Adaptive Sampling Optimization for Magnetic Resonance Imaging by Bayesian Experimental Design

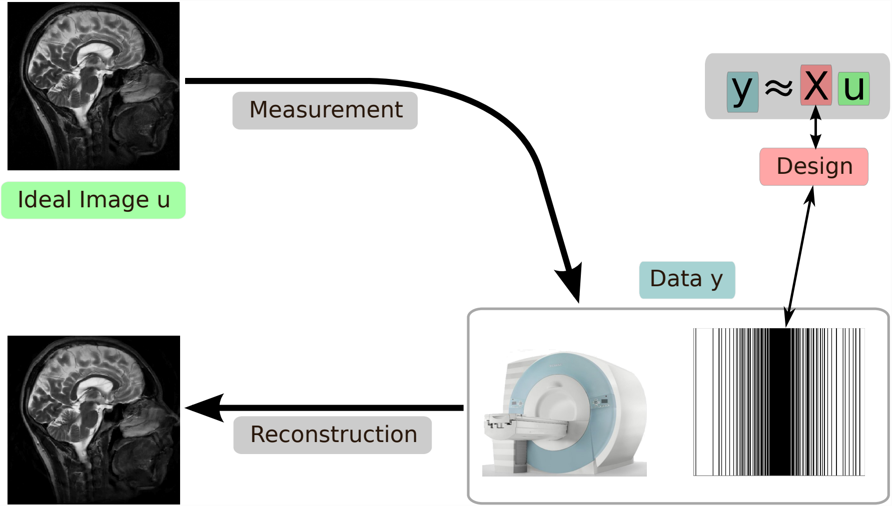
Matthias Seeger

joint work with Hannes Nickisch,
Rolf Pohmann, Bernhard Schölkopf

Laboratory for Probabilistic Machine Learning
Ecole Polytechnique Fédérale de Lausanne
http://lapmal.epfl.ch/

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Image Reconstruction

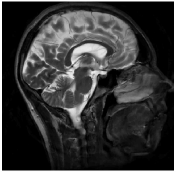


Sampling Optimization

scan time $\propto$ # phase encodes

$\mathbf{y} \approx \mathbf{X} \mathbf{u}$

$\boldsymbol{X} \leftarrow ?$

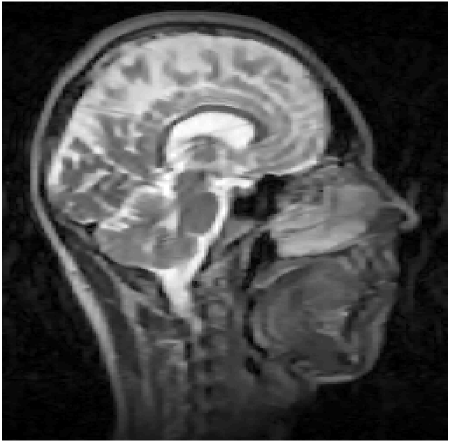


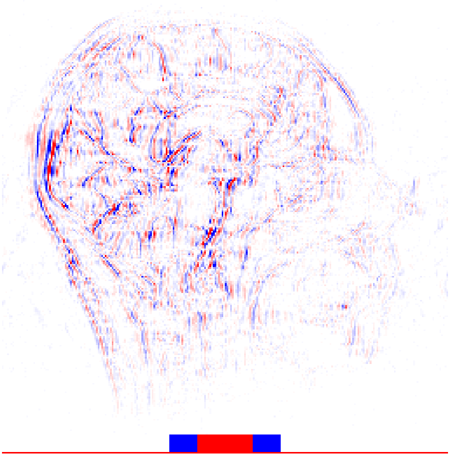
Reconstruction

Why Sampling Optimization?

MAP reconstruction, 1/4 Nyquist (64 encodes) $L_2 = 7.38$

- Low pass: Dense sampling of low frequencies



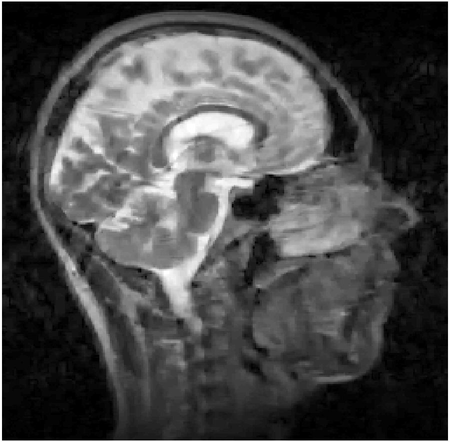


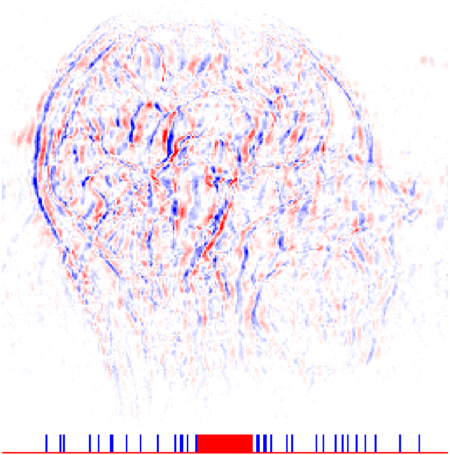
Why Sampling Optimization?

