

Approches rapides du bayésien variationnel pour des problèmes de grande dimension

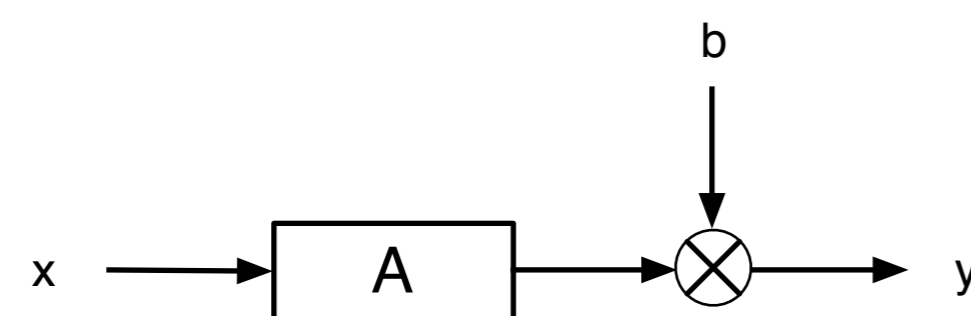
Yuling ZHENG, Thomas RODET, Aurélia FRAYSSE

Laboratoire des Signaux et Systèmes
CNRS, Université Paris Sud, Supélec, 3 rue Joliot Curie, 91190 Gif Sur Yvette, France

Résumé : Nous introduisons ici deux algorithmes d'estimation non supervisés basés sur la méthodologie bayésienne variationnelle. Nous montrons aussi l'application pratique de ces algorithmes sur un problème de super résolution d'images en utilisant un *a priori* TV.

Problème inverse mal posé

▷ **Modèle direct :**



$$y = Ax + b \quad (1)$$

◊ L'objectif : estimer $x \Leftarrow y$: problème inverse, généralement **mal posé** car A : non inversible ou/et mal-conditionnée

▷ **Approches bayésiennes non-supervisées :** loi *a posteriori* compliquée

▷ **Solutions :**

◊ Markov chain Monte Carlo (MCMC) : **coûteuse en temps de calcul**

◊ **Méthode Bayésienne variationnelle**

Méthodologie bayésienne variationnelle

▷ **Principe :** approximation séparable $q = \prod_i q_i$ de l'*a posteriori*

▷ **Problème d'optimisation à résoudre :**

$$q^{opt} = \arg \max_q \mathcal{F}(q), \quad \text{s.c. } q \text{ p.d.f. séparable}$$

où $\mathcal{F}(q) = \int q(\mathbf{w}) \log \frac{p(\mathbf{y}, \mathbf{w})}{q(\mathbf{w})} d\mathbf{w}$ est concave.

▷ **Solution analytique mais non explicite :** $q(\mathbf{w}) = \prod_i q_i(w_i)$, où

$$q_i(w_i) = \frac{1}{K_i} \exp \left(\langle \log p(\mathbf{y}, \mathbf{w}) \rangle_{\prod_{j \neq i} q_j(w_j)} \right). \quad (2)$$

⇒ Algorithme alterné : pas efficace

Méthodes des sous-espaces dans \mathbb{R}^N [1]

