Bayesian Inference for Image Reconstruction

So Nakashima and Hila Barel

Supervisors: Isaac Perez Castillo, Tim Westwood

London Mathematical Laboratory

Content

- Introduction to Inverse Problems
- Bayesian Inference
- Expectation Propagation Algorithm
- Results
- Summary

What is Image Reconstruction?

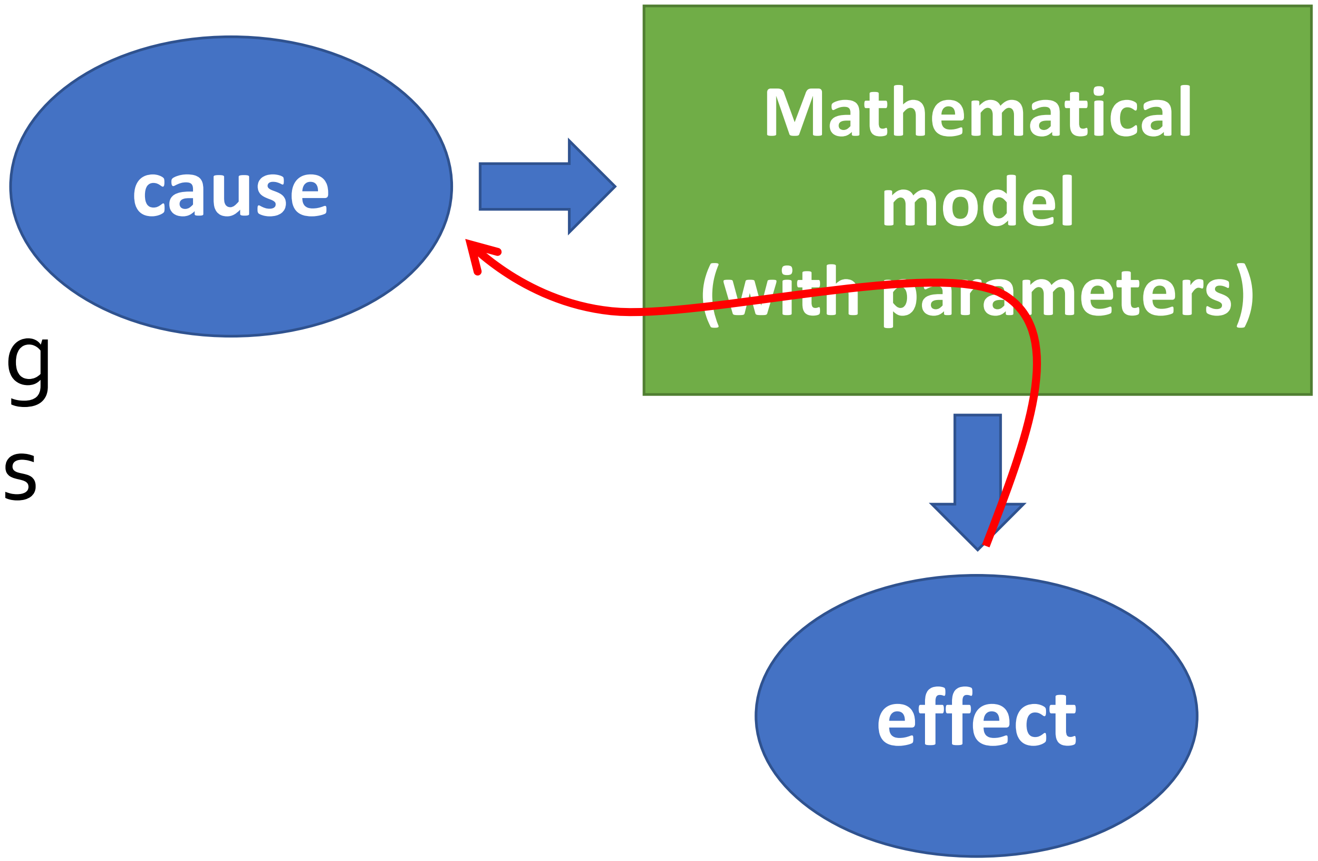
A type of an **inverse problem** where the challenge is to yield an estimate of a specific image from a set of **projections**.

