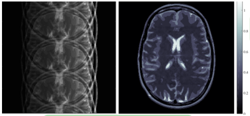
A Plug-and-Play Algorithm for Data-Driven Uncertainty Quantification in Computational Imaging

Michael Tang
Joint work with Audrey Repetti

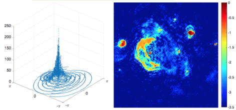
Maxwell Institute for Mathematical Sciences
University of Edinburgh and Heriot-Watt University

Workshop on Recent Advances in Iterative Reconstruction
22-23 May 2023

OBJECTIVE: Find an estimate $x^\dagger$ of $\overline{x}$ from y



$y = \mathcal{D}\left(\Phi \overline{x}\right)$



Inverse problems

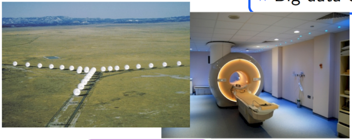
Methodologies

Bayesian inference

- ⋆ Maximum *a posteriori* estimate
- ⋆ Uncertainty quantification

Optimisation methods

- ⋆ Scalable and flexible
- ⋆ Big data context



Applications

MAP Estimate

- FORWARD MODEL: $\boldsymbol{y} = \mathcal{D}(\boldsymbol{\Phi}\overline{\boldsymbol{x}})$
  - $\boldsymbol{y}$ and $\boldsymbol{x}$ are related by a generative statistical model $\mathrm{p}(\boldsymbol{y}|\boldsymbol{x})$
- BAYESIAN FRAMEWORK

$$\begin{aligned}\boldsymbol{x}^\dagger \in \operatorname*{Argmax}_{\boldsymbol{x}\in\mathbb{R}^N} \mathrm{p}(\boldsymbol{x}|\boldsymbol{y}) &= \operatorname*{Argmin}_{\boldsymbol{x}\in\mathbb{R}^N} \left\{ f_{\boldsymbol{y}}(\boldsymbol{x}) = -\log \mathrm{p}(\boldsymbol{x}|\boldsymbol{y}) \right\} \\ &= \operatorname*{Argmin}_{\boldsymbol{x}\in\mathbb{R}^N} -\log \mathrm{p}(\boldsymbol{y}|\boldsymbol{x}) - \log \mathrm{p}(\boldsymbol{x}) \quad \text{(Bayes' formula)} \\ &= \operatorname*{Argmin}_{\boldsymbol{x}\in\mathbb{R}^N} \underbrace{-\log \mathrm{p}\big(\boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{x})\big)}_{=h_{\boldsymbol{y}}(\boldsymbol{x})} \quad \underbrace{-\log \mathrm{p}(\boldsymbol{x})}_{=g(\boldsymbol{x})}\end{aligned}$$

Data fidelity term

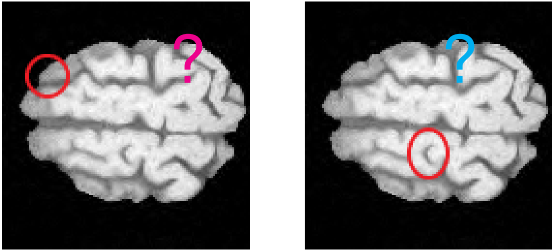
Regularisation term

E.g., $h_{\boldsymbol{y}}(\boldsymbol{x}) = \iota_{\mathcal{B}_2(\boldsymbol{y},\varepsilon)}(\boldsymbol{\Phi}\boldsymbol{x}),\ \frac{1}{\sigma^2}\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|^2$

E.g., $g(\boldsymbol{x}) = \iota_{[0,+\infty[^N}(\boldsymbol{x}),\ \lambda\|\boldsymbol{\Psi}^\dagger\boldsymbol{x}\|_1$

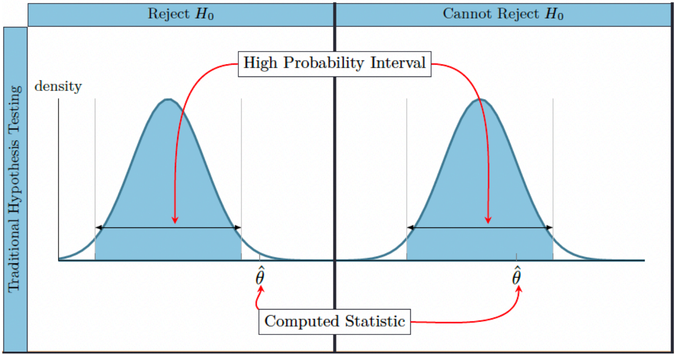
Uncertainty

In general the inverse problem is ill-posed or ill-conditioned.



