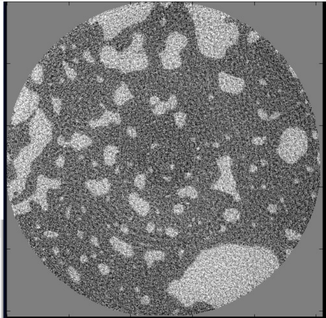
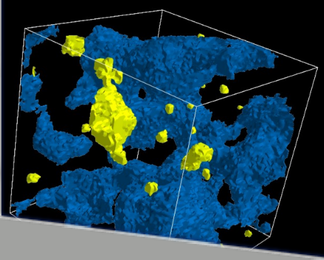
Belief Propagation Reconstruction for Discrete X-ray Tomography

Emmanuelle Gouillart
Joint unit CNRS/Saint-Gobain
Surface of Glass and Interfaces

Florent Krzakala (ESPCI), Marc Mézard (ENS Paris), Lenka Zdeborova (IPhT)

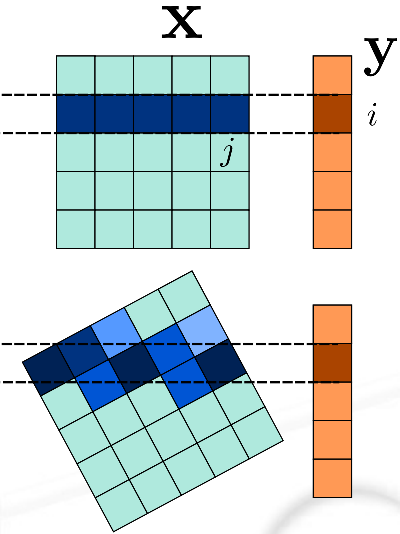




Acknowledgements

- ESRF team : Elodie Boller, Alessandro Mirone, Paul Tafforeau, Alexander Rack
- algorithms : Emilie Chouzenoux, Hugues Talbot, Gaël Varoquaux
- binary data : David Bouttes, Damien Vandembroucq
- funding : ESRF, Agence Nationale de la Recherche (EDDAM project)

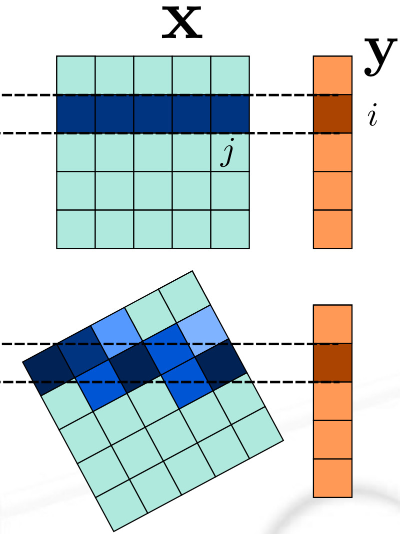
Tomographic reconstruction : an inverse problem



Projections of the sample

along several directions

Tomographic reconstruction : an inverse problem



Projections of the sample along several directions

$$\mathbf{y} = \mathbf{A}\,\mathbf{x} + \mathbf{w}$$

$$\mathbf{y} \in \mathbb{R}^M, \quad \mathbf{x} \in \mathbb{R}^N$$

$\mathbf{w}$: noise

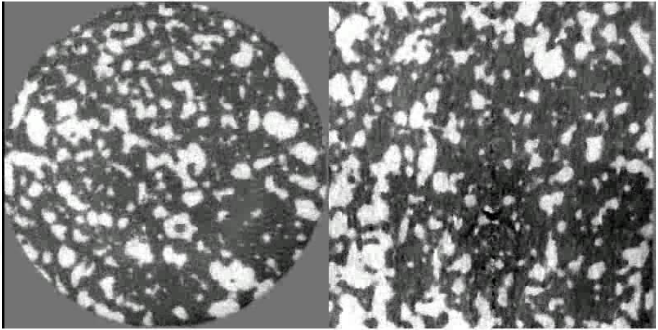
$$\alpha = \frac{M}{N}$$

α : undersampling rate

We want to find x knowing A and y

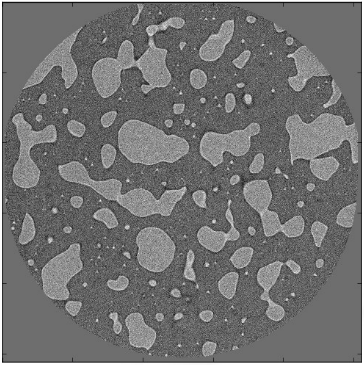
In-situ tomography : image fast transformations

David Bouttes, Damien Vandembroucq (PMMH)
coarsening in phase-separated glasses at high temperature

