a = (\sigma_{ab}, \sigma_{ac}, \cdots)$$

$$\sigma_{ab} = \sigma_{ba} = \pm 1$$

$$f_a(\vec{\sigma}_a = (\sigma_{ab}, \cdots)) \to \sum_{\sigma'_{ab}} G_{ab}(\sigma_{ab}, \sigma'_{ab}) f_a(\sigma'_{ab}, \cdots)$$

$$\sum_{\sigma_{ab}} G_{ab}(\sigma_{ab}, \sigma') G_{ba}(\sigma_{ab}, \sigma'') = \delta(\sigma', \sigma'')$$

The partition function is invariant under any G-gauge!

$$Z = \sum_{\vec{\sigma}} \prod_a f_a(\vec{\sigma}_a) = \sum_{\vec{\sigma}} \prod_a \left(\sum_{\vec{\sigma}'_a} f_a(\vec{\sigma}'_a) \prod_{b \in a} G_{ab}(\sigma_{ab}, \sigma'_{ab}) \right)$$

Belief Propagation as a Gauge Fixing

Chertkov, Chernyak '06

$$Z=\sum_{\vec{\sigma}}\prod_a f_a\left(\vec{\sigma}_a\right)=\sum_{\boldsymbol{\sigma}}\prod_a\left(\sum_{\vec{\sigma}'_a} f_a(\vec{\sigma}'_a)\prod_{b\in a} G_{ab}(\sigma_{ab},\sigma'_{ab})\right)$$

$$Z=\underbrace{Z_0(G)}_{\text{ground state}}+\underbrace{\sum_{\text{all possible colorings of the graph}} Z_c(G)}_{\vec{\sigma}\neq+\vec{1},\ \text{excited states}}$$

ground state $\vec{\sigma}=+\vec{1}$

$\vec{\sigma}\neq+\vec{1}$, excited states

Belief Propagation Gauge

$\forall a\ \&\ \forall b\in a:$

$$\sum_{\vec{\sigma}'_a} f_a(\vec{\sigma}')G^{(bp)}_{ab}(\sigma_{ab}=-1,\sigma'_{ab})\prod_{c\in a}^{c\neq b} G^{(bp)}_{ac}(+1,\sigma'_{ac})=0$$

No loose BLUE=colored edges at any vertex of the graph!

Holographic Gadgets & Gauges

Belief Propagation as a Gauge Fixing (II)

$\forall a \;\&\; \forall b \in a:$

$$\left\{\begin{array}{l}\sum_{\vec{\sigma}'_a} f_a(\vec{\sigma}') G^{(bp)}_{ab}(-1,\sigma'_{ab}) \prod_{c\in a}^{c\neq b} G^{(bp)}_{ac}(+1,\sigma'_{ac}) = 0 \\ \sum_{\sigma_{ab}} G_{ab}(\sigma_{ab},\sigma') G_{ba}(\sigma_{ab},\sigma'') = \delta(\sigma',\sigma'')\end{array}\right. \Rightarrow \left\{\begin{array}{l} G^{(bp)}_{ba}(+1,\sigma'_{ab}) = \rho_a^{-1} \overbrace{\sum_{\vec{\sigma}'_a\backslash\sigma'_{ab}} f_a(\vec{\sigma}') \prod_{c\in a}^{c\neq b} G^{(bp)}_{ac}(+1,\sigma'_{ac})}^{sum-product} \\ \rho_a = \sum_{\vec{\sigma}'_a} f_a(\vec{\sigma}') \prod_{c\in a} G^{(bp)}_{ac}(+1,\sigma'_{ac})\end{array}\right.$$

Belief Propagation in terms of Messages

$$G^{(bp)}_{ab}(+1,\sigma) = \frac{\exp(\sigma\eta_{ab})}{2\sqrt{\cosh(\eta_{ab}+\eta_{ba})}}, \quad G^{(bp)}_{ab}(-1,\sigma) = \sigma\frac{\exp(-\sigma\eta_{ba})}{2\sqrt{\cosh(\eta_{ab}+\eta_{ba})}} \Longrightarrow$$

$$\sum_{\vec{\sigma}_a\backslash\sigma_{ab}} f_a(\vec{\sigma}_a)\exp\left(\sum_{c\in a}\sigma_{ac}\eta_{ac}\right)\left(\sigma_{ab} - \tanh(\eta_{ab}+\eta_{ba})\right) = 0$$

$$b_a(\vec{\sigma}_a) = \frac{f_a(\vec{\sigma}_a)\exp\left(\sum_{b\in a}\sigma_{ab}\eta_{ab}\right)}{\sum_{\vec{\sigma}_a} f_a(\vec{\sigma}_a)\exp\left(\sum_{b\in a}\sigma_{ab}\eta_{ab}\right)}, \quad b_{ab}(\sigma) = \frac{\exp(\sigma(\eta_{ab}+\eta_{ba}))}{\sum_\sigma \exp(\sigma(\eta_{ab}+\eta_{ba}))}$$