MAP reconstruction, 1/4 Nyquist (64 encodes) $L_2 = 8.12(\pm 0.53)$

- Random (variable density)

Lustig, Donoho, Pauli: MRM 58(6)



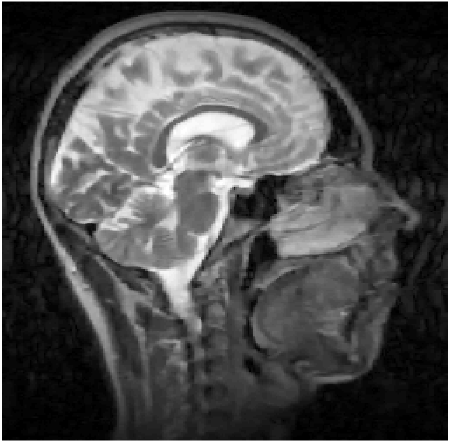


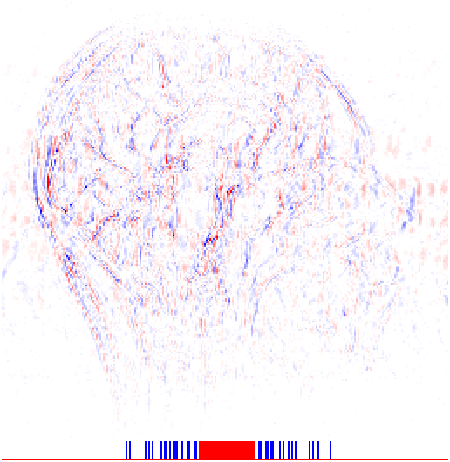
Why Sampling Optimization?

MAP reconstruction, 1/4 Nyquist (64 encodes) $L_2 = 5.66$

- Bayesian optimized

Seeger *et.al.*, MRM 63(1)

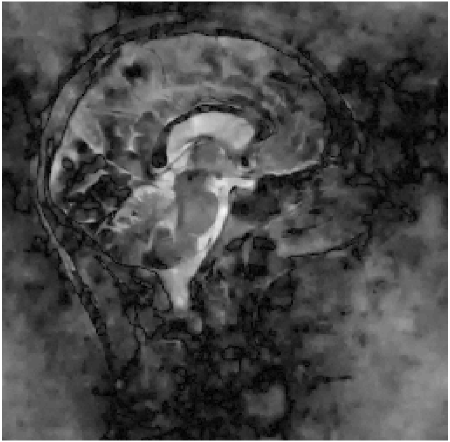


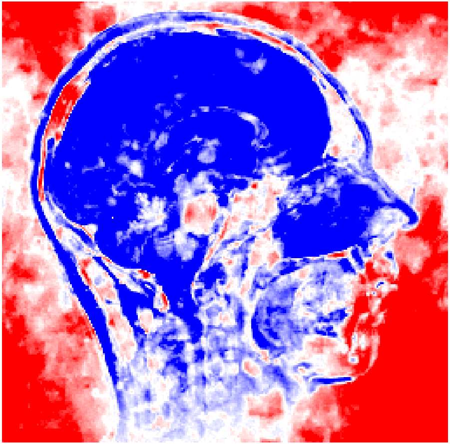


Why Sampling Optimization?

MAP reconstruction, 1/4 Nyquist (64 encodes) $L_2 = 46.48(\pm 8.01)$

- Uniformly random Fourier coefficients





Reconstruction is Ill Posed

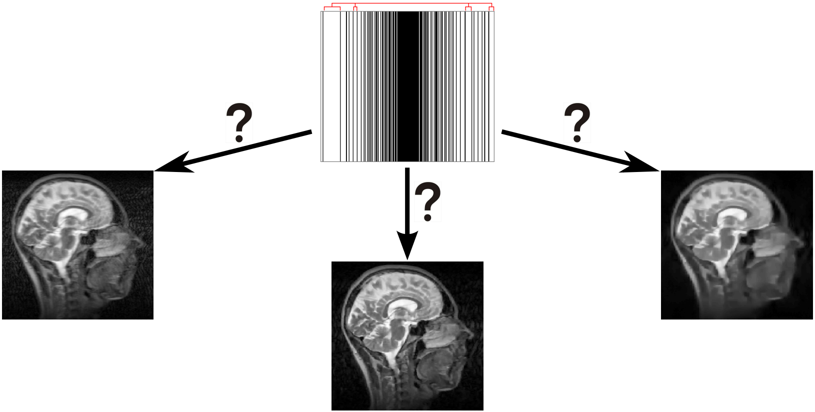


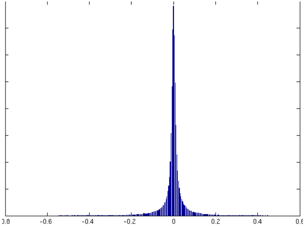
Image Statistics

Whatever images are . . .

they are not Gaussian!