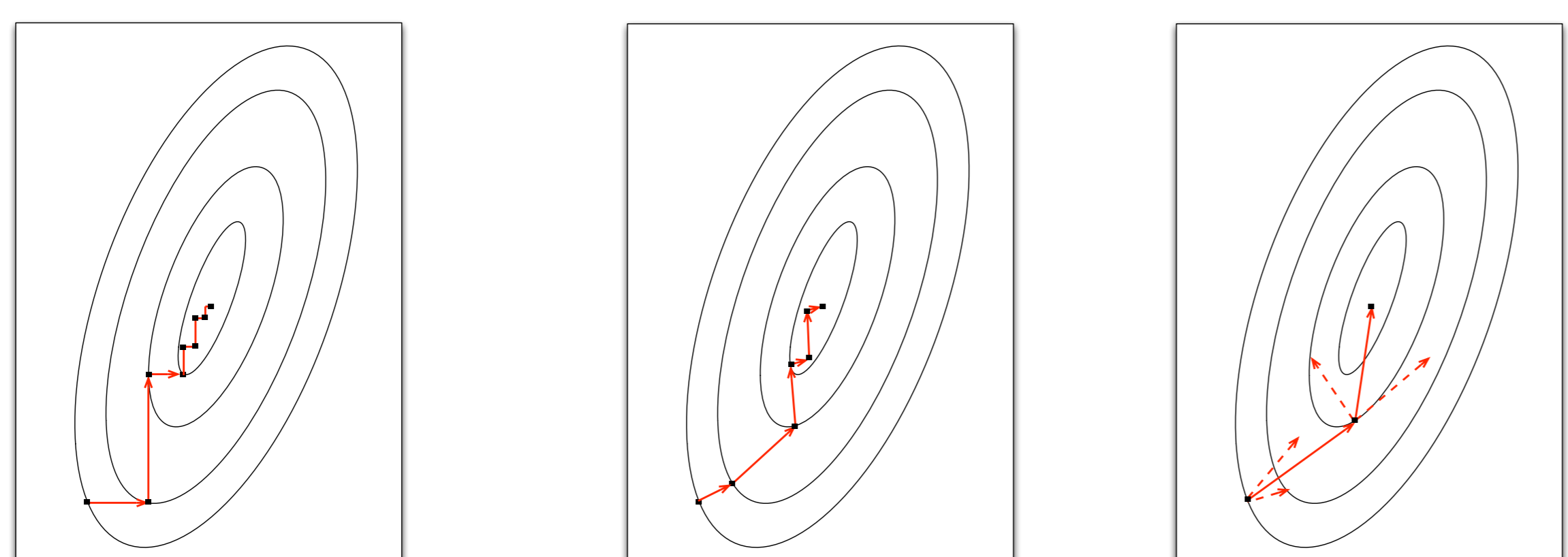
▷ **L'équation de mise à jour :**

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \delta^k = \mathbf{x}^k + \mathbf{D}^k \mathbf{s}^k$$

où $\mathbf{D}^k = [d_1^k, \dots, d_M^k]$: directions définissant le sous-espace,
 $\mathbf{s}^k = [s_1^k, \dots, s_M^k]$: pas suivant chaque direction.

$M = 2$
Pour le compromis entre la rapidité et la complexité

▷ **Comparaison** de trois algorithmes d'optimisation :



(a) Algo alterné

(b) Algo du gradient

(c) Algo du sous-espace

Méthode proposée

▷ **L'équation de mise à jour :**

$$q^{k+1}(\mathbf{x}) = K_k q^k(\mathbf{x}) \exp(\delta^k) \quad \text{où} \quad \delta^k = s_1 d_1^k(\mathbf{x}) + s_2 d_2^k(\mathbf{x}) \quad (3)$$

▷ **Sous-espaces considérés :**

$$SG : d_1^k(\mathbf{x}) = df(q^k, \mathbf{x}), \quad d_2^k(\mathbf{x}) = df(q^{k-1}, \mathbf{x}), \quad (4)$$

$$GM : d_1^k(\mathbf{x}) = df(q^k, \mathbf{x}), \quad d_2^k(\mathbf{x}) = \delta^{k-1}(\mathbf{x}), \quad (5)$$

où $df(q^k, \mathbf{x}), df(q^{k-1}, \mathbf{x})$: gradients

$\delta^{k-1}(\mathbf{x})$: direction précédente.

⇒ $\mathbf{s} = [s_1, s_2]$, pas de l'algorithme

▷ **Détermination du pas bidimensionnel** à l'aide du développement de Taylor d'ordre deux

Application à la super-résolution

▷ **L'objectif :** Reconstruire une image haute-résolution \Leftarrow une série d'images basse-résolution

▷ Utilisation du modèle (1), bruit gaussien : $b \sim \mathcal{N}(\mathbf{0}, \gamma_b^{-1} \mathbf{I})$

▷ Loi *a priori* pour \mathbf{x} : variation totale

$$p(\mathbf{x}|\gamma_p) \approx \tilde{p}(\mathbf{x}|\gamma_p) = c \gamma_p^{N/2} \exp \left[-\gamma_p \sum_{i=1}^N \sqrt{(\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2} \right],$$

où Δ_i^h, Δ_i^v : les différences d'ordre un (horizontale, verticale)

▷ Loi *a priori* pour les hyperparamètres γ_b et γ_p : Jeffreys

▷ Loi *a posteriori*

$$p(\mathbf{x}, \gamma_b, \gamma_p | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{x}, \gamma_b) \tilde{p}(\mathbf{x} | \gamma_p) p(\gamma_b) p(\gamma_p)$$

Compliquée ⇒ Méthode Bayésienne variationnelle

Difficulté liée au TV pour utiliser le bayésien variationnel

▷ **Critère incalculable**

Solution : Minoration-Maximisation (MM)

$$\tilde{p}(\mathbf{x}|\gamma_p) \geq M(\mathbf{x}, \gamma_p | \boldsymbol{\lambda}) = c \gamma_p^{N/2} \exp \left[-\gamma_p \sum_{i=1}^N \frac{(\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2 + \lambda_i}{2\sqrt{\lambda_i}} \right], \quad (6)$$

où $(\lambda_i)_{i=1, \dots, N}$: variables auxiliaires positives

⇒ \mathcal{F}^L : minoration de l'énergie libre négative

▷ **Algorithme alterné** pour maximiser \mathcal{F}^L :

