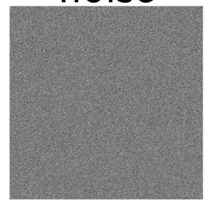
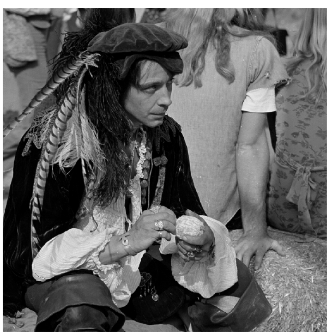
Imaging inverse problems

noise





$u^\star$

H

y

Imaging inverse problems

- Forward model:
$$y = Hu + w, \tag{1}$$
where,
  - $u \in \mathbb{R}^d$ unknown image, $y \in \mathbb{R}^d$ observed data and $d \in \mathbb{N}$,
  - H a circulant block matrix of dimension $d \times d$ obtained from a blur kernel h and ...
  - $w \sim \mathcal{N}(0, \sigma^2 Id)$ noise, $\sigma^2 > 0$.
- Deconvolution problems: Estimating u from y.

Deconvolution problems can be broadly classified in 3 groups: Non-blind, blind and semi-blind.

Imaging inverse problems

- Non-blind problems: H is known and u is unknown.
- Semi-blind problems: $H \in \mathcal{K}$ where,

$$\mathcal{K} = \left\{ H(\alpha) : \mathbb{R}^d \longrightarrow \mathbb{R}^d, \alpha \in \Theta_\alpha \right\}.$$

  - *Pros:* Introduces more structure.
  - *Cons:* The problem is non-linear w.r.t α
- Blind problems: Both H and u are completely unknown.

These deconvolution problems are ill-posed and additional information about u must be considered.

Motivation...

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u

(b) y $(19.81$ dB)

We simulate y with $\alpha = (\alpha_h, \alpha_v) = (0.4, 0.3)$

Motivation...

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u (b) y (19.81dB) (c) $\hat{u}(0.4, 0.3)$ (29.5dB)

We simulate y with $\alpha = (\alpha_h, \alpha_v) = (0.4, 0.3)$

HERIOT WATT UNIVERSITY

Motivation...

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u

(b) y (19.81dB)

(c) $\hat{u}(0.3, 0.2)$ (28.9dB)

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u

(b) y (19.81dB)

(c) $\hat{u}(0.2, 0.2)$ (28.0dB)

Motivation...

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u

(b) y (19.81dB)

(c) $\hat{u}(0.1, 0.1)$ (27.2dB)

- considering a parametric Gaussian forward model $H(\alpha)$

(a) u

(b) y (19.81dB)

(c) $\hat{u}(0.8, 0.7)$ (25.7dB)

Motivation...

- considering a parametric Gaussian forward model $H(\alpha)$

(a) $\hat{u}(0.4, 0.3)$ (29.5dB)

(b) $\hat{u}(0.1, 0.1)$ (27.2dB)

(c) $\hat{u}(0.8, 0.7)$ (25.7dB)

HERIOT WATT UNIVERSITY

Main Objectives

1. Estimate α and σ^2 from the observed data y.

2. Restore the image u from the observed data y, given the estimates of $\hat{\alpha}$ and $\hat{\sigma}^2$.

Problem formulation in Bayesian framework

From the Bayes theorem, the posterior distribution is

$$p(u|y,\theta,\alpha,\sigma^2) = \frac{p(y|u,\alpha,\sigma^2)p(u|\theta)}{p(y|\theta,\alpha,\sigma^2)},$$

where,

- Likelihood function

$$p(y|u,\alpha,\sigma^2) \propto \exp\left(-f^y_{\alpha,\sigma^2}(u)\right), \quad f^y_{\alpha,\sigma^2}(u) = \frac{1}{2\sigma^2}||y - H(\alpha)u||^2$$

- Prior distribution

$$p(u|\theta) = \exp\left(-\theta g(u)\right)/Z(\theta), \quad Z(\theta) = \int_{\mathbb{R}^d} e^{-\theta g(\tilde{u})} du$$

$u \mapsto g(u)$ is the *TV norm* and θ the regularisation parameter.