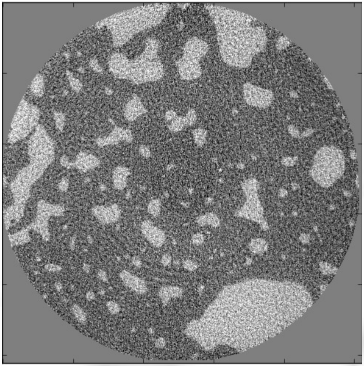


Problems for degraded acquisition

Courtesy David Bouttes



ex-situ

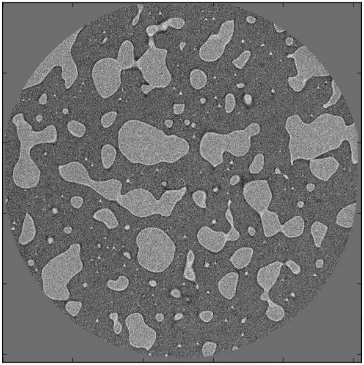


in-situ : less measurements, more noise

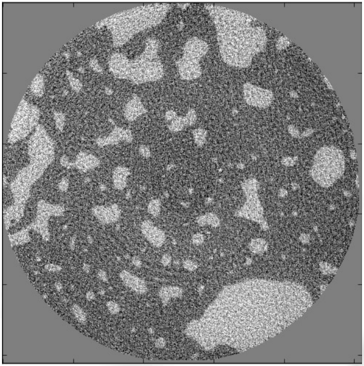
$\mathbf{y} = \mathbf{A}\,\mathbf{x} + \mathbf{w}$: undetermined / ill-conditioned system
$\rightarrow$ a subspace of $A^\mathsf{T} A$ is wrongly reconstructed

Problems for degraded acquisition

Courtesy David Bouttes



ex-situ



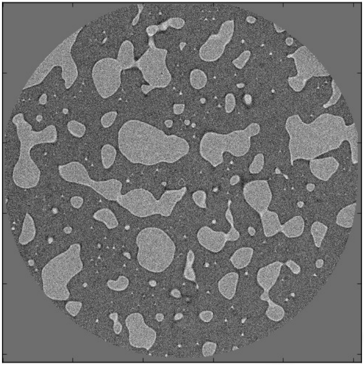
in-situ : less measurements, more noise

$\mathbf{y} = \mathbf{A}\,\mathbf{x} + \mathbf{w}$: undetermined / ill-conditioned system
$\rightarrow$ a subspace of $A^\mathsf{T} A$ is wrongly reconstructed

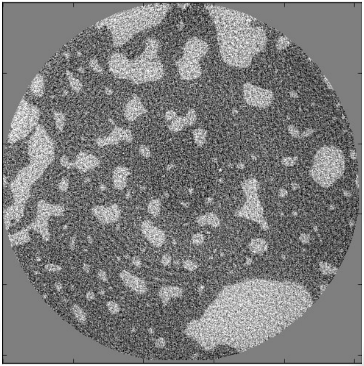
Can we use additional constraints to improve the reconstruction ?

Problems for degraded acquisition

Courtesy David Bouttes



ex-situ



in-situ : less measurements, more noise

Can we use additional constraints to improve the reconstruction ?

Idea : require the solution $\mathbf{x}$ to be a **binary image**

- $\mathbf{x} \in [0, 1]^N$ (after normalization)
- spatial regularity (domains)

Bayesian formulation of the reconstruction problem

Probabilistic formulation of the inverse problem.

$$\mathcal{P}(\mathbf{x}|\mathbf{y}) \propto \mathcal{P}(\mathbf{y}|\mathbf{x})\ \mathcal{P}(\mathbf{x}) \qquad (\star)$$

“Posterior”	Forward model	Binary prior
Quantity of interest	Tomo geometry	Expectations on $\mathbf{x}$

Several approaches to reconstruct $\mathbf{x}$

- $\mathcal{P}(\mathbf{x}|\mathbf{y}) \ \rightarrow$ expectation of x_i
- **or** compute the **maximum a posteriori** (MAP)

$$\mathrm{argmax}_{\mathbf{x}} \mathcal{P}(\mathbf{x}|\mathbf{y})$$

$\Rightarrow$ solve an optimization problem

MAP approach : convex relaxation of binary problem

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \beta\, \mathbf{TV}(\mathbf{x}) + \mathcal{I}\left([0,1]^N\right)$$

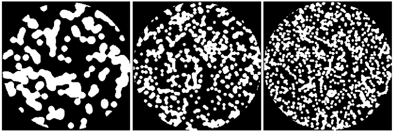
- $\mathbf{TV}(\mathbf{x}) = \sum_i \left\|(\nabla \mathbf{x})_i\right\|_2 = \sum_i \sqrt{(\partial_1 \mathbf{x}_i)^2 + (\partial_2 \mathbf{x}_i)^2}$

  total variation

  penalization of gradients $\rightarrow$ piecewise constant images
- $\mathcal{I}\left([0,1]^N\right)$: convex relaxation of $\mathcal{I}\left(\{0,1\}^N\right)$
- Non-smooth convex optimization : iterative method to find the solution
  - Numerous studies of TV regularization for tomography : [Sidky & Pan, Tang2009, Jia2010]
  - Now available at the ESRF synchrotron
  - Other kinds of spatial regularization may be used (non-convex)

Probabilistic approach : binary model

$\mathbf{x} \in \{-1, 1\}^n$ + domains
$\mathcal{P}(\mathbf{x})$: Ising model
(Markov random field)



