Méthode MCMC plug-and-play avec a priori génératif profond

Florentin COEURDOUX[1] Nicolas DOBIGEON[2] Pierre CHAINAIS[2]
[1]IRIT, INP-ENSEEIHT, 31071 Toulouse, France
[2]Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, 59000 Lille, France

Résumé – Les méthodes Plug-and-Play (PnP) forment une classe d'algorithmes itératifs pour la résolution de problèmes inverses dans lesquels la régularisation est assurée par le recours à un débruiteur générique. Bien que produisant de très bons résultats, ces méthodes d'optimisation PnP ne produisent souvent que des estimateurs ponctuels et non une caractérisation complète de la distribution a posteriori. Nous proposons une nouvelle famille de méthodes PnP stochastiques en exploitant un algorithme d'échantillonnage de Gibbs couplé à un réseau de neurones génératif. En plus d'un estimateur ponctuel, l'approche proposée fournit des intervalles de confiance pour un coût de calcul très modéré. L'efficacité des échantillonneurs proposés est évaluée grâce à des simulations de problèmes de reconstruction d'image. Les performances des estimateurs proposés se comparent favorablement à celles des algorithmes d'optimisation et MCMC les plus récents.

Abstract – Plug-and-Play (PnP) methods are a class of iterative algorithms for inverse problems solving in which regularization is provided by a generic denoiser. Although producing very good results, these PnP optimization methods produce only point estimators and not a complete characterization of the posterior distribution. We propose a new family of stochastic PnP methods by exploiting a Gibbs sampling algorithm coupled with a generative neural network. In addition to a point estimator, the proposed approach provides confidence intervals for a moderate computational cost. The efficiency of the proposed samplers is evaluated through simulations of image reconstruction problems. The performances of the proposed estimators compare favorably with those of the most recent optimization and MCMC algorithms

1 Introduction

Les problèmes inverses sont généralement formulés comme une tâche de déduction d'un objet d'intérêt inconnu $\mathbf{x} \in \mathbb{R}^N$ à partir d'une mesure bruitée $\mathbf{y} \in \mathbb{R}^M$. L'inversion est alors formulée généralement comme un problème d'optimisation

$$\min_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}) + g(\mathbf{x}) \qquad (1)$$

où $f(\cdot, \mathbf{y})$ désigne l'attache à la donnée observée $\mathbf{y}$, et $g(\cdot)$ est un terme de régularisation résumant la connaissance a priori sur $\mathbf{x}$. La nature mal posée ou mal conditionnée du problème implique qu'il n'est souvent pas possible de retrouver $\mathbf{x}$ de manière unique et stable à partir de la seule observation $\mathbf{y}$. De nombreux travaux ont porté sur la tâche complexe de la conception d'une régularisation pertinente et efficace telle que la variation totale ou la parcimonie d'une représentation. Certaines méthodes proposent d'entraîner directement un réseau de neurones sur une grande base d'images de référence pour résoudre directement le problème d'inversion ou être utilisé comme régularisation.

Le domaine évolue actuellement vers des méthodes qui combinent ces réseaux avec des algorithmes d'optimisation traditionnels, ce qui permet d'améliorer leur interprétabilité et leur portée générale. L'intégration de ces méthodes a conduit à la méthode Plug-and-Play (PnP) [14], qui remplace le problème d'optimisation par deux sous-problèmes plus simples. À l'instar de la méthode des directions alternées des multiplicateurs (ADMM) [1], les méthodes PnP reposent sur une stratégie de séparation des variables. Il est ainsi possible de traiter séparément le terme d'attache aux données et le terme de régularisation. Lors de l'inférence, la régularisation intervient en pratique dans un sous-problème qui se traduit par une tâche de débruitage. Cette tâche peut être confiée à un débruiteur dédié, pouvant recourir à l'apprentissage profond. L'efficacité des méthodes PnP en fait une référence qui a permis d'atteindre des performances de pointe dans une grande variété d'applications [18].

Malgré les progrès réalisés, les algorithmes PnP actuels ne produisent qu'un estimateur ponctuel et non une information complète sur la distribution a posteriori de la variable d'intérêt $\mathbf{x}$ sachant les observations $\mathbf{y}$. Dans un cadre d'inférence bayésienne, les algorithmes MCMC ont l'avantage de fournir une description complète de la distribution a posteriori. Cependant les méthodes MCMC peuvent devenir particulièrement coûteuses en grande dimension. Pour surmonter cette limitation, plusieurs méthodes ont vu le jour, notamment l'algorithme split-Gibbs sampling (SGS) [15]. Inspiré des stratégies de séparation des variables usuelles en optimisation, l'algorithme SGS permet de traiter séparément les termes associés à la vraisemblance et à la loi a priori. L'échantillonnage selon la loi conditionnelle a posteriori définie par la loi a priori peut alors être interprété comme une tâche de débruitage stochastique.