- Image gradient super-Gaussian (“sparse”)



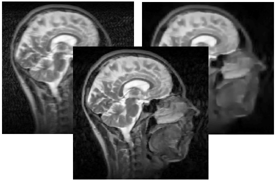


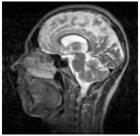
Use sparsity prior distribution $P(\boldsymbol{u})$

Posterior Distribution

- Likelihood $P(\boldsymbol{y}|\boldsymbol{u})$: Data fit
- Prior $P(\boldsymbol{u})$: Signal properties
- Posterior distribution $P(\boldsymbol{u}|\boldsymbol{y})$: Consistent information summary

$$P(\boldsymbol{y}|\boldsymbol{u})$$

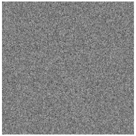


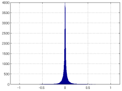


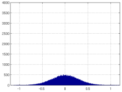
Posterior Distribution

- Likelihood $P(\boldsymbol{y}|\boldsymbol{u})$: Data fit
- Prior $P(\boldsymbol{u})$: Signal properties
- Posterior distribution $P(\boldsymbol{u}|\boldsymbol{y})$: Consistent information summary

$$P(\boldsymbol{y}|\boldsymbol{u}) \times P(\boldsymbol{u})$$



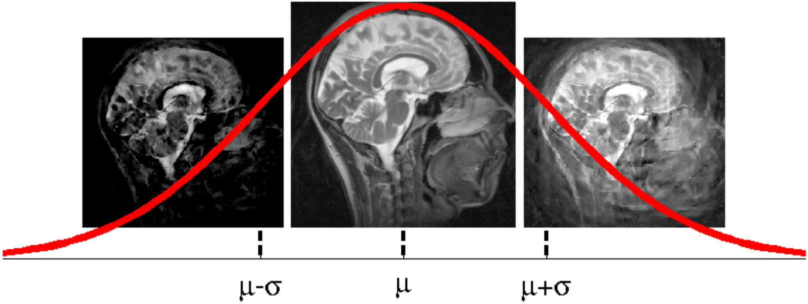




Posterior Distribution

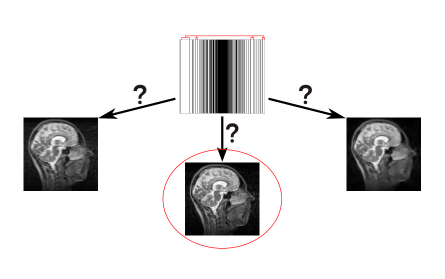
- Likelihood $P(\boldsymbol{y}|\boldsymbol{u})$: Data fit
- Prior $P(\boldsymbol{u})$: Signal properties
- Posterior distribution $P(\boldsymbol{u}|\boldsymbol{y})$: Consistent information summary

$$P(\boldsymbol{u}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{u}) \times P(\boldsymbol{u})}{P(\boldsymbol{y})}$$



Estimation and Bayesian Decision Making

Penalized Estimation

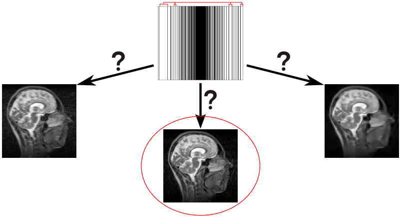


$$\hat{\boldsymbol{u}} = \operatorname{argmax} P(\boldsymbol{y}|\boldsymbol{u})P(\boldsymbol{u})$$

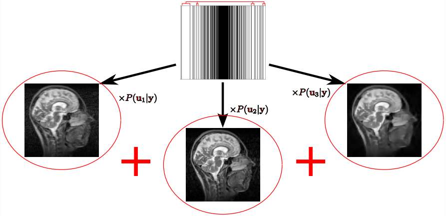
- Single best guess $\hat{\boldsymbol{u}}$
- Massive recent progress

Estimation and Bayesian Decision Making

Penalized Estimation



Bayesian Inference



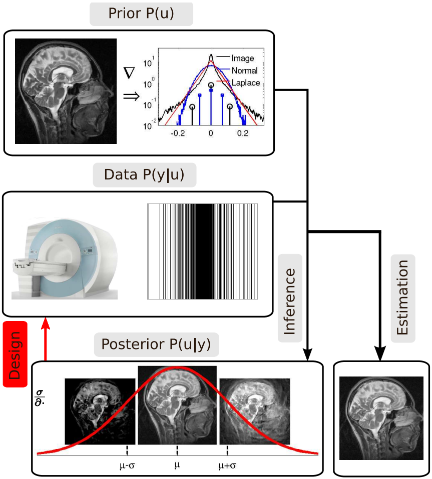
$$\hat{\boldsymbol{u}} = \operatorname{argmax} P(\boldsymbol{y}|\boldsymbol{u})P(\boldsymbol{u})$$

$$P(\boldsymbol{u}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{u})P(\boldsymbol{u})}{P(\boldsymbol{y})}$$