Why is it used for?

- Computed Tomography (CT-scan)
- Positron Emission Mammography
- ...

Inverse Problems

Wikipedia:
"the process of calculating from a set of observations the causal factors that produced them"



Inverse Problems

- While the **forward** problem has a unique solution, the **inverse** problem does not. It can have multiple solutions (even infinite number).
- Because of this one needs to make explicit any available a **priori information** on the model variables.
- The better the prior information is, the **less** measurements we need.

Example: Tomographic Reconstruction

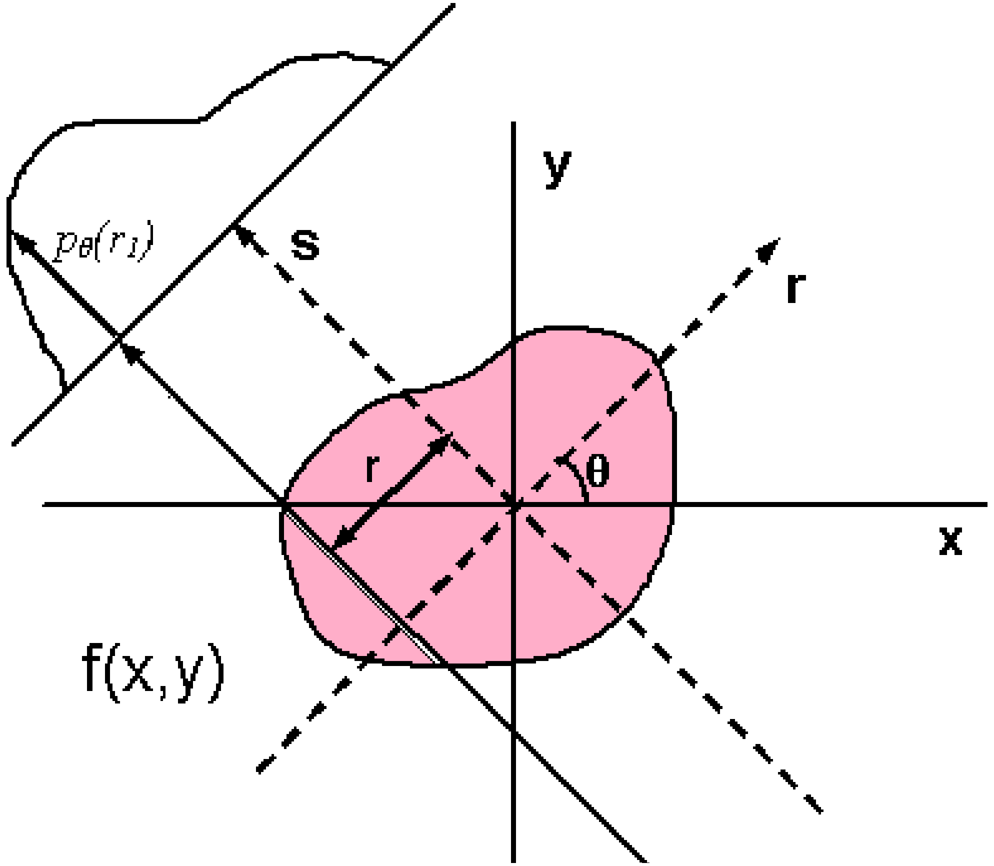
Tomography:
Imaging technique in diagnostic medicine.
It consists in illuminating an object (an internal organ) from different directions and detecting transmitted rays.

Reconstruction:
Finding the 2D cross-section of the object (unknown variables) from the set of projections (measurements).

Tomographic Reconstruction

The projection along the ray (r,θ):

$$p_\theta(r) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - r)\,dx\,dy$$



Where *f(x,y)* is the image intensity at position *(x,y)*

Discrete Tomographic Reconstruction

Formulation:

$$\boldsymbol{y} = F \cdot \boldsymbol{x} + \boldsymbol{w}$$

$$x_i - image\ pixels, \qquad i = 1, \dots N$$
$$y_j - projections, \qquad j = 1, \dots, M$$
$$F \ - projection\ matrix\ (M \times N)$$
$$w_j - measurements\ noise$$

Given **y**, we wish to estimate **x**

Discrete Tomographic Reconstruction

Formulation:

$$\boldsymbol{y} = F \cdot \boldsymbol{x} + \boldsymbol{w}$$

Typically M<N,

→ ill-posed problem.
→ Need additional prior information on **x**.