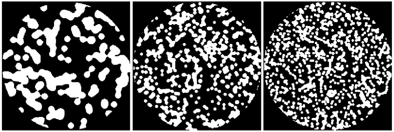
Neighboring pixels more likely to have the same value :

$$\begin{aligned}\mathcal{P}(\mathbf{x}) &\propto e^{\sum_{\langle i,j \rangle} J_{ij} \delta_{x_i, x_j}} \\ &\propto \prod_{\mu} e^{\sum_{(ij) \in \mu} J_{\mu} \delta_{x_i, x_j}}\end{aligned}$$

(factorization along light rays μ)

Probabilistic approach : binary model

$\mathbf{x} \in \{-1, 1\}^n$ + domains
$\mathcal{P}(\mathbf{x})$: Ising model
(Markov random field)



$$\begin{aligned}\mathcal{P}(\mathbf{x}) &\propto e^{\sum_{\langle i,j\rangle} J_{ij}\delta_{x_i,x_j}} \\ &\propto \prod_{\mu} e^{\sum_{(ij)\in\mu} J_\mu \delta_{x_i,x_j}}\end{aligned}$$

(factorization along light rays μ)

Posterior distribution

$$\mathcal{P}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z}\prod_{\mu=1}^{M}\left[\delta\left(y_\mu - \sum_{i\in\mu} x_i\right) e^{J_\mu \sum_{(ij)\in\mu} \delta_{x_i,x_j}}\right]$$

A product of **factors** involving different subsets of pixels

An intractable combinatorial problem

How to compute the marginals of pixel values $\mathcal{P}(x_i)$?

$$\mathcal{P}(x_i) = \sum_{\mathbf{x}' \in \{-1,1\}^N ; x'_i = x_i} \mathcal{P}(\mathbf{x}'|\mathbf{y})$$

Sampling $\mathcal{P}(\mathbf{x}|\mathbf{y})$ over 2^N images is not possible ! $(N \sim 10^5 - 10^6)$

- Gibbs sampling [Liao & Herman] : slow for large images
- Can the expression of the marginal be simplified, using the graph structure ? (ex : using graph cuts for segmentation)

An intractable combinatorial problem

How to compute the marginals of pixel values $\mathcal{P}(x_i)$?

$$\mathcal{P}(x_i) = \sum_{\mathbf{x}' \in \{-1,1\}^N ; x'_i = x_i} \mathcal{P}(\mathbf{x}' | \mathbf{y})$$

Sampling $\mathcal{P}(\mathbf{x}|\mathbf{y})$ over 2^N images is not possible ! ($N \sim 10^5 - 10^6$)

- Gibbs sampling [Liao & Herman] : slow for large images
- Can the expression of the marginal be simplified, using the graph structure ? (ex : using graph cuts for segmentation)

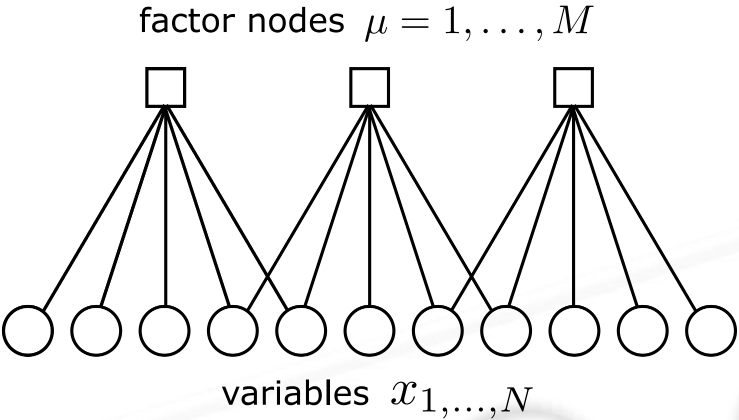
Use **belief propagation**

What is belief propagation ?

A recursive technique for **statistical inference** on **factor graphs** (e.g. Bayesian networks, Markov random fields)

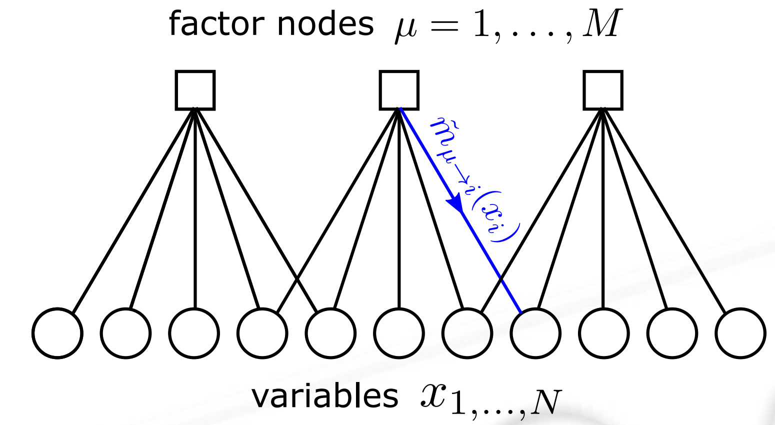
- Different names for related ideas : belief propagation [Pearl 82], sum-product [Gallager 62, LDPC codes], (approximate) message passing, ...
- Factor the computation of marginals of variables as a function of other marginals (supposed to be known).
- Exact and fast on trees
- Often a good approximation on other factor graphs
- Applications : error correcting codes (LDPC), satisfiability, computer vision, ...

