Bayesian formulation of Regularization by denoising. Application to image restoration

Diarra FALL
Institut Denis Poisson, Unité Mixte de Recherche
CNRS, Université d'Orléans, Université de Tours, FRANCE.

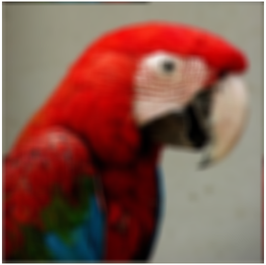
Joint work with: Elhadji C. Faye and Nicolas Dobigeon

Mathematical Foundations of AI - Paris, 25/03/2025

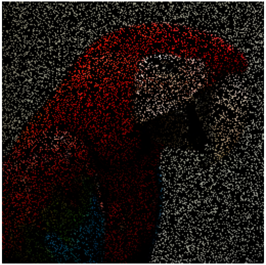
Image Restoration

- Goal : restore an original image from its observed degraded version.

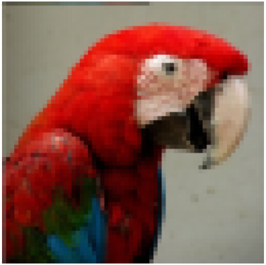
Deblurring



Inpainting



Super-resolution



Problem Statement

Consider image restoration as an inverse problem :

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

- $\mathbf{x} \in \mathbb{R}^n$: true unknown image,
- $\mathbf{y} \in \mathbb{R}^m$: measured data,
- $\mathbf{A} \in \mathbb{R}^{m \times n}$: degradation matrix,
- $\mathbf{n} \in \mathbb{R}^n$: additive noise.

- Estimating $\mathbf{x}$ from $\mathbf{y}$ is generally an ill-posed or, at least, ill-conditioned problem.
- $\Rightarrow$ Additional information is needed.

Bayesian Modeling

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n},$$

- A Bayesian model assumes that the parameter is a **random** quantity.
- Ingredients :
  - **Prior** over parameter : $\mathbf{x} \sim p(\mathbf{x})$.
  - **Likelihood** : $\mathbf{y}|\mathbf{x} \sim p(\mathbf{y}|\mathbf{x})$.
  - **Posterior** : compute the parameter distribution given the data.
    $\rightarrow$ *Bayes' theorem :*

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x})p(\mathbf{y}|\mathbf{x})}{\int p(\mathbf{x})p(\mathbf{y}|\mathbf{x})d\mathbf{x}}.$$

Bayesian Modeling

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n},$$

- A Bayesian model assumes that the parameter is a **random** quantity.
- Ingredients :
  - **Prior** over parameter : $\mathbf{x} \sim p(\mathbf{x})$.
  - **Likelihood** : $\mathbf{y}|\mathbf{x} \sim p(\mathbf{y}|\mathbf{x})$.
  - **Posterior** : compute the parameter distribution given the data.
    $\rightarrow$ *Bayes' theorem :*

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x})p(\mathbf{y}|\mathbf{x})}{\int p(\mathbf{x})p(\mathbf{y}|\mathbf{x})d\mathbf{x}}.$$

- One can compute an *estimate* of the parameter
  - *Maximum a posteriori* : $\hat{\mathbf{x}}_{\text{MAP}} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{x}|\mathbf{y})$.
  - *Posterior mean* : $\hat{\mathbf{x}}_{\text{PM}} = \mathbb{E}[\mathbf{x}|\mathbf{y}] = \int \mathbf{x}p(\mathbf{x}|\mathbf{y})d\mathbf{x}$.
- and perform uncertainty quantification through credible intervals/ pixelwise standard deviations.

Image Restoration Formulation

- In the following

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left[-f(\mathbf{x},\mathbf{y})\right],$$
$$p(\mathbf{x}) \propto \exp\left[-\beta g(\mathbf{x})\right],$$

such that

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\left[-f(\mathbf{x},\mathbf{y}) - \beta g(\mathbf{x})\right]. \quad (2)$$

Image Restoration Formulation

- In the following

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left[-f(\mathbf{x},\mathbf{y})\right],$$
$$p(\mathbf{x}) \propto \exp\left[-\beta g(\mathbf{x})\right],$$

such that

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\left[-f(\mathbf{x},\mathbf{y}) - \beta g(\mathbf{x})\right]. \tag{2}$$

- How to choose the prior potential g ? $\rightarrow$ Deep learning based prior.
- Regularization by Denoising (RED) [Romano, Elad, Milanfar, 2017].

  Use an explicit regularization term :

$$g_{\mathrm{red}}(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T(\mathbf{x} - \mathcal{D}_\nu(\mathbf{x})), \tag{3}$$

where $\mathcal{D}_\nu : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a (deep) denoiser designed for the removal of additive white Gaussian noise.