- Single best guess $\hat{\boldsymbol{u}}$
- Massive recent progress

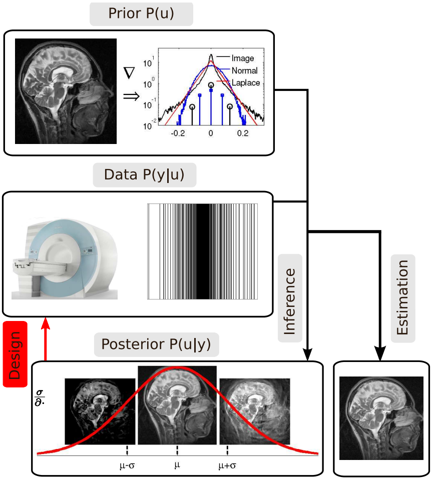
- Quantify your uncertainty
- Optimal decision making

Bayesian Experimental Design

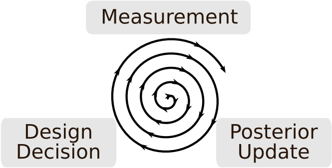


- Posterior: Uncertainty in reconstruction
- Experimental design: Find poorly determined directions
- Sequential search with interjacent partial measurements

Bayesian Experimental Design



- Posterior: Uncertainty in reconstruction
- Experimental design: Find poorly determined directions
- Sequential search with interjacent partial measurements



Main Message

- Compressive sampling for anatomical MRI?
  Needs design optimization.
  Random sampling plus ℓ_1 magic suboptimal.
- Design optimization:
  Decision making (about $\boldsymbol{X}$) under uncertainty (about $\boldsymbol{u}$).
  Needs Bayesian inference beyond MAP estimation.
- Bayesian inference? You must be kidding!
  - MCMC: Undirected random walks, not based on optimization
  - Many orders of magnitude slower than estimation (gap increasing)

Main Message

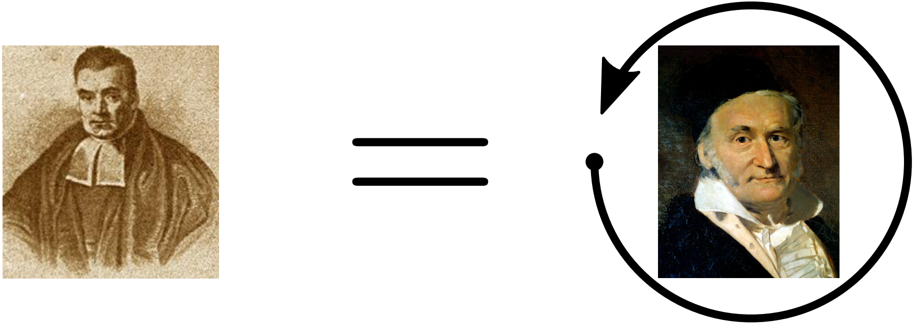
- Compressive sampling for anatomical MRI?
  Needs design optimization.
  Random sampling plus ℓ_1 magic suboptimal.
- Design optimization:
  Decision making (about $\boldsymbol{X}$) under uncertainty (about $\boldsymbol{u}$).
  Needs Bayesian inference beyond MAP estimation.
- Bayesian inference? You must be kidding!
  - MCMC: Undirected random walks, not based on optimization
  - Many orders of magnitude slower than estimation (gap increasing)

Main Message

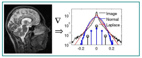
- Compressive sampling for anatomical MRI?
  Needs design optimization.
  Random sampling plus ℓ_1 magic suboptimal.
- Design optimization:
  Decision making (about $\boldsymbol{X}$) under uncertainty (about $\boldsymbol{u}$).
  Needs Bayesian inference beyond MAP estimation.
- Bayesian inference? You must be kidding!
  - MCMC: Undirected random walks, not based on optimization
  - Many orders of magnitude slower than estimation (gap increasing)

Main Message

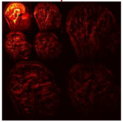
- Variational Bayesian inference:
  - Convex optimization
  - Driven by MAP (least squares) estimation



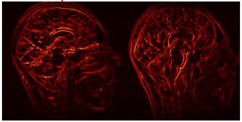
Sparse Linear Image Model



$$P(\mathbf{u}) \propto \prod_{i=1}^{q} t_i(s_i) = e^{-\tau_w \|\mathbf{B}_w \mathbf{u}\|_1} \times e^{-\tau_{tv} \|\mathbf{B}_{tv} \mathbf{u}\|_1}, \qquad \mathbf{s} = \mathbf{B}\mathbf{u}$$