Messages defined on factor graph



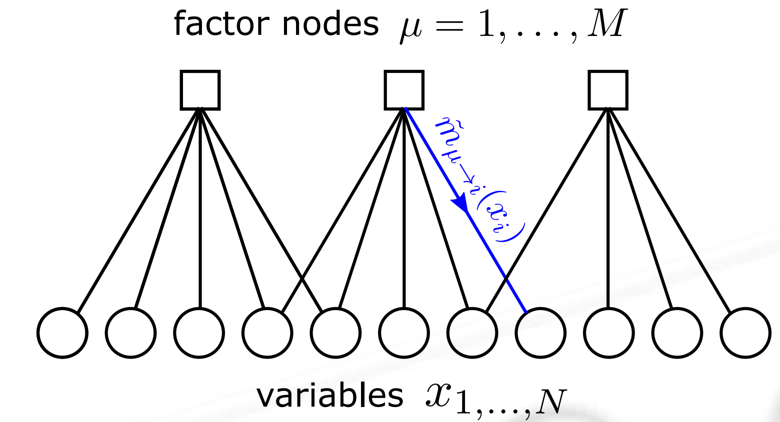
Factor graph : bipartite graph of variables (pixels) and (probability) factors

Messages defined on factor graph



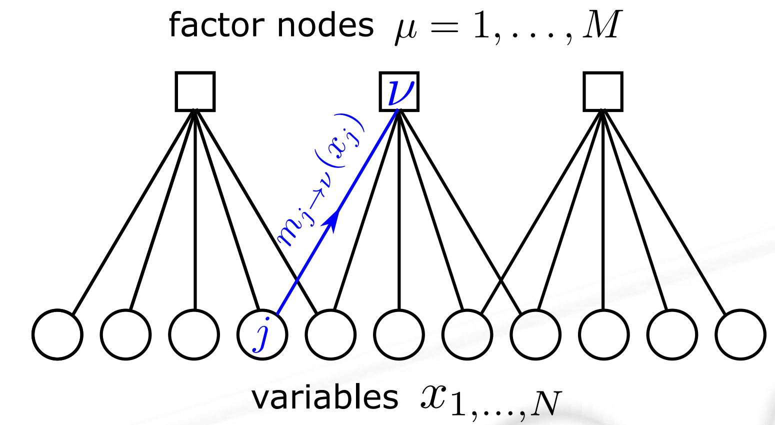
$\tilde{m}_{\mu \to i}(x_i)$: marginal of the variable x_i in absence of all constraints but μ.

Messages defined on factor graph



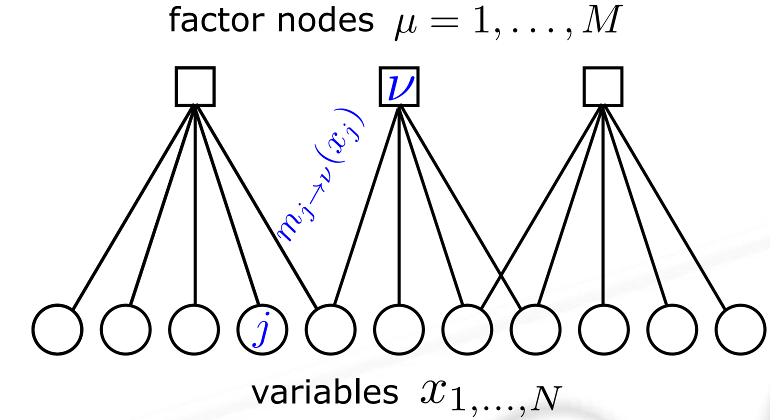
$\tilde{m}_{\mu\to i}(x_i)$: marginal of the variable x_i in absence of all constraints but μ.

Messages defined on factor graph



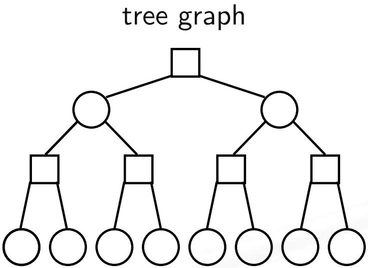
$m_{j\to\nu}(x_j)$: marginal of the variable x_j in absence of constraint ν