Proposed Probabilistic counterpart of RED

- We introduce a prior distribution defined from the RED potential $g_{\mathrm{red}}(\cdot)$ and define

$$p_{\mathrm{red}}(\mathbf{x}) \propto \exp\left[-\beta g_{\mathrm{red}}(\mathbf{x})\right] = \exp\left[-\frac{\beta}{2}\mathbf{x}^\top\left(\mathbf{x} - \mathsf{D}_\nu(\mathbf{x})\right)\right]. \quad (4)$$

- For $p_{\mathrm{red}}(\cdot)$ to be a valid p.d.f., i.e. $\int_{\mathbb{R}^n} p_{\mathrm{red}}(\mathbf{x})d\mathbf{x} < \infty$, certain conditions must be satisfied.

Proposed Probabilistic counterpart of RED

- We introduce a prior distribution defined from the RED potential $g_{\mathrm{red}}(\cdot)$ and define

$$p_{\mathrm{red}}(\mathbf{x}) \propto \exp\left[-\beta g_{\mathrm{red}}(\mathbf{x})\right] = \exp\left[-\frac{\beta}{2}\mathbf{x}^\top\left(\mathbf{x} - \mathrm{D}_\nu(\mathbf{x})\right)\right]. \quad (4)$$

- For $p_{\mathrm{red}}(\cdot)$ to be a valid p.d.f., i.e. $\int_{\mathbb{R}^n} p_{\mathrm{red}}(\mathbf{x})d\mathbf{x} < \infty$, certain conditions must be satisfied.

Assumption 1

The Hessian matrix $\mathbf{\Lambda}(\mathbf{x}) = \mathbf{I}_n - \nabla \mathrm{D}_\nu(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^n$, has at least one non-zero eigenvalue.

- Assumption 1 is violated only in trivial cases.

Proposition 1

If Assumption 1 + Some conditions on the denoiser hold, then

$$\int_{\mathbb{R}^n} p_{\mathrm{red}}(\mathbf{x})d\mathbf{x} < +\infty$$

and $p_{\mathrm{red}}(\cdot)$ in (4) defines a proper p.d.f.

Proposed Sampling Scheme

- Goal : sample from the posterior

$$\pi(\mathbf{x}) = p(\mathbf{x}|\mathbf{y}) \propto \exp\left[-f(\mathbf{x},\mathbf{y}) - \beta g_{\mathrm{red}}(\mathbf{x})\right]. \tag{5}$$

- Leverage an asymptotically exact data augmentation (AXDA) as introduced by [Vono, Dobigeon, Chainais, 2019].
- Idea : introduce an auxiliary variable $\mathbf{z} \in \mathbb{R}^n$ and consider the augmented posterior distribution

$$\begin{aligned} \pi_\rho(\mathbf{x},\mathbf{z}) &= p(\mathbf{x},\mathbf{z}|\mathbf{y};\rho^2) \qquad (6) \\ &\propto \exp\left[-f(\mathbf{x},\mathbf{y}) - \beta g_{\mathrm{red}}(\mathbf{z}) - \frac{1}{2\rho^2}||\mathbf{x}-\mathbf{z}||^2\right], \end{aligned}$$

$\rho > 0$: parameter controlling the dissimilarity between $\mathbf{x}$ and $\mathbf{z}$.

Proposed Sampling Scheme

- The SGS alternatively samples according to the two conditionals to generate samples asymptotically distributed according to $\pi_\rho(\mathbf{x}, \mathbf{z})$:

$$p_\rho(\mathbf{x}|\mathbf{y}, \mathbf{z}) \propto \exp\left[-f(\mathbf{x}, \mathbf{y}) - \frac{1}{2\rho^2}||\mathbf{x} - \mathbf{z}||^2\right] \tag{7}$$

$$p_\rho(\mathbf{z}|\mathbf{x}) \propto \exp\left[-\beta g_{\text{red}}(\mathbf{z}) - \frac{1}{2\rho^2}||\mathbf{x} - \mathbf{z}||^2\right]. \tag{8}$$

- Advantages of splitting :
  - ✓ The separation of the two components $f(\cdot, \cdot)$ and $g_{\text{red}}(\cdot)$.
  - ✓ Simpler, scalable and more efficient sampling schemes.

Proposed Sampling Scheme

- The SGS alternatively samples according to the two conditionals to generate samples asymptotically distributed according to $\pi_\rho(\mathbf{x}, \mathbf{z})$:

$$p_\rho(\mathbf{x}|\mathbf{y}, \mathbf{z}) \propto \exp\left[-f(\mathbf{x}, \mathbf{y}) - \frac{1}{2\rho^2}||\mathbf{x} - \mathbf{z}||^2\right] \quad (7)$$

$$p_\rho(\mathbf{z}|\mathbf{x}) \propto \exp\left[-\beta g_{\mathrm{red}}(\mathbf{z}) - \frac{1}{2\rho^2}||\mathbf{x} - \mathbf{z}||^2\right]. \quad (8)$$

- Advantages of splitting :
  - ✓ The separation of the two components $f(\cdot, \cdot)$ and $g_{\mathrm{red}}(\cdot)$.
  - ✓ Simpler, scalable and more efficient sampling schemes.
- Letting