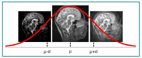


wavelet



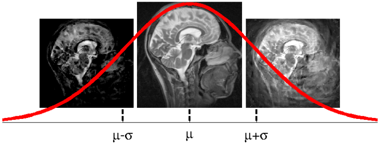
gradient

$$P(\mathbf{y}|\mathbf{u}) = \mathcal{N}(\mathbf{y}|\mathbf{X}\mathbf{u}, \sigma^2 \mathbf{I})$$



$$P(\mathbf{u}|\mathbf{y}) \propto P(\mathbf{u}) P(\mathbf{y}|\mathbf{u})$$

Variational Bayesian Inference

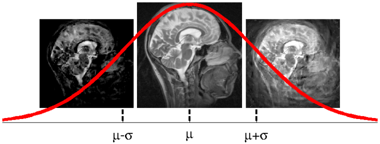


$$P(\boldsymbol{u}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{u}) \times P(\boldsymbol{u})}{P(\boldsymbol{y})}$$

Variational Inference Approximation

- Write intractable integration as (intractable) optimization
- Relax to tractable optimization problem

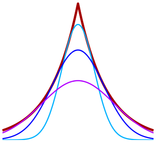
Variational Bayesian Inference



$$P(\boldsymbol{u}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{u}) \times P(\boldsymbol{u})}{P(\boldsymbol{y})}$$

- Variational Relaxation: Bound the master function

$$\underbrace{-\log P(\boldsymbol{y}) = -\log \int P(\boldsymbol{u},\boldsymbol{y})\, d\boldsymbol{u}}_{\text{Moment generating function}} \le \min_{\gamma} \min_{\boldsymbol{u}} \phi(\boldsymbol{u},\gamma)$$

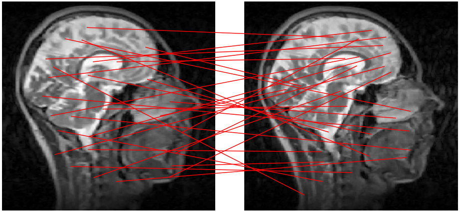


  - Approximate posterior $P(\boldsymbol{u}|\boldsymbol{y})$ by Gaussian
  - Integration $\Rightarrow$ Convex optimization

Decoupling by Concavity

$$-\log \int P(\boldsymbol{u}, \boldsymbol{y})\, d\boldsymbol{u} \le \min_{\gamma} \min_{\boldsymbol{u}} \phi(\boldsymbol{u}, \gamma)$$

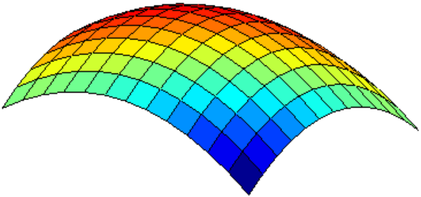
- Dependencies in posterior $P(\boldsymbol{u}|\boldsymbol{y})$
  $\Rightarrow$ Difficult couplings in criterion ϕ
- Critically coupled part is concave
- Upper bound by tangent plane



Decoupling by Concavity

$$-\log \int P(\boldsymbol{u}, \boldsymbol{y})\, d\boldsymbol{u} \leq \min_{\gamma} \min_{\boldsymbol{u}} \phi(\boldsymbol{u}, \gamma)$$

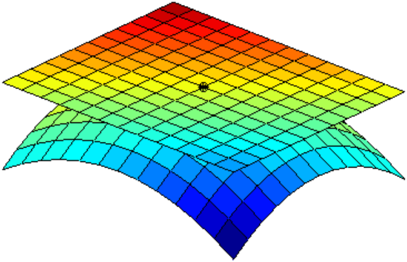
- Dependencies in posterior $P(\boldsymbol{u}|\boldsymbol{y})$
  $\Rightarrow$ Difficult couplings in criterion ϕ
- Critically coupled part is concave
- Upper bound by tangent plane



Decoupling by Concavity

$$-\log \int P(\boldsymbol{u}, \boldsymbol{y})\, d\boldsymbol{u} \leq \min_{\gamma} \min_{\boldsymbol{u}} \phi(\boldsymbol{u}, \gamma) = \min_{\boldsymbol{z}} \underbrace{\min_{\gamma} \min_{\boldsymbol{u}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)}_{\text{Decoupled problem}}$$

- Dependencies in posterior $P(\boldsymbol{u}|\boldsymbol{y})$
  $\Rightarrow$ Difficult couplings in criterion ϕ
- Critically coupled part is concave
- Upper bound by tangent plane



Double Loop Algorithm

$$\min_{\boldsymbol{u}} (\min_{\gamma} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma))$$

Double loop algorithm

- Inner loop optimization: $\min_{\gamma} \min_{\boldsymbol{u}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$
- Outer loop update: $\min_{\boldsymbol{z}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$

Double Loop Algorithm

$$\min_{\boldsymbol{u}}(\min_{\gamma} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)) = \min_{\boldsymbol{u}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{u}\|^2 + \mathcal{R}_{\boldsymbol{z}}(\boldsymbol{u})$$