Les performances des modèles génératifs profonds en font aujourd'hui des candidats de choix pour remplacer les débruiteurs déterministes des méthodes d'optimisation PnP. Les réseaux antagonistes (GAN), les auto-encodeurs variationnels (VAE) et récemment les modèles de diffusion (DDPM) [12] ont démontré des performances remarquables sur des tâches de débruitage [13, 11].

Nous proposons une méthode d'échantillonnage PnP qui

Ce travail a été soutenu par Artificial Natural Intelligence Toulouse Institute (ANITI, ANR-19-PI3A-0004), la Chaire IA Sherlock (ANR-20-CHIA-0031-01), le programme national d'investissement d'avenir ULNE (ANR-16-IDEX-0004) et la Région Hauts-de-France.

tire parti à la fois des performances des modèles génératifs profonds et de la versatilité de l'algorithme SGS. Cette méthode s'appuie sur une loi a priori qui n'est jamais explicitement spécifiée mais implicitement encodée dans un modèle génératif préalablement entraîné. Les résultats de restauration obtenus sont d'une qualité comparable à celle des méthodes d'optimisation de l'état de l'art. Ils fournissent de surcroît des intervalles de crédibilité contrôlés, le tout pour un coût de calcul modéré.

La partie 2 présente l'algorithme PnP-SGS, instancié dans le cas particulier d'un modèle DDPM [12]. La partie 3 illustre l'approche proposée pour la résolution d'une tâche de reconstruction d'image. Enfin, la section 3.2 présente les résultats expérimentaux avant de conclure.

2 PnP-SGS : une variante stochastique du PnP

2.1 Algorithme SGS

La formulation bayésienne du problème inverse associé à (1) revient à considérer une loi a posteriori

$$\pi(\mathbf{x}) \triangleq p(\mathbf{x} \mid \mathbf{y}) \propto \exp[-f(\mathbf{x}, \mathbf{y}) - g(\mathbf{x})]. \tag{2}$$

Les termes $\exp[-f(\cdot, \mathbf{y})]$ et $\exp[-g(\cdot)]$ correspondent respectivement à la vraisemblance et à la loi priori, aux constantes de normalisation près. La méthode split-Gibbs sampling (SGS) [15] consiste à introduire une variable auxiliaire $\mathbf{z}$ pour définir un modèle augmenté dont la loi a posteriori s'écrit

$$\begin{aligned} \pi_\rho(\mathbf{x}, \mathbf{z}) &\triangleq p(\mathbf{x}, \mathbf{z} \mid \mathbf{y}) \\ &\propto \exp\left[-f(\mathbf{x}, \mathbf{y}) - g(\mathbf{z}) - \frac{\|\mathbf{x} - \mathbf{z}\|_2^2}{2\rho^2}\right] \end{aligned} \tag{3}$$

où ρ est un paramètre positif contrôlant la dissimilarité entre $\mathbf{x}$ et $\mathbf{z}$. Dans la limite $\rho \to 0$, la distribution marginale de $\mathbf{x}$ sous π_ρ dans (3) coïncide avec la distribution cible π dans (2), caractérisant un modèle augmenté asymptotiquement exact (AXDA) [16]. L'algorithme SGS échantillonne selon cette loi jointe grâce à un algorithme de Gibbs de manière alternée. Les lois conditionnelles a posteriori s'écrivent alors

$$p(\mathbf{x} \mid \mathbf{y}, \mathbf{z}) \propto \exp\left[-f(\mathbf{y}, \mathbf{x}) - \frac{1}{2\rho^2}\|\mathbf{x} - \mathbf{z}\|_2^2\right] \tag{4}$$

$$p(\mathbf{z} \mid \mathbf{x}) \propto \exp\left[-g(\mathbf{z}) - \frac{1}{2\rho^2}\|\mathbf{x} - \mathbf{z}\|_2^2\right] \tag{5}$$

Cette séparation des variables permet de dissocier f et g, ce qui conduit à un algorithme d'échantillonnage plus simple et plus efficace. Dans le cas d'une fonction $f(\cdot, \mathbf{y})$ quadratique, l'étape (4) consiste à échantillonner selon une loi normale [17]. Pour d'autres types de potentiels associés à un bruit de mesure non-gaussien, d'autres stratégies peuvent être considérées, voir [7]. Il est aisé de remarquer que (5) correspond à la loi a posteriori associée à une tâche de débruitage bayésien dont l'objectif serait d'estimer la quantité $\mathbf{z}$ à partir d'une observation bruitée $\mathbf{x}$ entachée d'un bruit additif gaussien. Nous proposons de réaliser cet échantillonnage à l'aide d'un modèle génératif profond utilisé comme débruiteur PnP. A titre d'illustration, nous décrivons par la suite l'utilisation pratique d'un modèle de diffusion (DDPM) [11]. Cependant, il est important de souligner que le cadre proposé reste très général et permet de recourir à un modèle génératif quelconque pré-entraîné.