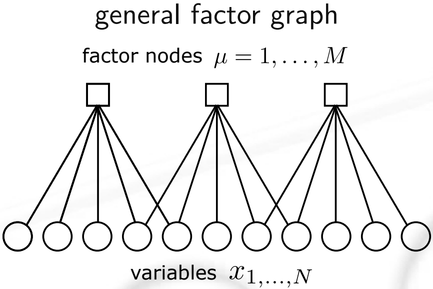
Messages defined on factor graph



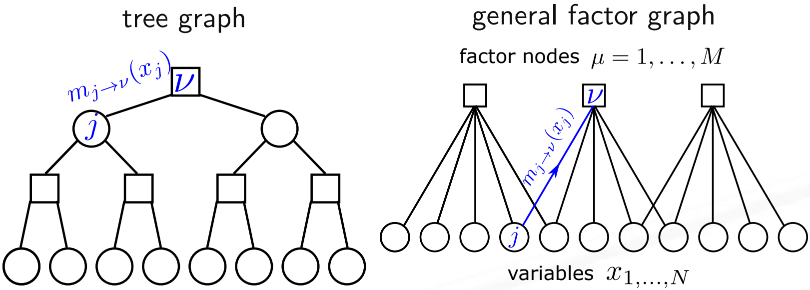
$m_{j\to\nu}(x_j)$: marginal of the variable x_j in absence of constraint ν

Relations between messages





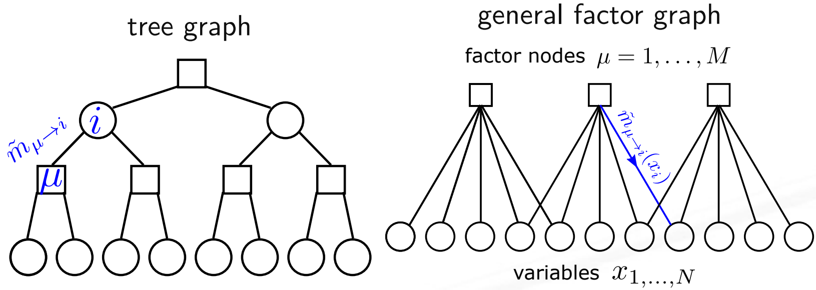
Relations between messages



$$m_{j\to\nu}(x_j) \propto \prod_{\mu\in j\neq\nu} \tilde{m}_{\mu\to j}(x_j)$$

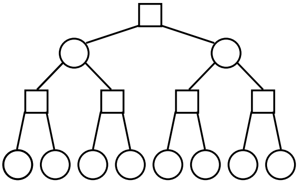
exact for a tree, not for a generic factor graph

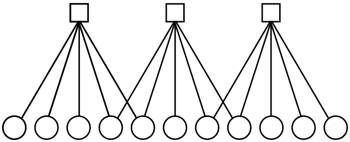
Relations between messages



$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \mathcal{P}_\mu(\{x_j\}, x_i) \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

Relations between messages



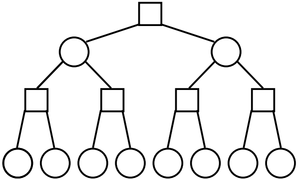


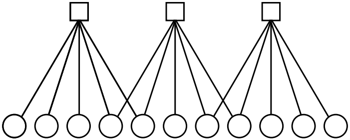
Belief propagation : iterate to solve fixed-point equations

$$m_{j\to\nu}(x_j) \propto \prod_{\mu \in j \neq \nu} \tilde{m}_{\mu\to j}(x_j)$$

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \mathcal{P}_\mu(\{x_j\}, x_i) \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

Relations between messages





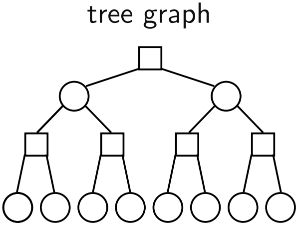
Belief propagation : iterate to solve fixed-point equations

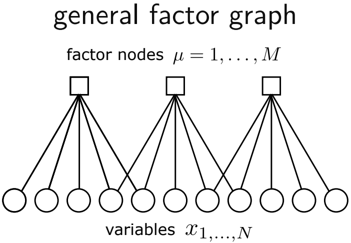
$$m_{j\to\nu}(x_j) \propto \prod_{\mu \in j \neq \nu} \tilde{m}_{\mu\to j}(x_j)$$

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \mathcal{P}_\mu(\{x_j\}, x_i) \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

- Trees : convergence to exact marginals [Pearl 82]

Relations between messages





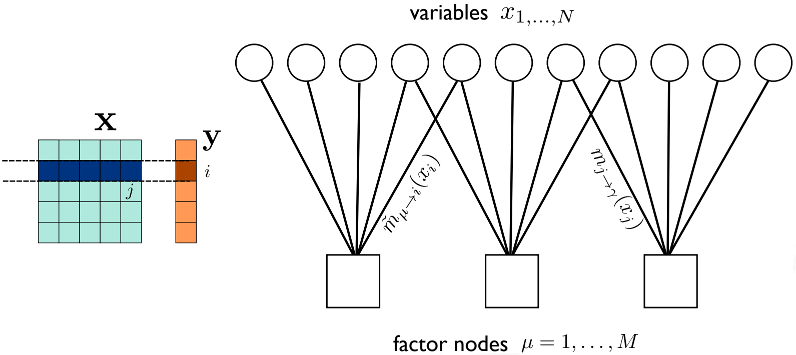
Belief propagation : iterate to solve fixed-point equations

$$m_{j\to\nu}(x_j) \propto \prod_{\mu \in j \neq \nu} \tilde{m}_{\mu\to j}(x_j)$$

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \mathcal{P}_\mu(\{x_j\}, x_i) \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

- Generic graphs : **"loopy" belief propagation** [Murphy 99]. Convergence not guaranteed ; computes approximation of marginals. May work very well even for images (very loopy graphs !)