Double loop algorithm

- Inner loop optimization: $\min_{\gamma} \min_{\boldsymbol{u}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$
  Penalized Least Squares
- Outer loop update: $\min_{\boldsymbol{z}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$

Double Loop Algorithm

$$\min_{\boldsymbol{z}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$$

Double loop algorithm

- Inner loop optimization: $\min_{\gamma} \min_{\boldsymbol{u}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$
  Penalized Least Squares
- Outer loop update: $\min_{\boldsymbol{z}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$

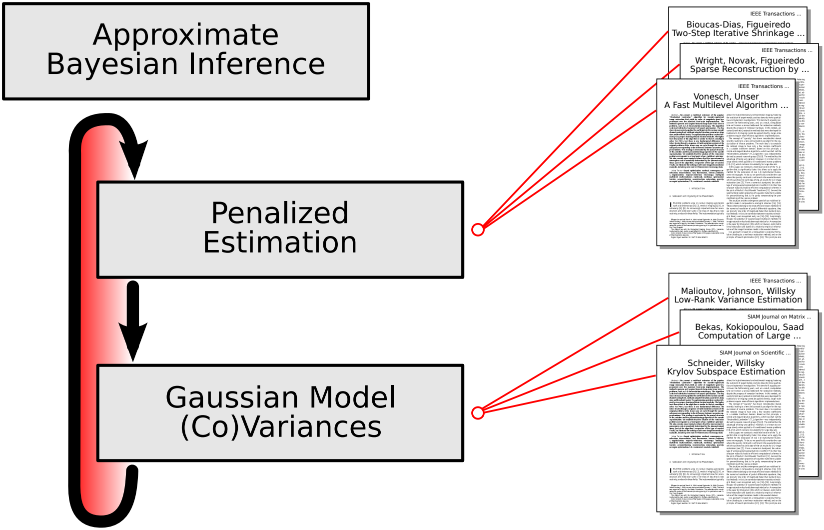
Double Loop Algorithm

$$\text{Tangent plane} : \boldsymbol{z} \leftarrow \nabla_{\gamma^{-1}} \phi_{\cap}(\gamma^{-1}) = \text{diag}(\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T)$$
$$\boldsymbol{A} = \sigma^{-2}\boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{B}^T(\text{diag}\,\gamma)^{-1}\boldsymbol{B}$$

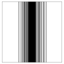
Double loop algorithm

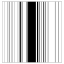
- Inner loop optimization: $\min_{\gamma} \min_{\boldsymbol{u}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$
  Penalized Least Squares
- Outer loop update: $\min_{\boldsymbol{z}} \phi_{\boldsymbol{z}}(\boldsymbol{u}, \gamma)$
  Gaussian (Co)Variances

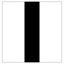
Standing on Shoulders

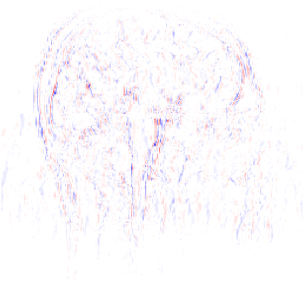


Optimizing Cartesian MRI

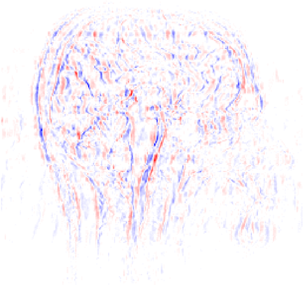
Bayes Optim. 