Holographic Gadgets & Gauges

Loop Series:

Chertkov,Chernyak '06

Exact (!!) expression in terms of BP

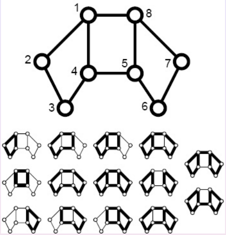
$$Z = \sum_{\vec{\sigma}_\sigma} \prod_a f_a(\vec{\sigma}_a) = Z_0 \left(1 + \sum_C r(C)\right)$$

$$r(C) = \frac{\prod_{a \in C} \mu_a}{\prod_{(ab) \in C} (1 - m_{ab}^2)} = \prod_{a \in C} \tilde{\mu}_a$$

$C \in$ Generalized Loops $=$ Loops without loose ends

$$m_{ab} = \sum_{\vec{\sigma}_a} b_a^{(bp)}(\vec{\sigma}_a)\sigma_{ab}$$

$$\mu_a = \sum_{\vec{\sigma}_a} b_a^{(bp)}(\vec{\sigma}_a) \prod_{b \in a, C} (\sigma_{ab} - m_{ab})$$



- The Loop Series is finite
- All terms in the series are calculated within BP
- BP is exact on a tree
- BP is a Gauge fixing condition. Other choices of Gauges would lead to different representation.

Holographic Gadgets & Gauges

Ice Model [vertexes of max degree 3]

#PL-3-NAE-ICE [Valiant '02]

- Input: A planar graph G = (V;E) of maximum degree 3.
- Output: The number of orientations (arrows) such that no node has all the edges directed towards it or away from it.

From arrows to binary variables

- Edge $\{a, b\}$ is broken in two by insertion of $a - b$ vertex
- Introduce binary variables s.t. if $a \to b \Rightarrow \pi_{a,a-b} = 0, \pi_{b,a-b} = 1$ $b \to a \Rightarrow \pi_{a,a-b} = 1, \pi_{b,a-b} = 0$

$$Z_{ice} = \sum_{\boldsymbol{\pi}'} \left(\prod_{a \in \mathcal{G}_0} f_a(\tilde{\boldsymbol{\pi}}_a) \right) \left(\prod_{\{a,b\} \in \mathcal{G}_1} g_{a-b}(\pi_{a,a-b}, \pi_{b,a-b}) \right)$$

$$f_a(\boldsymbol{\pi}'_a) = \begin{cases} 1, & \exists\ b, c \in \delta_{\mathcal{G}}(a), \quad \text{s.t.} \quad \pi_{a,a-b} \neq \pi_{a,a-c} \\ 0, & \text{otherwise} \end{cases}$$

$$g_{a-b}(\boldsymbol{\pi}'_a) = \begin{cases} 1 & \pi_{a,a-b} \neq \pi_{b,a-b} \\ 0, & \text{otherwise} \end{cases}$$

▸ Holographic Gadgets & Gauges

Ice Model [vertexes of max degree 3] II

General Gauge Transformation

$$f_a(\boldsymbol{\pi}_a) \to \tilde{f}_a(\boldsymbol{\pi}_a) = \sum_{\boldsymbol{\pi}'_a} \left(\prod_{b\sim a} G_{ab}(\pi_{ab}, \pi'_{ab}) \right) f_a(\boldsymbol{\pi}'_a)$$

$$\forall \{a, b\} \in \mathcal{G}_1 : \quad \sum_{\pi} G_{ab}(\pi, \pi') G_{ba}(\pi, \pi'') = \delta(\pi', \pi'')$$

$$Z = \sum_{\boldsymbol{\pi}} \prod_{a\in\mathcal{G}_0} \tilde{f}_a(\boldsymbol{\pi}_a) = \sum_{\boldsymbol{\pi}} \prod_{a\in\mathcal{G}_0} \left(\sum_{\boldsymbol{\pi}'_a} \left(\prod_{b\sim a} G_{ab}(\pi_{ab}, \pi'_{ab}) \right) f_a(\boldsymbol{\pi}_a) \right)$$

Gauge Transformation for the Ice model

$$G^{(ice)}_{a,a-b} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \tilde{g}_{a-b}(\boldsymbol{\pi}'_a) = \begin{cases} 1, & \pi_{a,a-b} = \pi_{b,a-b} = 0 \\ -1, & \pi_{a,a-b} = \pi_{b,a-b} = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{f}_a(\pi_{a,a-1}, \pi_{a,a-2}, \pi_{a,a-3}) = \frac{3}{\sqrt{2}} * \begin{cases} 1, & \pi_{a,a-1} = \pi_{a,a-2} = \pi_{a,a-3} = 0 \\ -1/3, & \sum_i \pi_{a,a-i} = 2 \\ 0, & \text{otherwise} \end{cases}$$