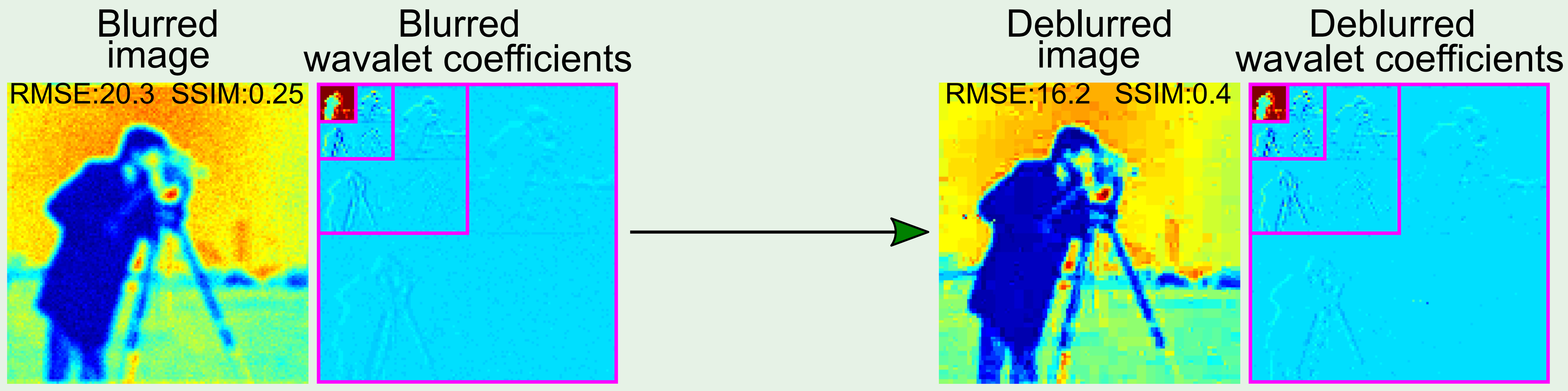
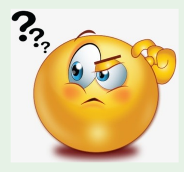
Unsupervised Expectation Propagation Method for Large-Scale Sparse Linear Inverse Problems

Dan Yao (dy2008@hw.ac.uk), Stephen McLaughlin, Yoann Altmann
School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK

Problem: Uncertainty Quantification in Solving Large Scale Sparse Linear Inverse Problems




But, how far the estimate might be from the 'true value'?


Uncertainty Quantification !

Goal

large scale observation $\rightarrow$ point estimate + uncertainty quantification

Challenges

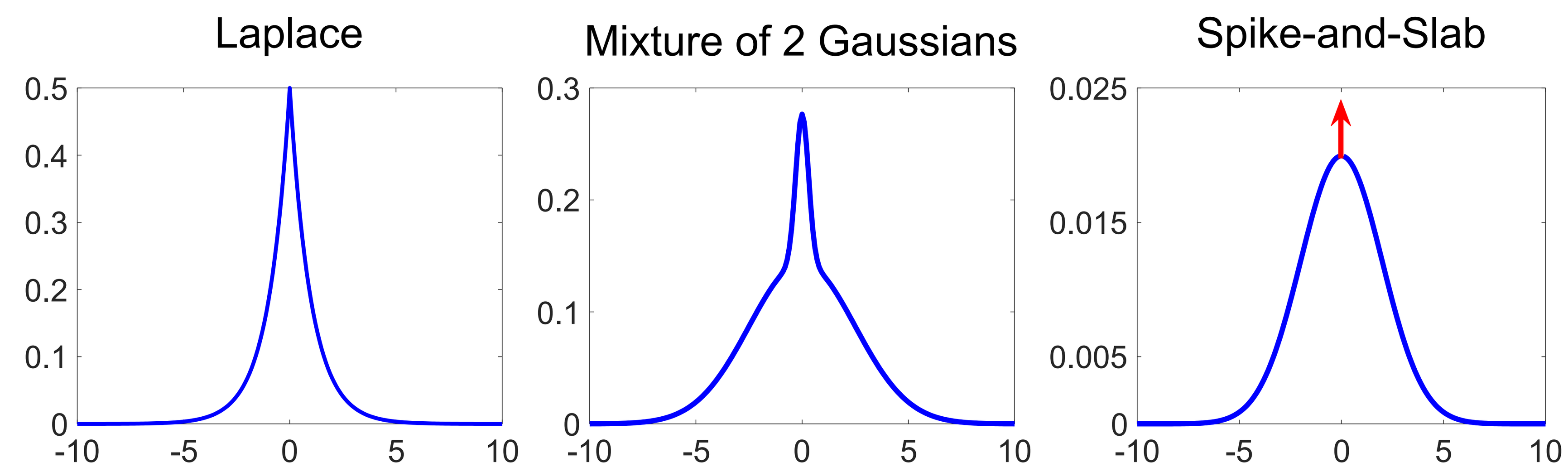
- **efficient computation** in solving large scale linear inverse problems (e.g. imaging problems)
- **optimal sparse solution** to large scale sparse linear inverse problems (e.g. in wavelet domain)
- **uncertainty quantification** for the point estimate
- **unsupervised hyperparameter estimation** when using Bayesian methods with sparsity enforcing priors