- Marginal likelihood

$$p(y|\theta,\alpha,\sigma^2) = \int_{\mathbb{R}^d} p(y|\tilde{u},\alpha,\sigma^2)p(\tilde{u}|\theta)d\tilde{u}.$$

Methodology

Our proposed method is based on the empirical Bayesian approach:

1. We evaluate the Maximum Marginal likelihood estimator from y,

$$(\hat{\theta}, \hat{\alpha}, \hat{\sigma}^2) \in \underset{\theta \in \Theta_\theta, \alpha \in \Theta_\alpha, \sigma \in \Theta_{\sigma^2}}{\operatorname{argmax}} \log p(y|\theta, \alpha, \sigma^2).$$

   where

$$p(y|\theta, \alpha, \sigma^2) = \int_{\mathbb{R}^d} \log p(y|\tilde{u}, \alpha, \sigma^2) p(\tilde{u}|\theta) d\tilde{u}.$$

2. Given $\hat{\theta}$, $\hat{\alpha}$ and $\hat{\sigma}^2$, we estimate u by maximum-a-posteriori estimation using the pseudo posterior $p(u|y, \hat{\theta}, \hat{\alpha}, \hat{\sigma}^2)$

$$\hat{u}_{MAP} = \underset{\tilde{u} \in \mathbb{R}^d}{\operatorname{argmin}} \left\{ f^y_{\hat{\alpha}, \hat{\sigma}^2}(\tilde{u}) + \hat{\theta} g(\tilde{u}) \right\}. \qquad (2)$$

HERIOT WATT UNIVERSITY

Step 1: Estimation of $\hat{\alpha}, \hat{\theta}$ and $\hat{\sigma}^2$

Projected gradient descent (PGD)

$$\begin{cases} \theta_{n+1} & = \Pi_{\Theta_\theta}\left[\theta_n + \delta_{n+1} \nabla_\theta \log(p(y|\theta,\alpha,\sigma^2))\right] \\ \\ \alpha_{n+1} & = \Pi_{\Theta_\alpha}\left[\alpha_n + \delta_{n+1} \nabla_\alpha \log(p(y|\theta,\alpha,\sigma^2))\right] \\ \\ \sigma^2_{n+1} & = \Pi_{\Theta_{\sigma^2}}\left[\sigma^2_n + \delta_{n+1} \nabla_{\sigma^2} \log(p(y|\theta,\alpha,\sigma^2))\right] \end{cases}, \qquad \forall n \in \mathbb{N}.$$

where,

$(\delta_n)_{n\in\mathbb{N}}$ decreasing sequence of step size

$\Theta_\alpha, \Theta_\theta$ and Θ_{σ^2} are a convex sets, and Π_Θ projection operator.

Step 1: Estimation of $\hat{\alpha}, \hat{\theta}$ and $\hat{\sigma}^2$

Projected gradient descent (PGD)

$$\begin{cases} \theta_{n+1} & = \Pi_{\Theta_\theta}\left[\theta_n + \delta_{n+1}\nabla_\theta \log(p(y|\theta,\alpha,\sigma^2))\right] \\ \alpha_{n+1} & = \Pi_{\Theta_\alpha}\left[\alpha_n + \delta_{n+1}\nabla_\alpha \log(p(y|\theta,\alpha,\sigma^2))\right] \\ \sigma^2_{n+1} & = \Pi_{\Theta_{\sigma^2}}\left[\sigma^2_n + \delta_{n+1}\nabla_{\sigma^2} \log(p(y|\theta,\alpha,\sigma^2))\right] \end{cases}, \qquad \forall n \in \mathbb{N}.$$

We recall that

$$p(y|u,\theta,\alpha,\sigma^2) = \int_{\mathbb{R}^d} p(y|\tilde{u},\alpha,\sigma^2)p(\tilde{u}|\theta)d\tilde{u}.$$

Thus, $\nabla_\theta \log p(y|u,\theta,\alpha,\sigma^2)$, $\nabla_\alpha \log p(y|u,\theta,\alpha,\sigma^2)$ and $\nabla_{\sigma^2} \log p(y|u,\theta,\alpha,\sigma^2)$ are intractable.