Bayesian Inference

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{x})}{p(\boldsymbol{y})} p_{\text{prior}}(\boldsymbol{x})$$

With better prior, we need less data

Example.

$\boldsymbol{x}$: value of pixels

$\boldsymbol{y}$: data\projections

$p_{\text{prior}}(\boldsymbol{x}) = \prod_i \frac{1}{Z} \, \mathrm{I}[x_i \in [-L, L]]$

Approximation methods

Estimator: $x_i^* = E_{p(x_i|\boldsymbol{y})}[x_i]$

$p(x_i \mid \boldsymbol{y}) = \int p(\boldsymbol{x} \mid \boldsymbol{y})p_{\text{prior}}(\boldsymbol{x})\text{d}\boldsymbol{x}_{\setminus \text{i}}$ is difficult to compute

Methods to approximate $p(x_i \mid \boldsymbol{y})$

- Belief Propagation (BP) ······ rely on Bethe approximation
- Markov Chain Monte Carlo (MCMC) ······ slow
- Expectation Propagation (EP)

Model for Image Reconstruction

q-value: $x_i = 0, 1, \ldots, q-1$

Measurement noise $\boldsymbol{w} = F\boldsymbol{x} - \boldsymbol{y} \sim N(0, I_M/\beta)$

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = \frac{1}{Z_1}\exp\left(-\frac{\beta}{2}(F\boldsymbol{x}-\boldsymbol{y})^t(F\boldsymbol{x}-\boldsymbol{y})\right)$$

$$p_{\mathrm{prior}}(\boldsymbol{x}) = \frac{1}{Z_2}\exp\left(-\frac{J}{2}\boldsymbol{x}^t L\boldsymbol{x}\right)\cdot\prod_i \psi_i(x_i)$$

Smoothness of image

$P(x_i = q) = \alpha_q$

EP algorithm

$$p_{\mathrm{prior}}(\boldsymbol{x}) = \frac{1}{Z_2} \exp\left(-\frac{J}{2} \boldsymbol{x}^t L \boldsymbol{x}\right) \cdot \prod_i \psi_i(x_i)$$

Approximation

$$q_{\mathrm{prior}}(\boldsymbol{x}) = \frac{1}{Z_2} \exp\left(-\frac{J}{2} \boldsymbol{x}^t L \boldsymbol{x}\right) \cdot \prod_i \phi_i(x_i \mid a_i, b_i)$$

$$\phi_i(x_i) \sim \mathrm{N}(a_i, b_i)$$

Then, everything is Gaussian

EP algorithm

Choose $q_{\text{prior}}(\boldsymbol{x})$ so that $q(\boldsymbol{x} \mid \boldsymbol{y})$ is *similar* to $p(\boldsymbol{x} \mid \boldsymbol{y})$

$q(\boldsymbol{x} \mid \boldsymbol{y}) \coloneqq p(\boldsymbol{y} \mid \boldsymbol{x})\, q_{\text{prior}}(\boldsymbol{x})$

Kullback–Leibler divergence

$$\mathrm{KL}(p||q) = \int p(x) \ln \frac{p(x)}{q(x)} \,\mathrm{d}x$$

But $\mathrm{KL}(p(\boldsymbol{x} \mid \boldsymbol{y})||q(\boldsymbol{x} \mid \boldsymbol{y}))$ is difficult to compute…

EP algorithm

$q^{(i)}(\boldsymbol{x} \mid \boldsymbol{y}) = p(\boldsymbol{y} \mid \boldsymbol{x})\psi_i(x_i)\prod_{j \neq i} \phi_j(x_j \mid a_j, b_j)$

Approximation of $p(\boldsymbol{x} \mid \boldsymbol{y})$

cf. $q(\boldsymbol{x} \mid \boldsymbol{y}) \coloneqq p(\boldsymbol{y} \mid \boldsymbol{x}) \prod_k \phi_k(x_k \mid a_i, b_i)$