Proposed Solution: Expectation Propagation (EP) for Unsupervised Approximate Bayesian Inference

Bayesian model

likelihood: $f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})$

sparsity-enforcing priors: $f_x(\boldsymbol{x}|\boldsymbol{\theta})$

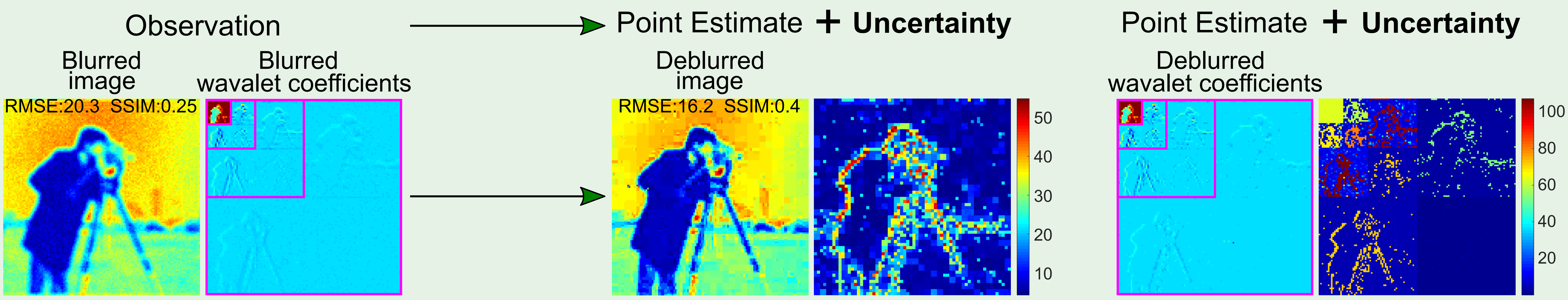


- posterior: $$p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta}) = \frac{f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})f_x(\boldsymbol{x}|\boldsymbol{\theta})}{\int f_{y|x}(\boldsymbol{y}|\mathbf{H}\boldsymbol{x})f_x(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}}$$
- MMSE (point estimate): $\mathbb{E}_p[\boldsymbol{x}] = \int \boldsymbol{x}p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}$
- covariance (uncertainty): $$\mathrm{Cov}_p(\boldsymbol{x}) = \int (\boldsymbol{x}-\mathbb{E}[\boldsymbol{x}])(\boldsymbol{x}-\mathbb{E}[\boldsymbol{x}])^T p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}$$

EP for approximate Bayesian inference

$$q(\boldsymbol{x}|\boldsymbol{\theta}) \underset{\text{minimization}}{\overset{\text{Kullback-Leibler divergence}}{\approx}} p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta})$$

- approximate posterior: $q(\boldsymbol{x}|\boldsymbol{\theta}) \approx p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta})$
- approximate MMSE: $\mathbb{E}_q[\boldsymbol{x}] \approx \mathbb{E}_p[\boldsymbol{x}]$
- approximate covariance: $\mathrm{Cov}_q(\boldsymbol{x}) \approx \mathrm{Cov}_p(\boldsymbol{x})$
- unsupervised hyperparamter estimation: $$\hat{\boldsymbol{\theta}} = \underset{\theta}{\mathrm{argmax}}\ \mathbb{E}_{q(\boldsymbol{x}|\boldsymbol{\theta}^{(t-1)})}[\log f(\boldsymbol{y}|\boldsymbol{\theta})]$$



References

[1] Dan Yao, Stephen McLaughlin, and Yoann Altmann. "Fast Scalable Image Restoration using Total Variation Priors and Expectation Propagation," IEEE Transactions on Image Processing, 2022, doi: 10.1109/TIP.2022.3202092.
[2] José Miguel Hernández-Lobato, Daniel Hernández-Lobato, and Alberto Suárez. "Expectation propagation in linear regression models with spike-and-slab priors," Machine Learning 99.3 (2015), pp. 437–487.
[3] Thomas P Minka. "Expectation propagation for approximate Bayesian inference," Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence. 2001, pp. 362–369.