Step 1: Gradients approximation

Using Fisher's Identity,

$$\nabla_{\theta} \log p(y|\theta, \alpha, \sigma^2) = -\underbrace{\int_{\mathbb{R}^d} g(u) p(u|y, \theta, \alpha, \sigma^2) du}_{\mathbb{E}_{u \sim u|y}[g(u)]} - \frac{d}{\theta}$$

$$\nabla_{\alpha} \log p(y|\theta, \alpha, \sigma^2) = -\underbrace{\int_{\mathbb{R}^d} \nabla_{\alpha} f^y_{\alpha,\sigma^2}(u) p(u|y, \theta, \alpha, \sigma^2) du}_{\mathbb{E}_{u \sim u|y}\left[\nabla_{\alpha} f^y_{\alpha,\sigma^2}(u)\right]}$$

$$\nabla_{\sigma^2} \log p(y|\theta, \alpha, \sigma^2) = -\underbrace{\int_{\mathbb{R}^d} \nabla_{\sigma^2} f^y_{\alpha,\sigma^2}(u) p(u|y, \theta, \alpha, \sigma^2) du}_{\mathbb{E}_{u \sim u|y}\left[\nabla_{\sigma^2} f^y_{\alpha,\sigma^2}(u)\right]} - \frac{d}{2\sigma^2}$$

Step 1: Gradients approximation

MYULA Markov kernel [Durmus et al., 2018]

$$X_{k+1} = (1 - \frac{\gamma}{\lambda})X_k - \gamma \nabla_x f^y_{\alpha,\sigma^2}(X_k) + \frac{\gamma}{\lambda} prox^{\lambda}_{g_\theta}(X_k) + \sqrt{2\gamma} Z_{k+1}$$

where $\lambda > 0, g_\theta = \theta g$ and,

$$prox^{\lambda}_{g_\theta}(u) = \operatorname*{argmin}_{v} g_\theta(v) + \frac{1}{2\lambda}||u - v||_2^2$$

Given the $(X_k)_{k \in \mathbb{N}}$, we obtain the following approximation

- $\mathbb{E}_{u \sim p(u|y,\theta,\alpha,\sigma^2)}\left[g(u)\right] \approx \frac{1}{m}\sum_{k=0}^{m} g(X_k)$
- $\mathbb{E}_{u \sim p(u|y,\theta,\alpha,\sigma^2)}\left[\nabla_\alpha f^y_{\alpha,\sigma^2}(u)\right] \approx \frac{1}{m}\sum_{k=0}^{m} \nabla_\alpha f^y_{\alpha,\sigma^2}(X_k)$
- $\mathbb{E}_{u \sim p(u|y,\theta,\alpha,\sigma^2)}\left[\nabla_{\sigma^2} f^y_{\alpha,\sigma^2}(u)\right] \approx \frac{1}{m}\sum_{k=0}^{m} \nabla_{\sigma^2} f^y_{\alpha,\sigma^2}(X_k)$

HERIOT WATT UNIVERSITY

Step 2: MAP estimate [Afonso et al., 2010]

Having $\hat{\theta}$, $\hat{\alpha}$ and $\hat{\sigma^2}$,

- Pseudo posterior

$$p(u|y, \hat{\theta}, \hat{\alpha}, \hat{\sigma^2}) \propto p(y|u, \hat{\alpha}, \hat{\sigma^2})p(u|\hat{\theta})$$

- ... which is then used to recover u by computing the MAP estimation of the following optimisation problem

$$\hat{u}_{MAP} = \underset{\tilde{u}\in\mathbb{R}^d}{\operatorname{argmin}} \left\{ f^y_{\hat{\alpha},\hat{\sigma^2}}(\tilde{u}) + \hat{\theta} g(\tilde{u}) \right\}.$$

HERIOT WATT UNIVERSITY

General algorithm

Algorithm 1 Algo.