MAPs are point estimates: No measure of uncertainty.

Hypothesis Testing



BUQO: Bayesian Uncertainty Quantification by Convex Optimization

Let H_0 be a hypothesis:

- $\widetilde{\mathcal{C}}_\alpha$ – HDP region (posterior is small (wrt to $\boldsymbol{x}^\dagger$))
- $\mathcal{S}$ – associated to, hypothesis H_0

Hyp H_0: A structure of interest is ABSENT

Suppose that $\widetilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$.

$\widetilde{\mathcal{C}}_\alpha$

$\mathrm{P}(\boldsymbol{x} \in \widetilde{\mathcal{C}}_\alpha | \boldsymbol{y}) \geqslant 1 - \alpha$

$\mathrm{P}(\boldsymbol{x} \in \mathcal{S} | \boldsymbol{y}) \leqslant \alpha$

$\mathcal{S}$

Then we can say that $\mathrm{P}(\boldsymbol{x} \in \mathcal{S} | \boldsymbol{y}) \leqslant \alpha$.

BUQO: Bayesian Uncertainty Quantification by Convex Optimization

Let H_0 be a hypothesis:

- $\widetilde{\mathcal{C}}_\alpha$ – HDP region (posterior is small (wrt to $\boldsymbol{x}^\dagger$))
- $\mathcal{S}$ – associated to, hypothesis H_0

Hyp H_0: A structure of interest is ABSENT

Suppose that $\widetilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$.

$\widetilde{\mathcal{C}}_\alpha$

$\mathrm{P}(\boldsymbol{x} \in \widetilde{\mathcal{C}}_\alpha | \boldsymbol{y}) \geqslant 1 - \alpha$

$\mathrm{P}(H_0 | \boldsymbol{y}) \leqslant \alpha$

$\mathcal{S}$

Then we can say that $\mathrm{P}(H_0|\boldsymbol{y}) \leqslant \alpha$.

BUQO: Bayesian Uncertainty Quantification by Convex Optimization

Let H_0 be a hypothesis:

- $\widetilde{\mathcal{C}}_\alpha$ – HDP region (posterior is small (wrt to $\boldsymbol{x}^\dagger$))
- $\mathcal{S}$ – associated to, hypothesis H_0

Hyp H_0: A structure of interest is ABSENT

Suppose that $\widetilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$.

$\widetilde{\mathcal{C}}_\alpha$

$\mathrm{P}(\boldsymbol{x} \in \widetilde{\mathcal{C}}_\alpha | \boldsymbol{y}) \geqslant 1 - \alpha$

$\mathrm{P}(H_0 | \boldsymbol{y}) \leqslant \alpha$

$\mathcal{S}$

I.e., REJECT H_0 with confidence α.

BUQO: Variational $\mathcal{S}$

In [Repetti, Pereyra, Wiaux, 2019] $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_3$ where

$$\begin{cases} \mathcal{S}_1 = [0, +\infty)^N & \text{positivity} \\ \mathcal{S}_2 = \{\boldsymbol{x} \in \mathbb{R}^N \mid \boldsymbol{M}\boldsymbol{x} = \boldsymbol{L}\boldsymbol{M}^c\boldsymbol{x} + \boldsymbol{\tau}\} & \text{smooth} \\ \mathcal{S}_3 = \{\boldsymbol{x} \in \mathbb{R}^N \mid \boldsymbol{M}\boldsymbol{x} \in \mathcal{B}_2(\mu, \theta)\} & \text{bdd energy} \end{cases}$$