Belief propagation for binary tomography

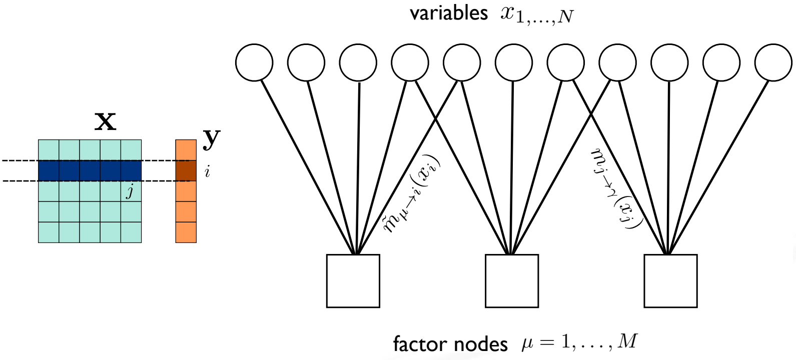


$$m_{i\to\gamma}(x_i) \propto \prod_{\mu \in i \neq \gamma} \tilde{m}_{\mu\to i}(x_i)$$

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \delta\Big(y_\mu - \sum_{j\in\mu} x_j\Big) \quad e^{J_\mu \sum_{(jk)\in\mu} x_j x_k}$$

$$\times \prod\nolimits_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

Belief propagation for binary tomography



$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} \delta\Big(y_\mu - \sum_{j\in\mu} x_j\Big) \quad e^{J_\mu \sum_{(jk)\in\mu} x_j x_k} \times \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j)$$

2^L terms, $L \sim \sqrt{N}$ ☹... BP works well for small number of variables / factor

Trick : use BP within BP !

Some formula massaging...

$$x_i = \pm 1 \ \rightarrow \ m_{i\to\mu}(x_i) \propto e^{x_i h_{i\to\mu}}$$

$$\begin{aligned}
\tilde{m}_{\mu\to i}(x_i) &\propto \sum_{x_j; j\in\mu\neq i} \delta(y_\mu - \sum_{j\in\mu} x_j) e^{J_\mu \sum_{(jk)\in\mu} x_j x_k} \prod_{j\in\mu\neq i} m_{j\to\mu}(x_j) \\
&\propto \sum_{x_j; j\in\mu\neq i} \delta(y_\mu - \sum_{j\in\mu} x_j) e^{J_\mu \sum_{(jk)\in\mu} x_j x_k + \sum_{j\in\mu\neq i} x_j h_{j\to\mu}} .
\end{aligned}$$

and a trick...

Replace the δ by a Lagrange multiplier (canonical formulation)

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} e^{H \sum_i x_i + J_\mu \sum_{(jk)\in\mu} x_j x_k + \sum_{j\in\mu\neq i} x_j h_{j\to\mu}} .$$

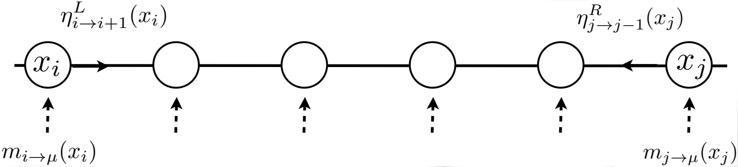
(several iterations to find H)

Trick : use BP within BP !

$$\tilde{m}_{\mu\to i}(x_i) \propto \sum_{x_j; j\in\mu\neq i} e^{H\sum_i x_i + J_\mu \sum_{(jk)\in\mu} x_j x_k + \sum_{j\in\mu\neq i} x_j h_{j\to\mu}} .$$

probability of a configuration : product of factors w/ 1 or 2 variables

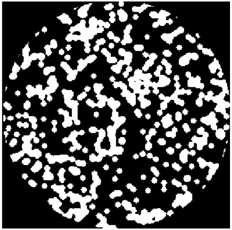
$\Rightarrow$ apply BP on a chain (easy)



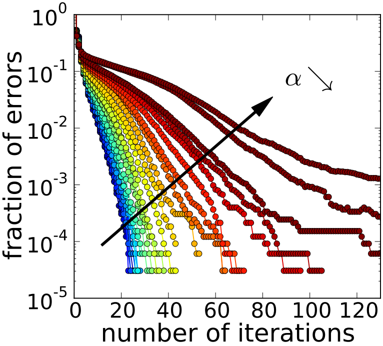
Code : `https://github.com/eddam/bp-for-tomo`

Convergence of the algorithm : it works !