Set $\Theta_\theta, \Theta_\alpha$ and Θ_{σ^2}, λ, N, $\{\gamma_n\}_{n=1}^N$ $\{\delta_n\}_{n=1}^N$, $\{m_n\}_{n=1}^N$.

Initialization: $\{\theta_0, \alpha_0, \sigma_0^2, X_0^0\}$.

for $n = 0 : N-1$ **do**

  $(X_k^n)_{k=1}^{m_n}$ from MYULA kernel

  $\theta_{n+1} = \Pi_{\Theta_\theta}\left[\theta_n + \dfrac{\delta_{n+1}}{m_n}\sum_{k=1}^{m_n}\left\{\frac{d}{q\theta_n} - g(X_k^n)\right\}\right]$

  $\alpha_{n+1} = \Pi_{\Theta_\alpha}\left[\alpha_n - \dfrac{\delta_{n+1}}{m_n}\sum_{k=1}^{m_n}\nabla_\alpha f^y_{\alpha_n,\sigma_n^2}(X_k^n)\right]$

  $\sigma^2_{n+1} = \Pi_{\Theta_{\sigma^2}}\left[\sigma_n^2 - \dfrac{\delta_{n+1}}{m_n}\sum_{k=1}^{m_n}\left\{\nabla_{\sigma^2} f^y_{\alpha_n,\sigma_n^2}(X_k^n) + \dfrac{d}{2\sigma_n^2}\right\}\right]$

end for

Compute $\bar{\theta}_N$, $\bar{\alpha}_N$ and $\bar{\sigma}_N^2$.

return $\bar{x}_{MAP} = \operatorname{argmax}_{x\in\mathbb{R}^d} p(x|y, \bar{\theta}, \bar{\alpha}, \bar{\sigma}^2)$

HERIOT WATT UNIVERSITY

Illustration of the method

- Gaussian blur operator.
  - In discrete domain we have, for all $i \in \mathcal{X}, j \in \mathcal{Y}$

$$h_{\alpha_h,\alpha_v}(i,j) = \frac{\alpha_h \alpha_v}{2\pi} \exp\left(-\frac{1}{2}\left({\alpha_h}^2 i^2 + {\alpha_v}^2 j^2\right)\right)$$

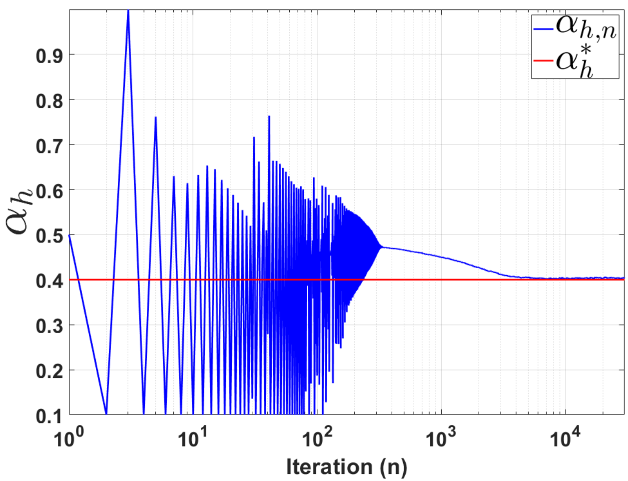
1. $\sum_{i\in\mathcal{X},j\in\mathcal{Y}} h_{\alpha_h,\alpha_v}(i,j) = 1,$ $\times$ Not satisfied
2. For all $i \in \mathcal{X}, j \in \mathcal{Y}$, $h_{\alpha_h,\alpha_v}(i,j) \geq 0.$ ✓ Satisfied!!