1. Maximiser la minoration \mathcal{F}^L par rapport à q

2. Maximiser la minoration \mathcal{F}^L par rapport à $(\lambda_i)_{i=1, \dots, N}$

Algorithme proposé

▷ **L'objectif :** déterminer l'approximation optimale totalement séparable

$$q(\mathbf{x}, \gamma_b, \gamma_p) = \prod_i q_i(\mathbf{x}_i) q_b(\gamma_b) q_p(\gamma_p). \quad (7)$$

1. Initialiser $(q_i^0)_{i=1, \dots, N}, q_b^0, q_p^0$ et $(\lambda_i^0)_{i=1, \dots, N}$

2. Mettre à jour les paramètres de q_i^{k+1} pour $i = 1, \dots, N$

a. Déterminer les sous-espaces selon (4) ou (5)

b. Calculer le pas bidimensionnel sous-optimal

c. Mettre à jour les paramètres de q_i^{k+1} en utilisant (3)

3. Mettre à jour les variables auxiliaires $(\lambda_i^{k+1})_{i=1, \dots, N}$

4. Mettre à jour les paramètres des lois des hyperparamètres : q_b^{k+1} et q_p^{k+1} en utilisant l'algorithme bayésien variationnel classique (2)

5. Retourner à 2 jusqu'à convergence

Simulations et Résultats

▷ **Simulations** des images basse-résolution :

◊ noyau de convolution uniforme de taille 3×3

◊ décimation d'un facteur 4

◊ niveau de bruit : SNR = 5, 25, 45 dB.

▷ **Résultats :**

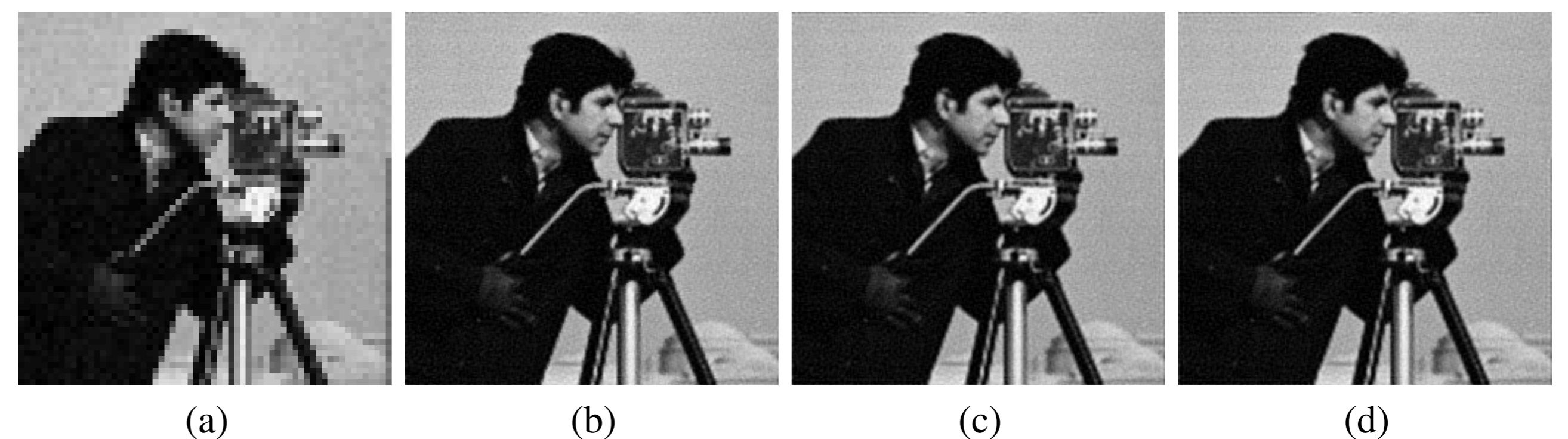


Fig. 1 : (a) Une des images basse-résolution (64×64 , SNR = 25dB), et les images haute-résolution (256×256) reconstruites par (b) Babacan [2], (c) SG, (d) GM.

Reconstructions similaires !!

Table 1 : NOMBRE D'ITÉRATIONS/TEMPS CPU(S).