Noisefree measures + perfect model : perfect recovery possible



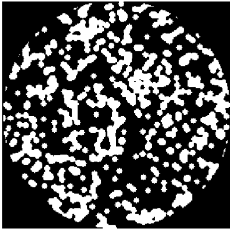
synthetic data
no noise, different α



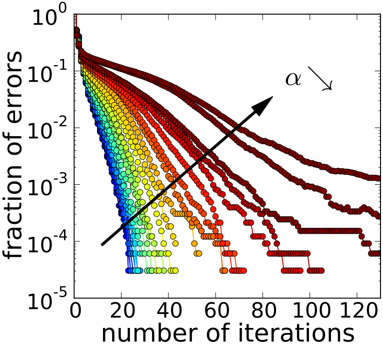
- Fast convergence when α large (enough measurements)
- Convergence slows down when the undersampling rate α decreases (less measurements)
- No convergence above critical α (convergence time diverges)

Convergence of the algorithm : it works !

Noisefree measures + perfect model : perfect recovery possible



synthetic data
no noise, different α

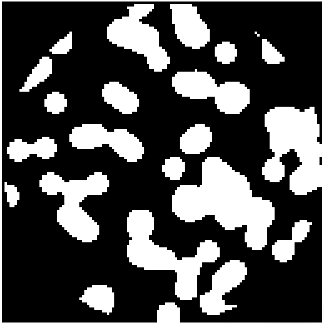


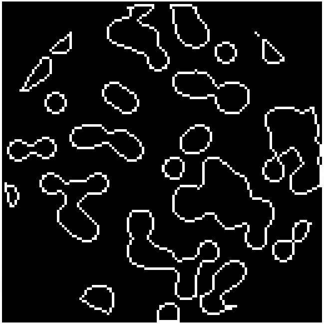
- Fast convergence when α large (enough measurements)
- Convergence slows down when the undersampling rate α decreases (less measurements)
- No convergence above critical α (convergence time diverges)

What is value of the critical α ?

Measuring the gradient-sparsity of images

Measuring the sparsity of the images



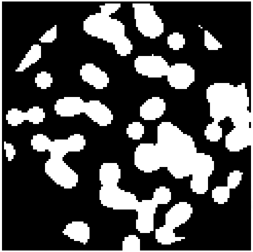


Compute the density of $1-$pixel wide boundaries

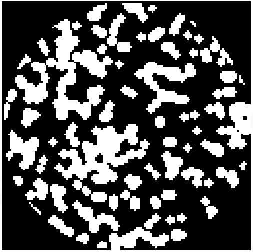
$$\rho = \frac{\#\text{ of pixels on boundaries}}{\#\text{ of searched-for pixels}}$$

Measuring the gradient-sparsity of images

$\rho = 0.05$

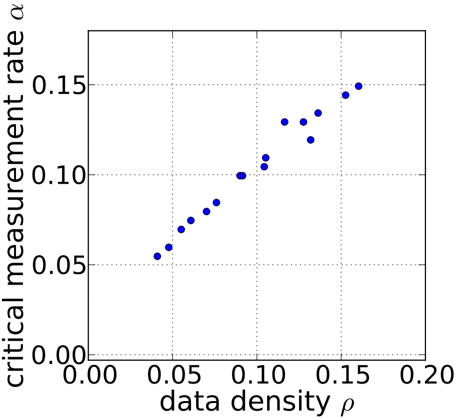


$\rho = 0.09$



$\rho = 0.17$

Phase diagram of noisefree recovery



$\alpha_c \simeq \rho$: almost optimal recovery ! !

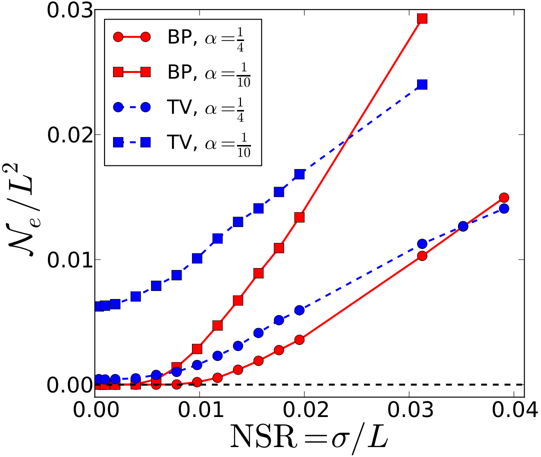
E. Gouillart et al., Inverse Problems, March 2013

Good robustness to noise (i.i.d. additive Gaussian noise)

Compare BP-tomo with TV regularization, for

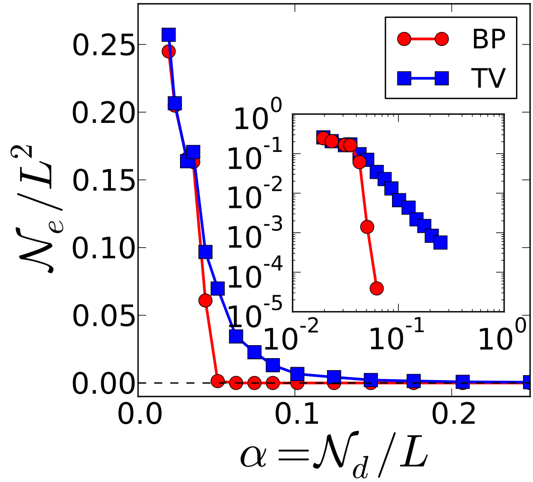
- different undersampling α
- different values of the noise σ

fraction of errors vs. noise amplitude



A sharp transition with the undersampling rate

fraction of errors vs. undersampling rate



For a given level of noise, choose α not too close to the transition (in applications : trade-off measurements / noise)