VD Random 

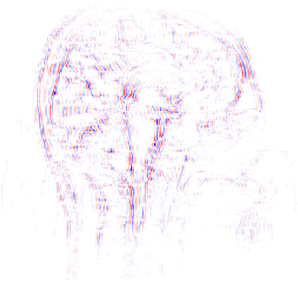
Low Pass 



Seeger *et.al.*, MRM 63(1), 2010

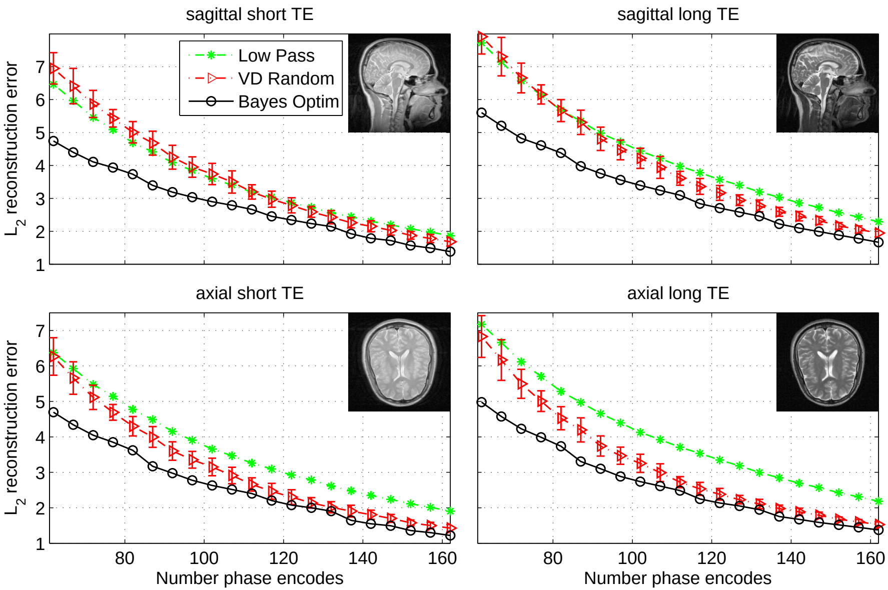


Lustig, Donoho, Pauli, MRM 58(6), 2007



Common MRI practice

Experimental Results: Test Set Errors



Bayesian Technology for Imaging

MRI applications other than sampling optimization

- Autocalibrating parallel MRI (plus k-space optimization) by robust Bayesian blind deconvolution
- Robust joint estimation / calibration: B_0 field map, relaxation, motion, ...
- Parallel transmit: Optimizing spokes design
- Dynamic / 3D MRI: Graphical models, belief propagation to factorize computations

Bayesian Technology for Imaging

MRI applications other than sampling optimization

- Autocalibrating parallel MRI (plus k-space optimization) by robust Bayesian blind deconvolution
- Robust joint estimation / calibration: B_0 field map, relaxation, motion, . . .
- Parallel transmit: Optimizing spokes design
- Dynamic / 3D MRI: Graphical models, belief propagation to factorize computations

Bayesian Technology for Imaging

MRI applications other than sampling optimization

- Autocalibrating parallel MRI (plus *k*-space optimization) by robust Bayesian blind deconvolution
- Robust joint estimation / calibration: B_0 field map, relaxation, motion, . . .
- Parallel transmit: Optimizing spokes design
- Dynamic / 3D MRI: Graphical models, belief propagation to factorize computations

Bayesian Technology for Imaging

MRI applications other than sampling optimization

- Autocalibrating parallel MRI (plus *k*-space optimization) by robust Bayesian blind deconvolution
- Robust joint estimation / calibration: B_0 field map, relaxation, motion, . . .
- Parallel transmit: Optimizing spokes design
- Dynamic / 3D MRI: Graphical models, belief propagation to factorize computations

Bayesian Technology for Imaging

MRI applications other than sampling optimization

- Autocalibrating parallel MRI (plus k-space optimization) by robust Bayesian blind deconvolution
- Robust joint estimation / calibration: B_0 field map, relaxation, motion, . . .
- Parallel transmit: Optimizing spokes design
- Dynamic / 3D MRI: Graphical models, belief propagation to factorize computations

Generic framework beyond magnetic resonance imaging

- You can do penalized estimation efficiently? You can do variational Bayesian inference!

Conclusions

- Modern nonlinear image reconstruction:
  Better images through robust low-level prior knowledge
- *k*-space optimization makes the difference:
  - Specific to reconstruction method
  - Specific to signal class (natural/MR images)
- Nonlinear Bayesian sampling optimization:
  General, goal-directed alternative to randomized trial-and-error
- Driven by scalable variational inference
  - Decoupling speeds up optimization dramatically
  - Penalized estimation technology inside

Conclusions

- Modern nonlinear image reconstruction:
  Better images through robust low-level prior knowledge
- k-space optimization makes the difference:
  - Specific to reconstruction method
  - Specific to signal class (natural/MR images)
- Nonlinear Bayesian sampling optimization:
  General, goal-directed alternative to randomized trial-and-error
- Driven by scalable variational inference
  - Decoupling speeds up optimization dramatically
  - Penalized estimation technology inside

Conclusions

- Modern nonlinear image reconstruction:
  Better images through robust low-level prior knowledge
- k-space optimization makes the difference:
  - Specific to reconstruction method
  - Specific to signal class (natural/MR images)
- Nonlinear Bayesian sampling optimization:
  General, goal-directed alternative to randomized trial-and-error
- Driven by scalable variational inference
  - Decoupling speeds up optimization dramatically
  - Penalized estimation technology inside

Collaborators

- Hannes Nickisch (now Philips Research, Hamburg)
- Rolf Pohmann, Bernhard Schölkopf (MPI Tübingen)