Comments

- Gaussian operator
- $\boldsymbol{L} = \boldsymbol{L}(\boldsymbol{M})$
- Tune $\boldsymbol{\tau}, \mu, \theta$

Uncertainty quantification

$$\text{Find } (\mathbf{x}^\ddagger_{\mathcal{S}}, \mathbf{x}^\ddagger_{\widetilde{\mathcal{C}}_\alpha}) \in \underset{(\mathbf{x}_{\mathcal{S}}, \mathbf{x}_{\widetilde{\mathcal{C}}_\alpha}) \in \mathcal{S} \times \widetilde{\mathcal{C}}_\alpha}{\operatorname{Argmin}} \frac{1}{2} \|\mathbf{x}_{\mathcal{S}} - \mathbf{x}_{\widetilde{\mathcal{C}}_\alpha}\|^2 \tag{1}$$

Theorem (Variational Hyp. Test)

Suppose that $\alpha \in]4\exp(-N/3), 1[$ *and let* $(\boldsymbol{x}^\ddagger_{\mathcal{S}}, \boldsymbol{x}^\ddagger_{\widetilde{\mathcal{C}}_\alpha}) \in \mathcal{S} \times \widetilde{\mathcal{C}}_\alpha$ *be a solution to* (1). *If* $\|\boldsymbol{x}^\ddagger_{\mathcal{S}} - \boldsymbol{x}^\ddagger_{\widetilde{\mathcal{C}}_\alpha}\| > 0$, *then* $\mathrm{P}(H_0 \mid \boldsymbol{y}) \leqslant \alpha$, *and hence* H_0 *is rejected.*

PnP-BUQO: Data-Driven $\mathcal{S}$

Structure-free

Let $\mathcal{G}$ be an inpainting operator,

$$\mathcal{S} = \{\text{valid } \boldsymbol{x} \mid \boldsymbol{x} \approx \mathcal{G}(\boldsymbol{M}, \boldsymbol{x})\}$$

Data-driven

- parameter-free

$$\text{Find} \quad \boldsymbol{x} \in \operatorname*{Argmin}_{\boldsymbol{x} \in \mathbb{R}^N} \frac{\zeta}{2} \|\boldsymbol{x} - \mathcal{G}(\boldsymbol{x})\|_2^2 + \iota_{\widetilde{\mathcal{C}}_\alpha}(\boldsymbol{x}) \qquad (2)$$

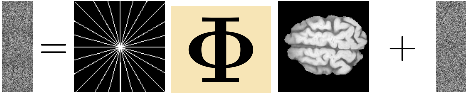
Corollary (Data-Driven Hyp. Test)

Suppose that $\alpha \in]4\exp(-N/3), 1[$ *and let* $\boldsymbol{x}^\ddagger \in \widetilde{\mathcal{C}}_\alpha$ *be a solution to (2). If* $\|\boldsymbol{x}^\ddagger - \mathcal{G}(\boldsymbol{x}^\ddagger)\| > 0$, *then* $\mathrm{P}(H_0 \mid \boldsymbol{y}) \leqslant \alpha$, *and hence* H_0 *is rejected.*

Simulations

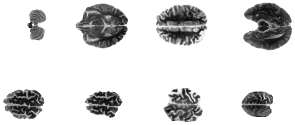
Inverse Problem

$$\boldsymbol{y} = [\ \widetilde{\boldsymbol{\Phi}}\ ]\ \boldsymbol{x} + \boldsymbol{w}$$



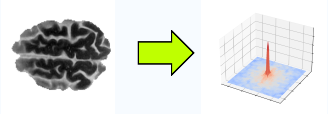
Solve for $\boldsymbol{x}$

BRATS21 MRI dataset



Measurement Operator:

(Discrete) Fourier (torch.fft.fft2)