Mean-field approximation to reduce memory cost

$$\#\left\{\tilde{m}_{\mu\to i}, m_{i\to\mu}\right\} = 2Nn_\theta \propto N^{3/2}$$

"Mean-field" approximation

Replace

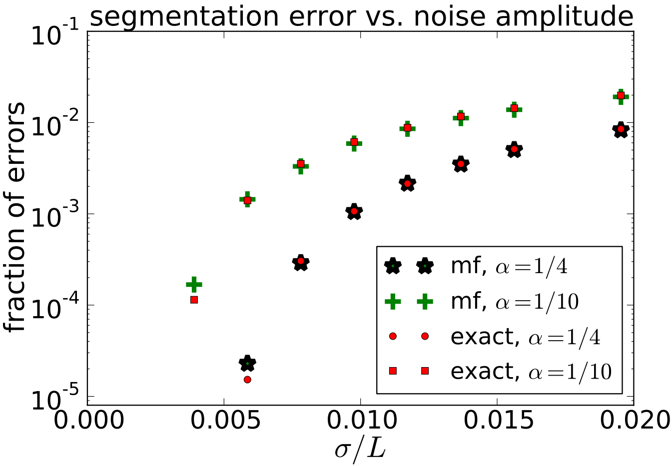
$$m_{i\to\mu} = \sum_{\nu\in i\neq\mu} \tilde{m}_{\nu\to i}$$

by

$$m_i = \frac{\#\{\nu \in i\} - 1}{\#\{\nu \in i\}} \sum_{\nu\in i} \tilde{m}_{\nu\to i}$$

$\tilde{m}_{\nu\to i}$ temporary variables computed when solving the Ising chains, but not kept in memory

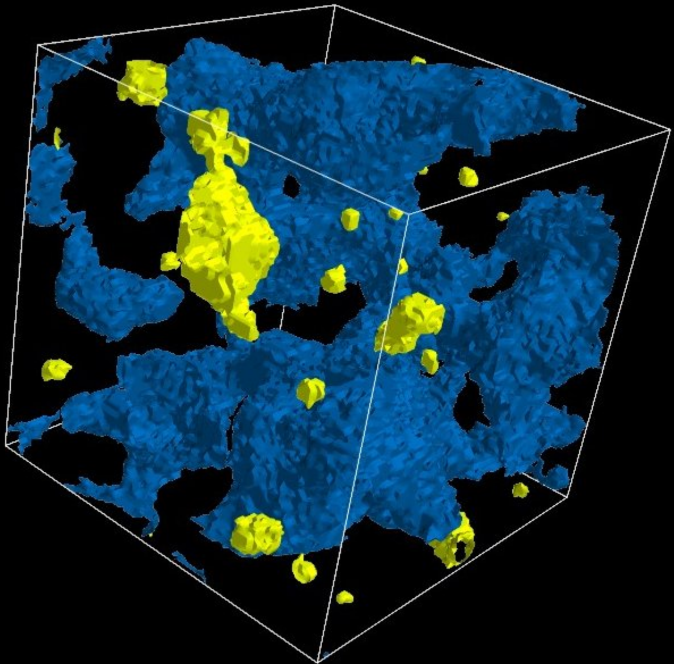
Mean-field approximation to reduce memory cost



- very low noise : slight degradation
- moderate noise : same result, or slight improvement

Conclusions

- **Belief-propagation** : an interesting algorithm to compute expectations for combinatorial problems described by a factor graph
  - Tomography : impressive performance on synthetic binary data
  - Sharp transition with undersampling rate α : stay far from critical α
- **Next steps**
  - Test with real data : prior model not perfect, forward model also not perfect... How robust is the algorithm ?
  - More colors : three, fours, ...phases



Bibliography I

- Robert Gallager, **Low-density parity-check codes**, Information Theory, IRE Transactions on **8** (1962), no. 1, 21–28.
- X. Jia, Y. Lou, R. Li, W.Y. Song, and S.B. Jiang, **Gpu-based fast cone beam ct reconstruction from undersampled and noisy projection data via total variation**, Medical physics **37** (2010), 1757.
- Kevin P Murphy, Yair Weiss, and Michael I Jordan, **Loopy belief propagation for approximate inference : An empirical study**, Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., 1999, pp. 467–475.

Bibliography II

- Judea Pearl, **Reverend bayes on inference engines : A distributed hierarchical approach**, Cognitive Systems Laboratory, School of Engineering and Applied Science, University of California, Los Angeles, 1982.
- E.Y. Sidky and X. Pan, **Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization**, Physics in medicine and biology **53** (2008), 4777.
- J. Tang, B.E. Nett, and G.H. Chen, **Performance comparison between total variation (tv)-based compressed sensing and statistical iterative reconstruction algorithms**, Physics in Medicine and Biology **54** (2009), 5781.