Données	PSNR	Babacan	Gradient	SG	GM	
Camera-man	5dB	11.64	30/15.7	169/10.3	111/9.3	64/5.5
	25dB	30.59	15/14.9	103/6.1	71/5.9	49/4.0
	45dB	40.62	31/86.8	205/11.8	139/11.3	91/7.3
Lena	5dB	14.81	26/56.6	210/47.9	134/43.1	68/21.7
	25dB	33.42	12/46.8	106/25.2	60/20.0	60/19.4
	45dB	38.72	29/296.4	298/72.6	201/67.8	104/33.7

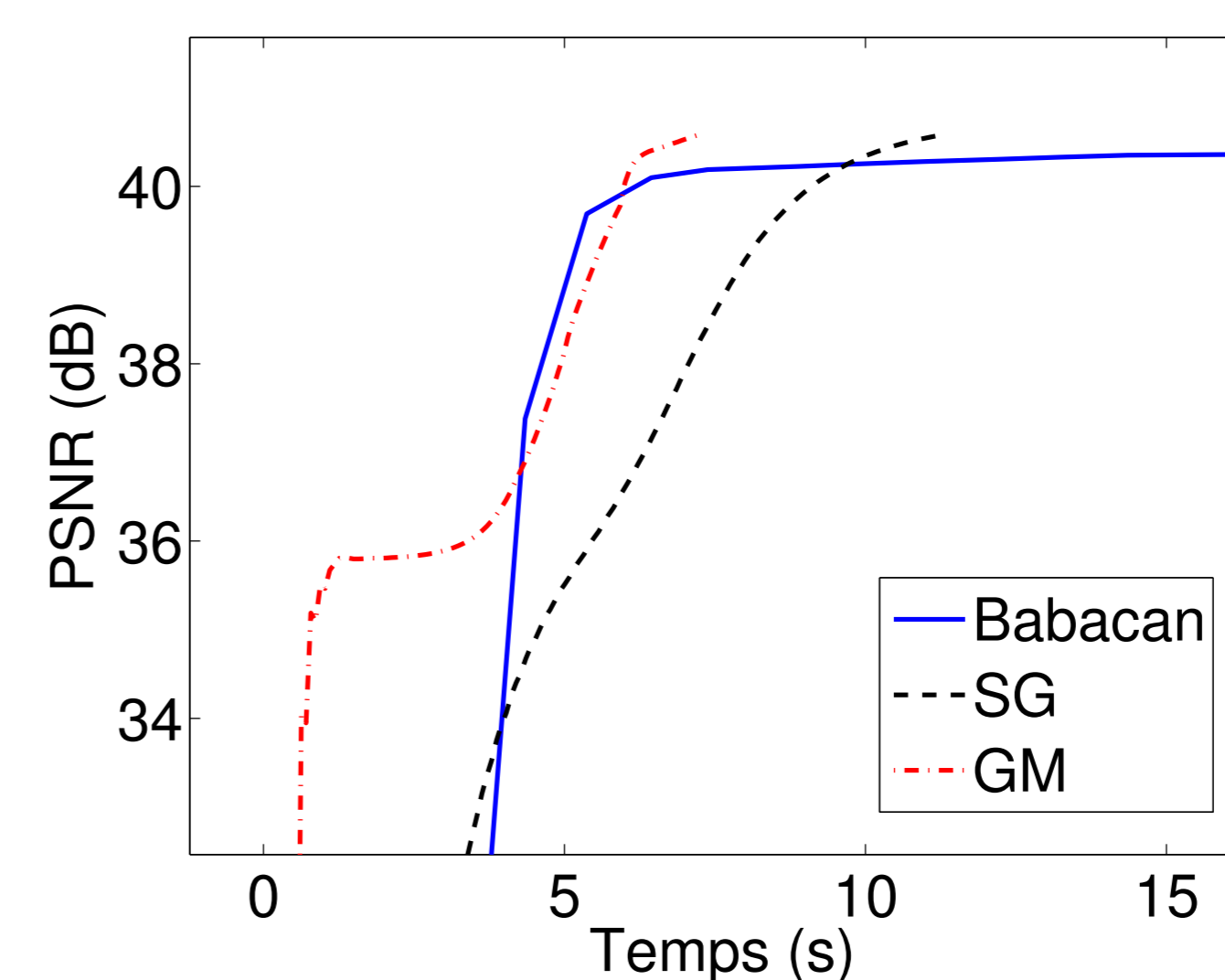


Fig. 2 : Cameraman pour un SNR = 45dB

▷ **Conclusions :**

– Nos deux approches convergent plus vite que celle de [2]

– Notre approche utilisant GM est plus efficace que celle utilisant SG

1. REFERENCES

[1] E. Chouzenoux, J. Idier, and S. Moussaoui, "A Majorize-Minimize strategy for subspace optimization applied to image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 18, pp. 1517–1528, 2011.

[2] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 984–999, 2011.