Correction of h_{α_h,α_v},

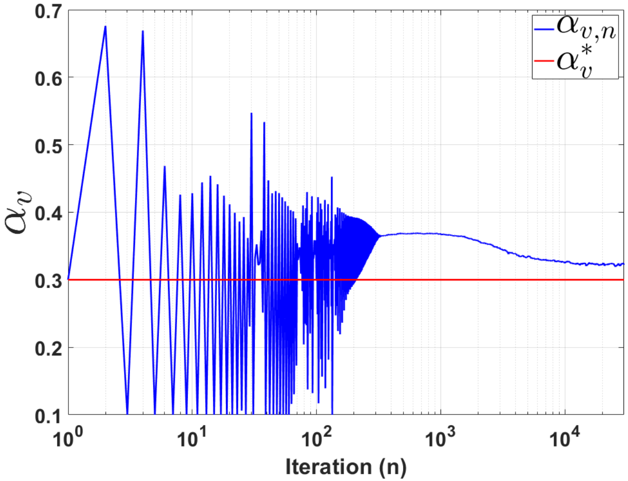
$$\tilde{h}_{\alpha_h,\alpha_v}(i,j) = \frac{\alpha_h \alpha_v}{2\pi} \exp\left(-\frac{1}{2}\left({\alpha_h}^2 i^2 + {\alpha_v}^2 j^2\right)\right) / C(\alpha_h, \alpha_v)$$

where $C(\alpha_h, \alpha_v) = \sum_{i\in\mathcal{X},j\in\mathcal{Y}} h_{\alpha_h,\alpha_v}(i,j)$