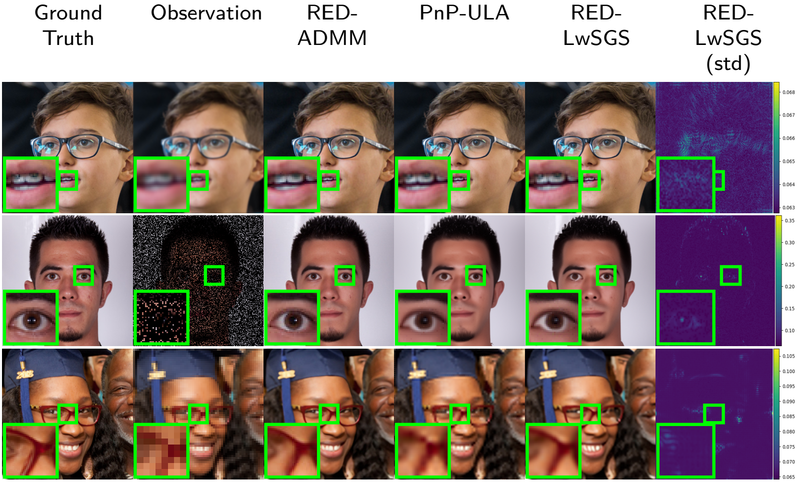
$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2\sigma^2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2, \quad (9)$$

  Sampling from (7) $\Leftrightarrow$ sampling from a high-dimensional Gaussian.
- Sampling from (8) not easy due to the regularization potential $g_{\mathrm{red}}(\cdot)$ involving the denoiser $\mathrm{D}_\nu(\cdot) \Rightarrow$ ULA step :

$$\mathbf{z}^{(t+1)} = \mathbf{z}^{(t)} + \gamma \nabla \log p_\rho\left(\mathbf{z}^{(t)} \mid \mathbf{x}\right) + \sqrt{2\gamma}\varepsilon^{(t)} \quad (10)$$

$\gamma > 0$: step-size ; $\left\{\varepsilon^{(t)}\right\}$: Gaussian random variables.

Numerical Experiments



FFHQ data set : images recovered by the compared methods for deblurring (top), inpainting (middle) and super-resolution (bottom).

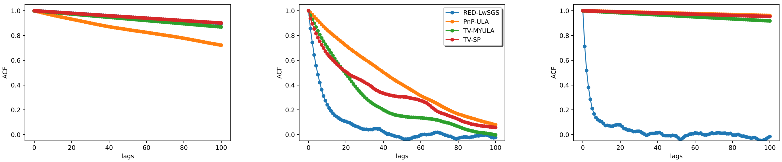
Performance Metrics

	PSNR(dB) ↑	**SSIM** ↑	**LPIPS** ($\times 10^{-2}$) ↓	**FID** ↓	**Time**(s) ↓
			MCMC-based methods		
RED-LwSGS	30.43±2.161	0.872±0.036	3.519±2.139	100.09±34.62	115±25
PnP-ULA	29.01±2.013	0.847±0.037	5.025±2.359	202.15±48.02	128±40
TV-MYULA	28.99±2.017	0.847±0.037	4.925±2.435	202.29±48.30	133±26
TV-SP	28.94±2.019	0.846±0.037	5.125±2.435	199.16±47.63	112±23
			Optimization-based methods		
RED-ADMM	30.49±2.222	<u>0.875</u>±0.036	<u>3.418</u>±2.038	95.48±33.51	3±1
RED-HQS	<u>30.54</u>±2.206	**0.876**±0.036	<u>3.418</u>±2.037	95.10±33.33	3±1
PnP-ADMM	30.13±2.184	0.867±0.037	3.510±2.140	110.04±36.31	3±1
DiffPIR	**30.99**±2.212	0.868±0.034	**1.112**±0.821	<u>65.16</u>±30.15	50±4
DPS	28.94±1.798	0.833±0.054	5.842±4.591	**18.32**±42.57	3±1
DPIR	30.21±2.279	0.869±0.038	2.720±3.910	105.47±36.21	3±1

Super-resolution : average performance over a test set of 100 images from FFHQ.

Convergence Property

Deblurring Inpainting Super-resolution



Auto-correlation functions for the three problems.

Conclusion

- Summary : the proposed method
  - ✓ Combines the advantages offered by the Bayesian and the deep NN frameworks.
    1. Use of deep neural networks data-driven priors.
    2. Uncertainty quantification.
  - ✓ PSNR performances similar to optimization-based algorithms while scaling efficiently to high dimension.
  - ✓ Theoretical guarantees of the proposed MCMC scheme (convergence guarantee, bias quantification).
- Ongoing work :
  - Extension of the proposed framework and its theoretical analysis to non-linear inverse problems.
  - Applications to medical imaging.

For more information :

- E.C. Faye, M.D. Fall and N. Dobigeon *"Regularization by denoising : Bayesian model and Langevin-within-split Gibbs sampling"*, IEEE Transactions on Image Processing, vol 34, pages 221-234, 2024.

Generalization of the proposed sampling scheme to various data-driven priors. Application on real data sets :

- E.C. Faye, M.D. Fall, S. Delchini and N. Dobigeon*"Bridging data-driven priors via the score function in Bayesian inverse problems - Efficient Monte Carlo sampling and comparative study"*, Submitted.

Thank you for your attention !