EP algorithm

Until convergence:

For each $i = 1, 2, \ldots, N$

$\underset{a_i, b_i}{\text{Min.}} \ \mathrm{KL}\left(q^{(i)}(\boldsymbol{x} \mid \boldsymbol{y}) \| q(\boldsymbol{x} \mid \boldsymbol{y})\right)$

Moment matching

Minimize $\mathrm{KL}\left(q^{(i)}(\boldsymbol{x} \mid \boldsymbol{y}) \mid q(\boldsymbol{x} \mid \boldsymbol{y}) \right)$

$\Leftrightarrow$ find q s.t.

$$E_{q^{(i)}}[x_i] = E_q[x_i]$$
$$E_{q^{(i)}}\left[x_i^2\right] = E_q\left[x_i^2\right]$$

Computable

Explicit formula in a_i, b_i

Summary of EP

Initialize a_i, b_i for $i = 1, 2, \ldots, N$

Until convergence:

for each $i = 1, 2, \ldots, N$

Moment Matching between $q^{(i)}(\boldsymbol{x} \mid \boldsymbol{y}),\ q(\boldsymbol{x} \mid \boldsymbol{y})$

Output $E_{q(\boldsymbol{x}|\boldsymbol{y})}[x_i]$ for $i = 1, 2, \ldots, N$

$$q(\boldsymbol{x} \mid \boldsymbol{y}) \coloneqq p(\boldsymbol{y} \mid \boldsymbol{x}) \prod_k \phi_k(x_k \mid a_i, b_i)$$
$$q^{(i)}(\boldsymbol{x} \mid \boldsymbol{y}) = p(\boldsymbol{y} \mid \boldsymbol{x}) \psi_i(x_i) \prod_{j \neq i} \phi_j(x_j \mid a_j, b_j)$$

Update of model parameters

Parameters in the model:

β: the inverse variance of measurement noise

J: smoothness of image

$\alpha_q : P(x_i = q)$

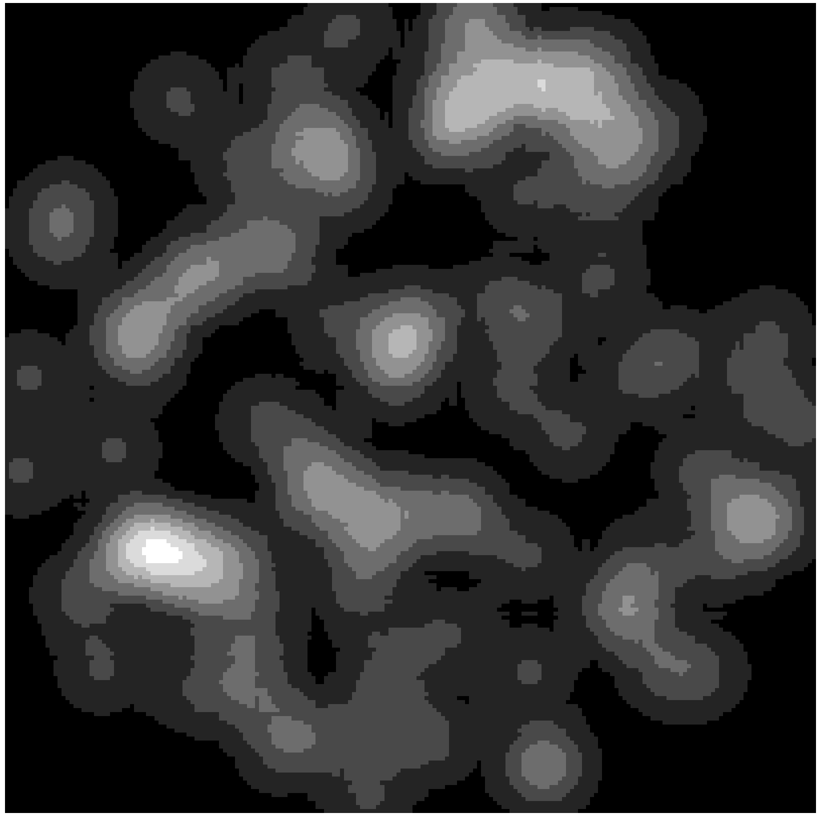
But these values are unknown in general.