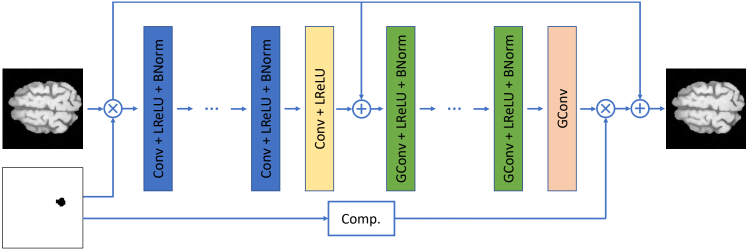
MAP Estimation

- Fidelity: Bounded energy
- Regularize: `db8` wavelet sparsity

Find $\boldsymbol{x}^\dagger$ in

$$\underset{\boldsymbol{x}\in\mathbb{R}^N}{\text{Argmin}}\ \ \iota_{[0,+\infty[^N}(\boldsymbol{x}) + \iota_{\mathcal{B}_2(\boldsymbol{y},\varepsilon)}(\boldsymbol{\Phi}\boldsymbol{x}) + \|\boldsymbol{\Psi}\boldsymbol{x}\|_1$$

Inpainting CNN: DnCNN-like Architecture with Gated Convolutions



DATASET: 90885 images and masks processed from the `BRATS21` dataset.

- optimiser = Adam
- epochs = 32
- learning rate = 0.001
- drop frequency =16
- drop ratio = 0.2

LOSS FUNCTION

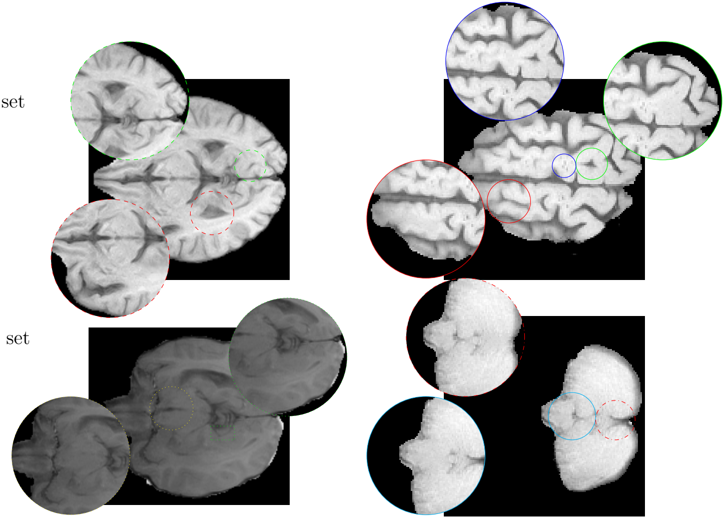
$$\mathcal{L}(\boldsymbol{\theta}) = \alpha \mathrm{MSE}(\boldsymbol{\theta})_+ + \beta \|\mathcal{M}\boldsymbol{\theta}\|_1 + \gamma \partial \mathrm{TV}(\boldsymbol{\theta}) + \delta \mathcal{L}_{\mathrm{percep.}}(\boldsymbol{\theta}) + \epsilon \mathcal{L}_{\mathrm{style}}(\boldsymbol{\theta})$$