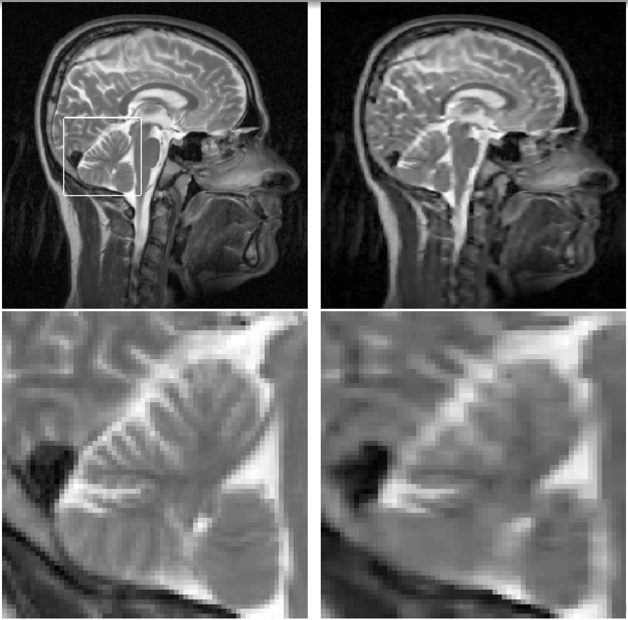
Matlab Code

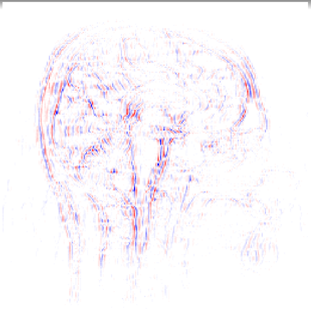
http://www.mloss.org/software/view/269/

More Information

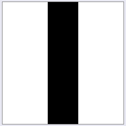
http://lapmal.epfl.ch/proj/ed_mri/index.shtml

Experimental Results: 1/4 Nyquist



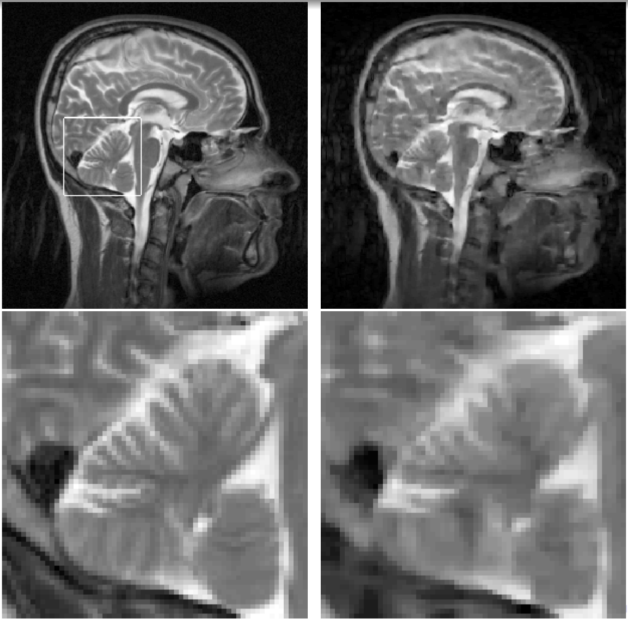


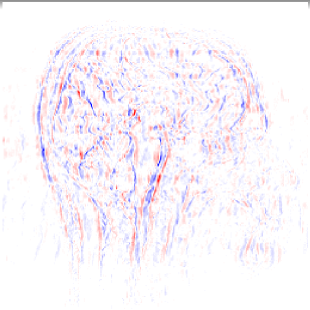
Low Pass



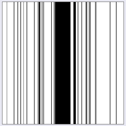
$\ell_2 = 8.0319$

Experimental Results: 1/4 Nyquist



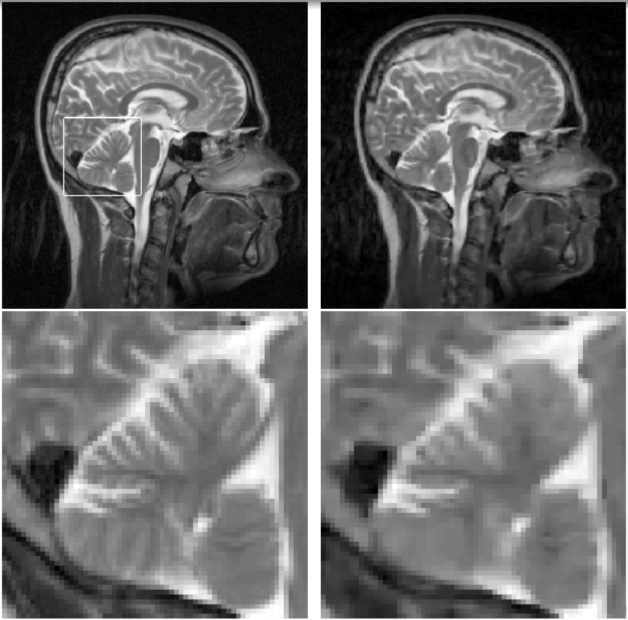


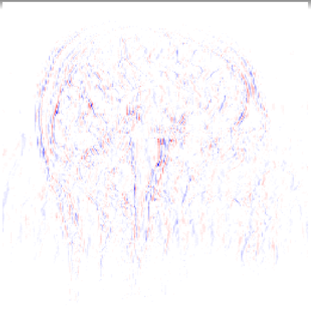
VD Random



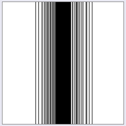
$\ell_2 = 8.8056$

Experimental Results: 1/4 Nyquist





Bayes Optim.



$\ell_2 = 6.0635$