We can estimate these parameters by EM-type method.

Results

Simulation:

- Simulated phantom images
- Image size ~ 10x10
  (due to computation limitation)
- Q = 4 levels
- Number of cluster in the image,
  ~25% of image pixels

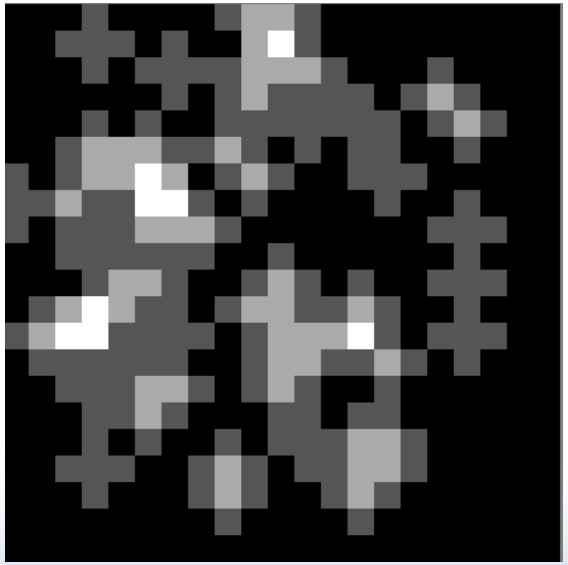


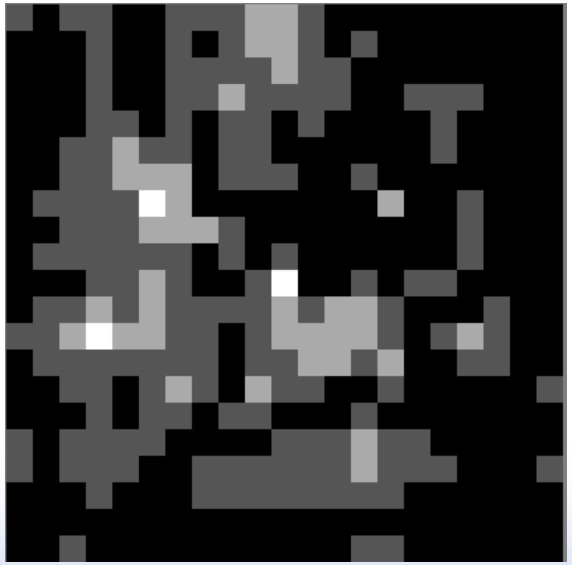
Results – example images

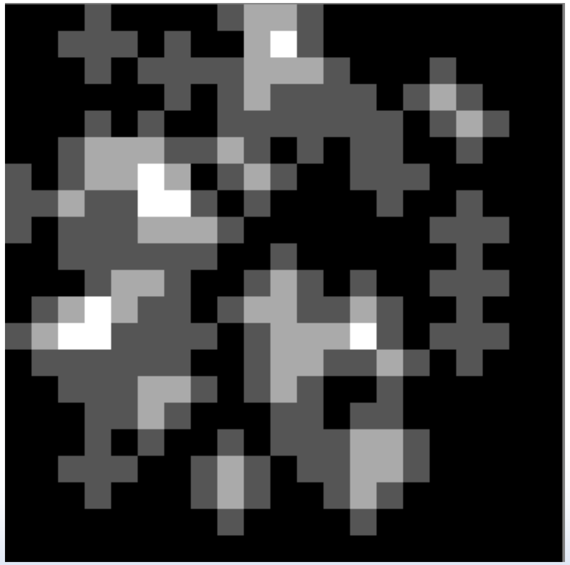
Left: original image

Middle: reconstructed image with $\alpha := \frac{M}{N} = 0.5$

Right: perfect reconstructed image with $\alpha = 0.7$