Results: α_v and α_h iterates

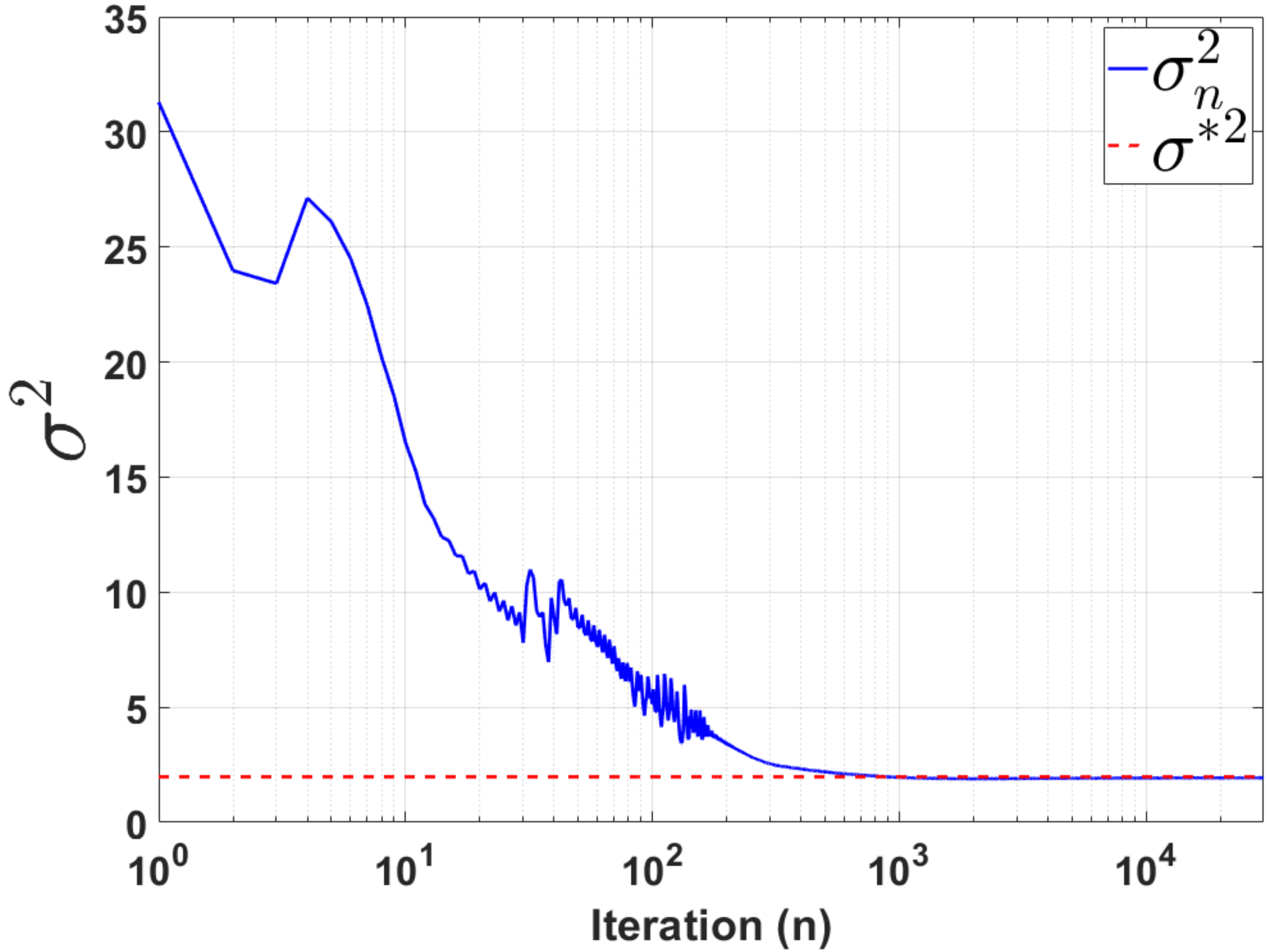


(a) α_h

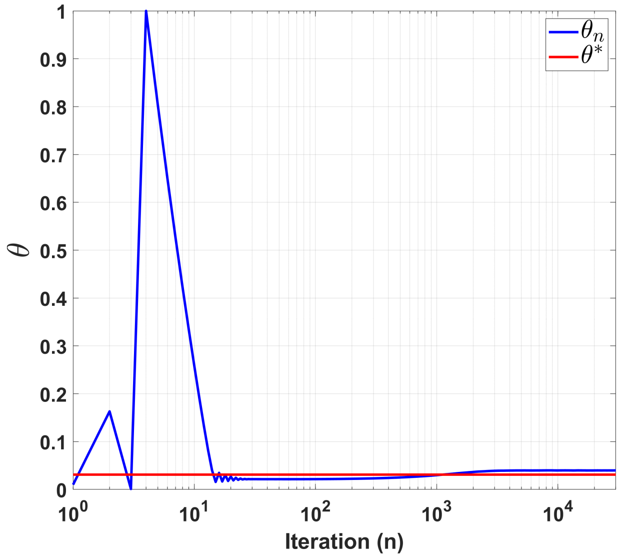


(b) α_v

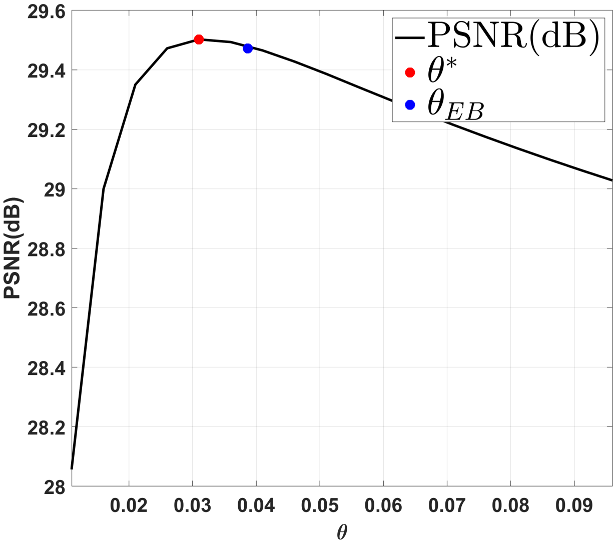
Results: σ^2 iterates



Results: θ iterates



(a) α_h



(b) α_v

Results: $\bar{\sigma}^2$ & $\sigma^{\star 2}$

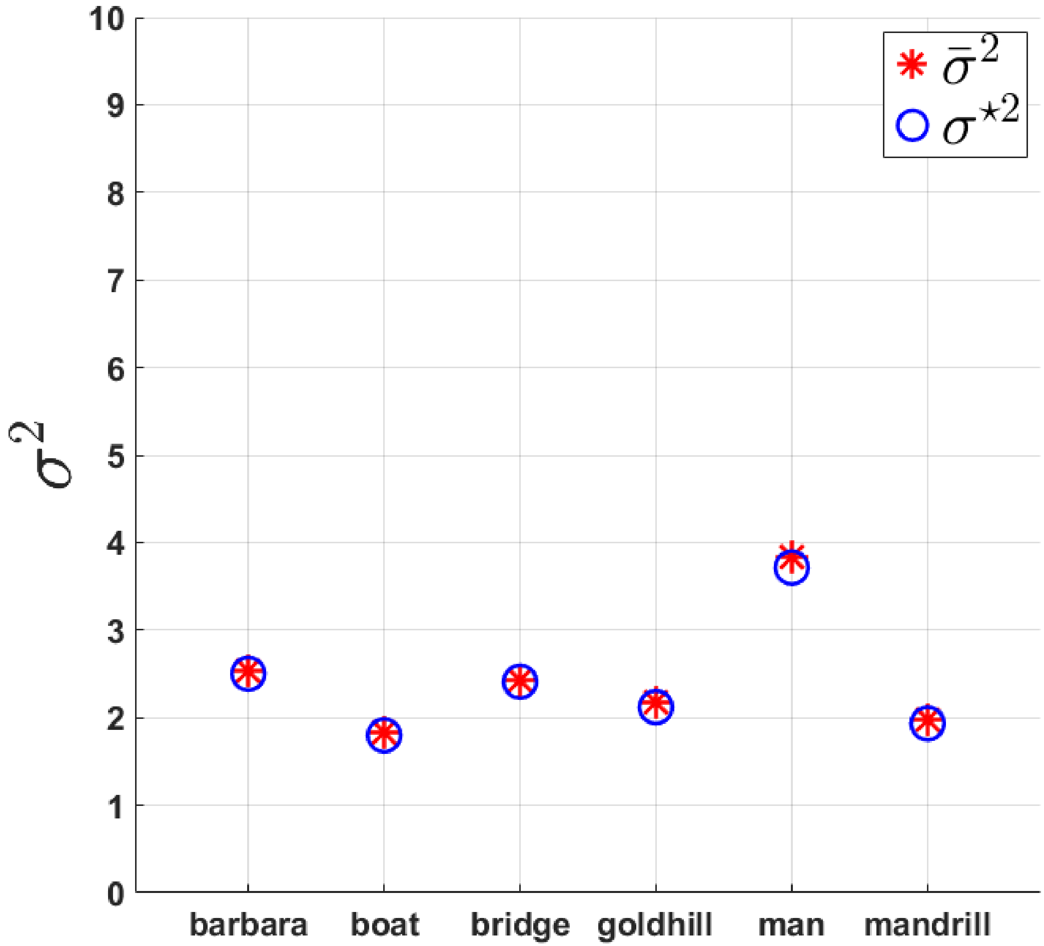


Figure 11: $\bar{\sigma}^2$ & $\sigma^{\star 2}$

Results: $\bar{\theta}$ & $\theta^{\star}$

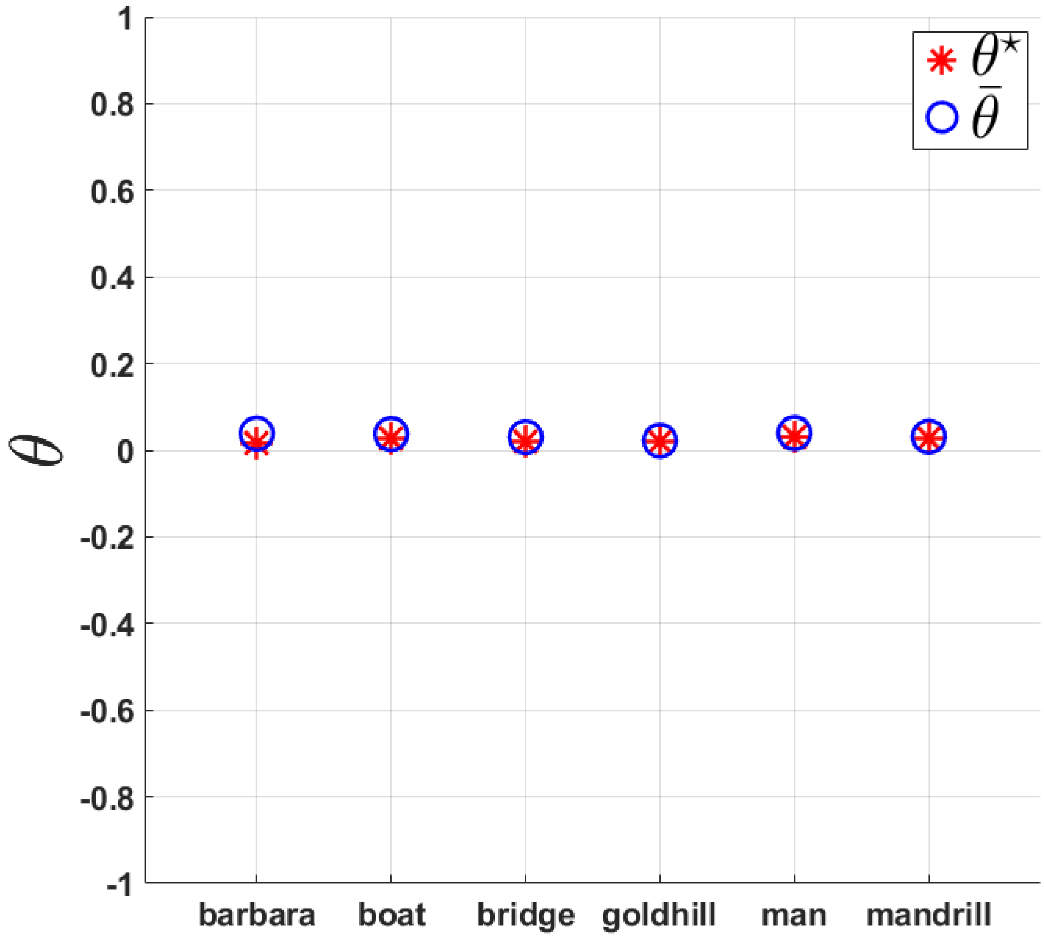


Figure 12: $\bar{\theta}$ & $\theta^{\star}$

Results: Summary

Table 1: Summary results