- $(\alpha, \beta, \gamma, \delta, \epsilon) = (2, 6, 3, 0.05, 240)$

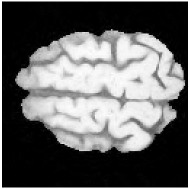
Inpainting: Train/Test Examples

Test set

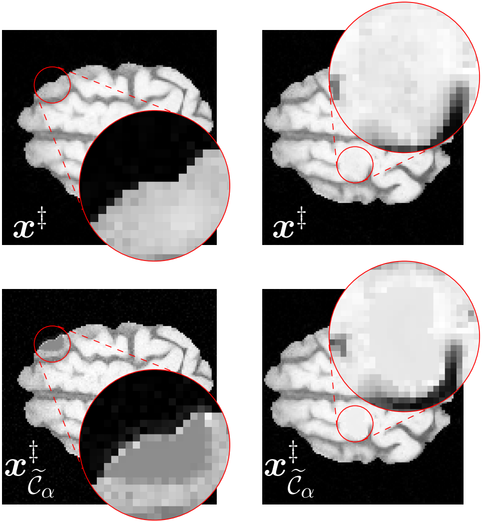
Train set



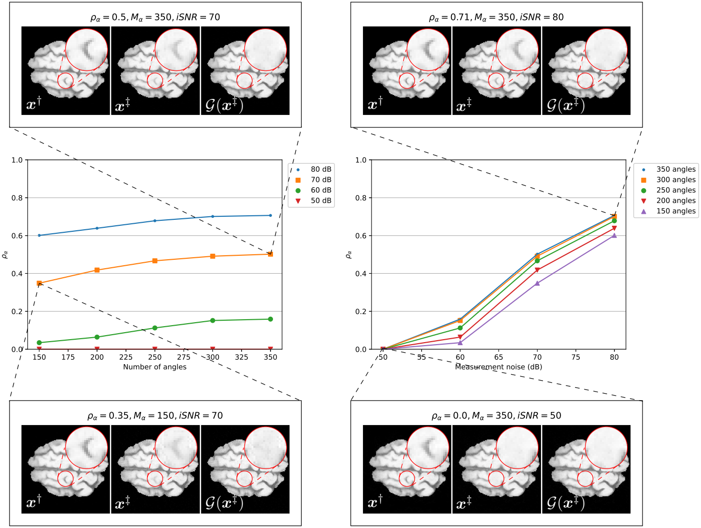
BUQO: Comparison with original method



$\boldsymbol{x}^\dagger$



BUQO: Simulation results



Future work

- Consider other formulations/algos.
- Scale up!
- Improve network architecture/loss/training to improve performance.
    - Training for bespoke artefact definitions/adaptive noise.

REFERENCES

- Marcelo Pereyra. *Maximum-a-posteriori estimation with bayesian confidence regions*. SIAM Journal on Imaging Sciences, 10(1):285-302, 2017
- Audrey Repetti, Marcelo Pereyra, and Yves Wiaux. *Scalable bayesian uncertainty quantification in imaging inverse problems via convex optimization*. SIAM Journal on Imaging Sciences, 12(1):87-118, 2019.
- MT and A. Repetti, A PnP approach to uncertainty quantification with data-driven inpainting operators, *submitted* (Preprint, 2023 arXiv:2304.11200).
- J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, Free-form image inpainting with gated convolution, in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 4470-4479 .

BUQO Algo

Iterations:

for $k = 0, 1, \ldots$ **do**

$$\widetilde{\boldsymbol{v}}_1^{(k)} = \boldsymbol{v}_1^{(k)} + \mu_{1,1} \boldsymbol{\Psi} \boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k)}$$

$$\boldsymbol{v}_1^{(k+1)} = \widetilde{\boldsymbol{v}}_1 - \mu_{1,1} \Pi_{\mathcal{B}_1(\boldsymbol{0}, \widetilde{\eta}_\alpha / \lambda)} \left(\mu_{1,1}^{-1} \widetilde{\boldsymbol{v}}_1^{(k)} \right)$$

$$\widetilde{\boldsymbol{v}}_2^{(k)} = \boldsymbol{v}_2^{(k)} + \mu_{1,2} \boldsymbol{\Phi} \boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k)}$$

$$\boldsymbol{v}_2^{(k+1)} = \widetilde{\boldsymbol{v}}_2 - \mu_{1,2} \Pi_{\mathcal{B}_2(\boldsymbol{y}, \varepsilon)} \left(\mu_{1,2}^{-1} \widetilde{\boldsymbol{v}}_2^{(k)} \right)$$

$$\widetilde{\boldsymbol{x}}_{\widetilde{\mathcal{C}}_\alpha}^{(k)} = \Pi_{[0,1]^N} \left((1 - \gamma\sigma) \boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k)} + \gamma\sigma \boldsymbol{x}_{\mathcal{S}}^{(k)} - \sigma \boldsymbol{\Psi}^\dagger \boldsymbol{v}_1^{(k+1)} - \sigma \boldsymbol{\Phi}^\dagger \boldsymbol{v}_2^{(k+1)} \right)$$

$$\boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k+1)} = 2 \widetilde{\boldsymbol{x}}_{\widetilde{\mathcal{C}}_\alpha}^{(k)} - \boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k)}$$