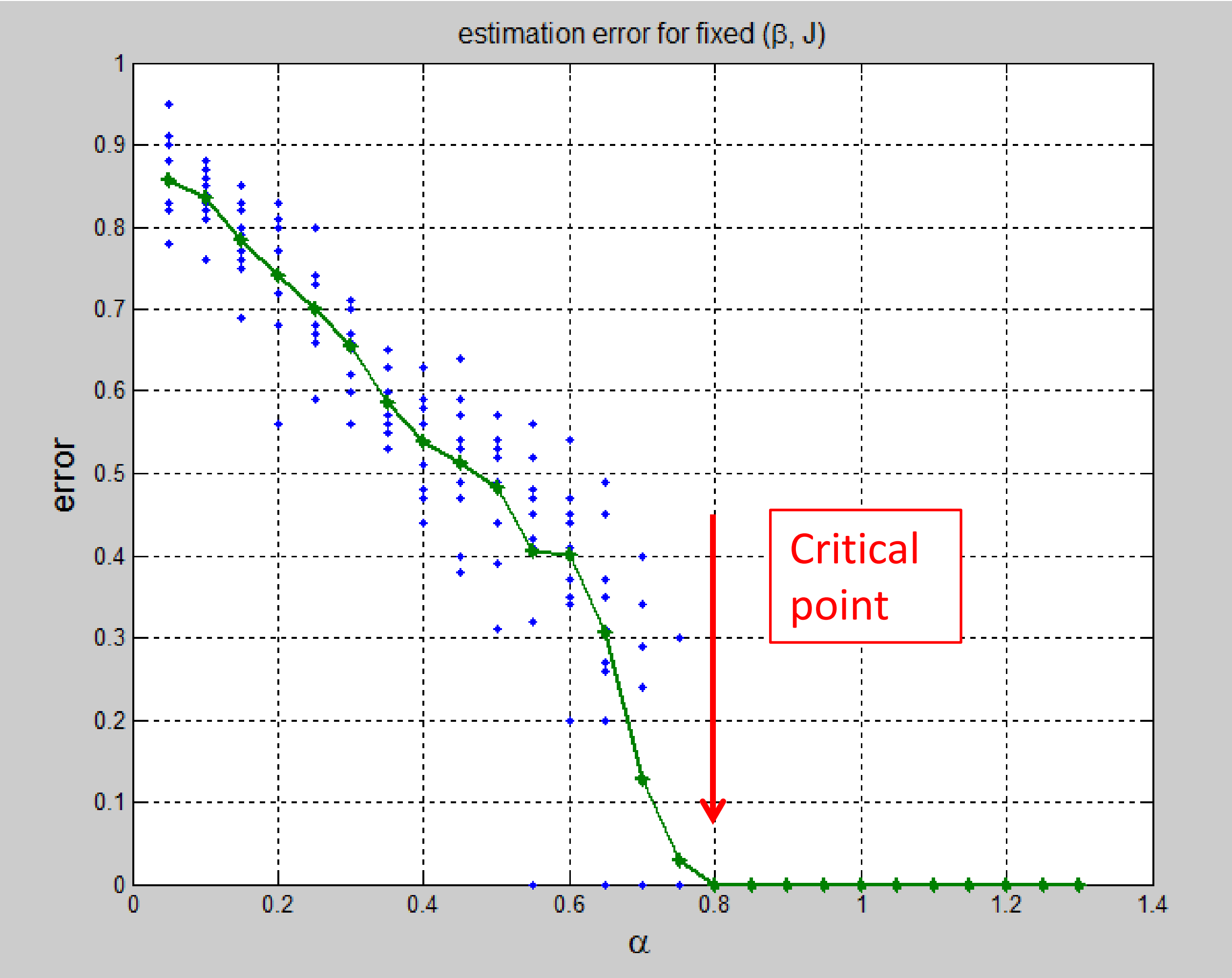






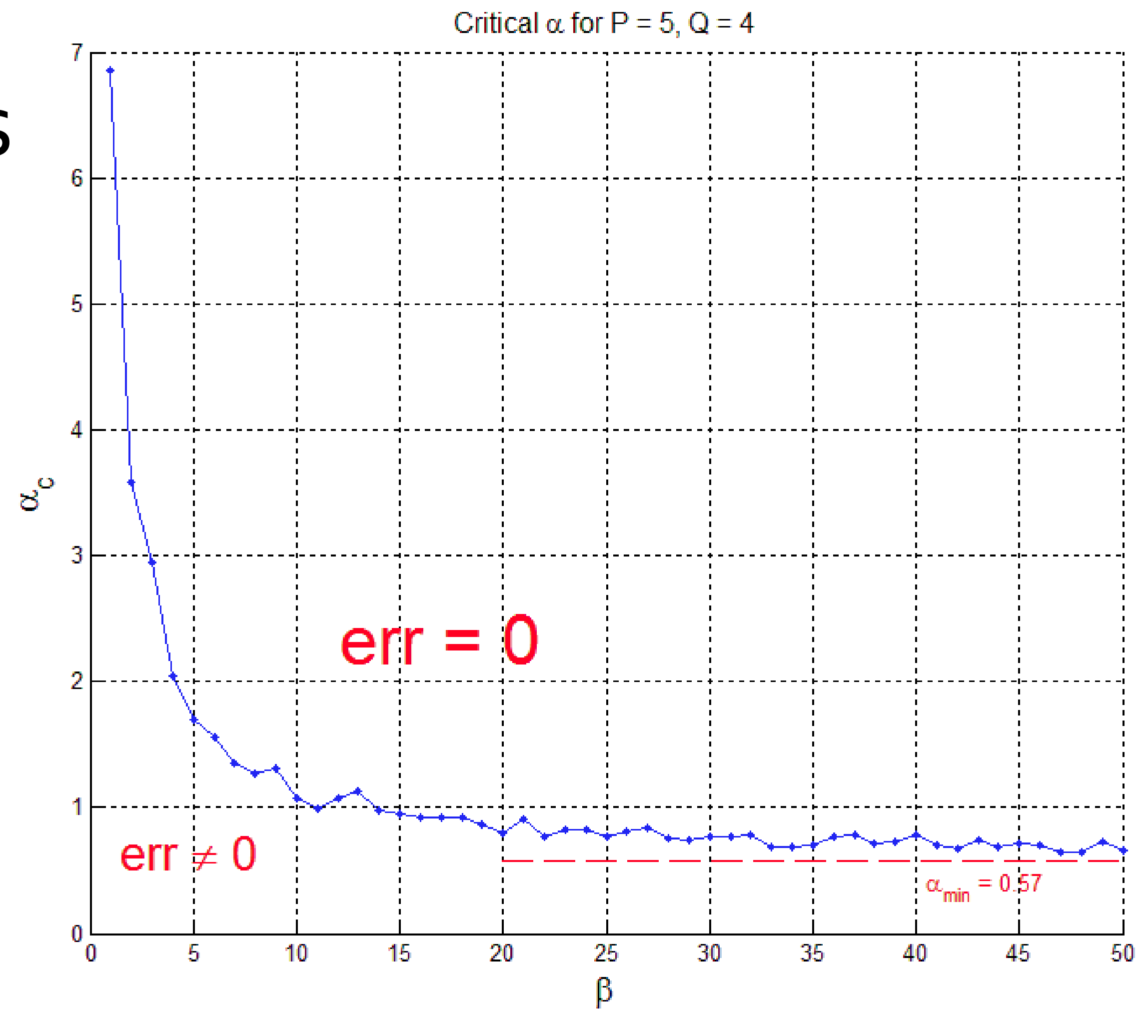
Results – how much information we need?

- Given (β, J) we can look at the estimation error as a function of number of measurements.
- The critical point is defined as the point that above which perfect reconstruction occurs.



Results – critical curve

- Critical curve characterizes the performance of the EP algorithm.
- The minimal information needed for perfect reconstruction can be extracted.



Summary

- Perfect reconstruction is achievable even in the presence of noise and $M < N$.

- Minimal information needed (α_c) changes with SNR level.

References

- Thomas Minka (August 2–5, 2001). "Expectation Propagation for Approximate Bayesian Inference". In Jack S. Breese, Daphne Koller. *UAI '01: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*. University of Washington, Seattle, Washington, USA.
- Muntoni, Anna Paola (2017), Statistical mechanics approaches to optimization and inference. PhD thesis.
- Wikipedia, Inverse Problems. https://en.wikipedia.org/wiki/Inverse_problem