Parameters and errors	SNR = 20dB	SNR = 30dB
$\bar{\alpha_h} \pm$ std $(\alpha_h^\star = 0.4)$	$0.39 \pm 0.30 \times 10^{-2}$	$0.40 \pm 0.02 \times 10^{-2}$
$\bar{\alpha_v} \pm$ std $(\alpha_v^\star = 0.3)$	$0.29 \pm 0.20 \times 10^{-2}$	$0.31 \pm 3.4 \times 10^{-2}$
$\frac{\lvert\bar{\theta}-\theta^*\rvert}{\theta^*} \pm$ std	0.74 ± 0.12	0.48 ± 0.17
$\frac{\lvert\bar{\sigma^2}-\sigma^{*2}\rvert}{\sigma^{*2}} \pm$ std	$0.01 \pm 7.4 \times 10^{-5}$	$0.02 \pm 9.4 \times 10^{-5}$
ℓ_1-error $\pm$std	$0.01 \pm 0.12 \times 10^{-3}$	$4.5 \times 10^{-3} \pm 0.08 \times 10^{-4}$

HERIOT WATT UNIVERSITY

Results: MAP estimate

(a) u

(b) y $(19.81$ dB)

We simulate y with $\alpha = (\alpha_h, \alpha_v) = (0.4, 0.3)$

Results: MAP estimate

(a) y (19.81 dB) (b) $\hat{u}$ (28.75dB) (c) $\hat{u}$ (**29.46dB**)

Results: MAP estimate

(a) u

(b) $\hat{u}$ (28.75dB)

(c) $\hat{u}$ (**29.46dB**)

Conclusion

To conclude:

- We proposed a provably empirical Bayesian approach to perform inference in semi-blind image deconvolution problems.
- The proposed SAPG algorithm can be deployed to any semi-blind inverse problems that is convex w.r.t. the unknown image