$$\widetilde{\boldsymbol{u}}_1^{(k)} = \boldsymbol{u}_1^{(k)} + \mu_{2,1} \overline{\boldsymbol{L}} \boldsymbol{x}_{\mathcal{S}}^{(k)}$$

$$\boldsymbol{u}_1^{(k+1)} = \widetilde{\boldsymbol{u}}_1^{(k)} - \mu_{2,1} \Pi_{[-\tau, \tau]^{N_M}} \left(\mu_{2,1}^{-1} \widetilde{\boldsymbol{u}}_1^{(k)} \right)$$

$$\widetilde{\boldsymbol{u}}_2^{(k)} = \boldsymbol{u}_2^{(k)} + \mu_{2,2} \boldsymbol{M} \boldsymbol{x}_{\mathcal{S}}^{(k)}$$

$$\boldsymbol{u}_2^{(k+1)} = \widetilde{\boldsymbol{u}}_2^{(k)} - \mu_{2,2} \Pi_{\mathcal{B}_2(\mu, \theta)} \left(\mu_{2,2}^{-1} \widetilde{\boldsymbol{u}}_2^{(k)} \right)$$

$$\widetilde{\boldsymbol{x}}_{\mathcal{S}}^{(k)} = \Pi_{[0,1]^N} \left((1 - \gamma\sigma) \boldsymbol{x}_{\mathcal{S}}^{(k)} + \gamma\sigma \boldsymbol{x}_{\widetilde{\mathcal{C}}_\alpha}^{(k)} - \sigma \overline{\boldsymbol{L}}^\dagger \boldsymbol{u}_1^{(k+1)} - \sigma \boldsymbol{M}^\dagger \boldsymbol{u}_2^{(k+1)} \right)$$

$$\boldsymbol{x}_{\mathcal{S}}^{(k+1)} = 2 \widetilde{\boldsymbol{x}}_{\mathcal{S}}^{(k)} - \boldsymbol{x}_{\mathcal{S}}^{(k)}$$

end for

Plug-and-Play Algorithm

Initialization: Let $h(\boldsymbol{x}) = r\|\boldsymbol{x} - \mathcal{G}(\boldsymbol{x})\|^2$, $L = \text{Lip}(\nabla h)$ and $\tau^{-1} - \sigma_1\|\boldsymbol{\Psi}\|^2 - \sigma_2\|\boldsymbol{\Phi}\|^2 > L/2$.

Iterations:

for $k = 0, 1, \ldots$ **do**

$$
\begin{aligned}
\widetilde{\boldsymbol{v}}_1^{(k)} &= \boldsymbol{v}_1^{(k)} + \sigma_1 \boldsymbol{\Psi x} \\
\boldsymbol{v}_1^{(k+1)} &= \widetilde{\boldsymbol{v}}_1^{(k)} - \sigma_1 \Pi_{\mathcal{B}_1(\mathbf{0}, \widetilde{\eta}_\alpha/\lambda)}(\sigma_1^{-1}\widetilde{\boldsymbol{v}}_1^{(k)}) \\
\widetilde{\boldsymbol{v}}_2^{(k)} &= \boldsymbol{v}_2^{(k)} + \sigma_2 \boldsymbol{\Phi x}^{(k)} \\
\boldsymbol{v}_2^{(k+1)} &= \widetilde{\boldsymbol{v}}_2^{(k)} - \sigma_2 \Pi_{\mathcal{B}_2(\boldsymbol{y}, \varepsilon)}(\sigma_2^{-1}\widetilde{\boldsymbol{v}}_1^{(k)}) \\
\widetilde{\boldsymbol{x}}^{(k)} &= \Pi_{[0,\infty[^N} \left(\boldsymbol{x}^{(k)} - \tau \left(\nabla h(x) + \boldsymbol{\Psi}^\dagger \boldsymbol{v}_1^{(k+1)} + \boldsymbol{\Phi}^\dagger \boldsymbol{v}_2^{(k+1)} \right) \right) \\
\boldsymbol{x}^{(k+1)} &= 2\widetilde{\boldsymbol{x}}^{(k)} - \boldsymbol{x}^{(k)}
\end{aligned}
$$

end for