2.2 Modèles probabilistes de diffusion par débruitage (DDPM)

DDPM comme modèle génératif – Un DDPM [5] est une architecture neuronale qui construit deux chaînes de Markov : une chaîne directe qui perturbe les données par un bruit et une chaîne rétrograde qui convertit le bruit en données. La première est conçue dans le but de transformer toute loi potentiellement complexe en une loi instrumentale simple, par ex. une loi normale multivariée centrée réduite. La seconde chaîne inverse la première en apprenant un noyau de transition paramétré par un réseau de neurones.

Formellement, étant donné un vecteur de données, supposé sans bruit, noté $\mathbf{x}_0$ de loi $\mathbf{x}_0 \sim q(\mathbf{x}_0)$, la chaîne de Markov directe produit une séquence de variables aléatoires $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T$ de noyau de transition $q(\mathbf{x}_t \mid \mathbf{x}_{t-1})$. Le choix le plus courant pour ce noyau de transition correspond à une perturbation gaussienne définie au pas de temps t par

$$q(\mathbf{x}_t \mid \mathbf{x}_{t-1}) = \mathcal{N}\left(\mathbf{x}_t; \sqrt{1 - b(t)}\mathbf{x}_{t-1}, b(t)\mathbf{I}\right) \tag{6}$$

où $t \mapsto b(t)$ est une fonction continue strictement croissante, typiquement une fonction linéaire. Ce processus direct injecte lentement du bruit dans les données jusqu'à ce que toutes les structures soient perdues. La chaîne rétrograde est paramétrée par une loi instrumentale choisie comme $p(\mathbf{x}_T) = \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I})$ et un noyau de transition $p_{\boldsymbol{\theta}}(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$. Le noyau de transition appris $p_{\boldsymbol{\theta}}(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$ prend la forme de

$$p_{\boldsymbol{\theta}}(\mathbf{x}_{t-1} \mid \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_{\boldsymbol{\theta}}(\mathbf{x}_t, t), \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)) \tag{7}$$

où la moyenne $\boldsymbol{\mu}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)$ et la matrice de covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)$ sont encodées par un réseau de neurones de paramètres $\boldsymbol{\theta}$. Ces paramètres sont ajustés en entraînant le réseau à éliminer progressivement le bruit pour cibler la donnée $\mathbf{x}_0$ grâce à la chaîne rétrograde. Pour générer de nouveaux échantillons, un DDPM entraîné commence par générer un vecteur de bruit $\mathbf{x}_T$ selon la loi instrumentale puis élimine progressivement dans le sens rétrograde le bruit injecté selon le processus induit par la chaîne directe.

DDPM comme débruiteur PnP – Il est important de remarquer qu'un DDPM préalablement entraîné peut être facilement utilisé comme débruiteur PnP stochastique. Contrairement à l'utilisation usuelle d'un DDPM comme générateur ci-dessus, nous commençons le processus de diffusion rétrograde, non pas à partir d'une réalisation de bruit $\mathbf{x}_T$, mais à partir d'une image bruitée $\mathbf{x}_t$ pour un certain t. Pour retrouver $\mathbf{x}_0$, il suffit d'appliquer le processus rétrograde définie par (7) à partir de l'étape correspondant au "temps" t. Ce processus est alors interprété comme t étapes successives de débruitage progressif, t restant un paramètre continu. La seule difficulté est d'alors de déterminer l'instant t, autrement dit le niveau de bruit, à partir duquel la chaîne rétrograde doit être appliquée. Pour cela, il suffit d'exprimer la loi de $\mathbf{x}_t$ sachant $\mathbf{x}_0$. Grâce à la factorisation induite par la chaîne de Markov directe et à la nature gaussienne du noyau de transition (6), la loi de l'image bruitée à l'instant t est

$$q(\mathbf{x}_t \mid \mathbf{x}_0) = \mathcal{N}\left(\mathbf{x}_t; \sqrt{\bar{\alpha}(t)}\mathbf{x}_0, (1 - \bar{\alpha}(t))\mathbf{I}\right) \tag{8}$$

avec $\bar{\alpha}(t) = \prod_{s=1}^{t}(1 - b(s))$. Par définition de la fonction $b(\cdot)$, la fonction $\alpha(\cdot) = 1 - \bar{\alpha}(\cdot)$ définie comme un produit de fonctions continues strictement monotones est inversible. La loi (8) montre que, pour tout t, l'image résultant de la chaîne directe contient un bruit gaussien d'une variance $\alpha(t)$.

2.3 PnP-SGS basé sur un DDPM

Comme indiqué précédemment, la seconde étape (5) de l'algorithme SGS s'interprète comme un débruitage stochastique de la valeur courante de $\mathbf{x}$. Cet échantillonnage selon (5) est ici réalisé grâce à un DPPM préalablement entraîné, conformément à la stratégie introduite au paragraphe précédent. Cette tâche nécessite une estimation du niveau de bruit afin de débuter la diffusion rétrograde (7) à partir d'un pas de temps $\hat{t}$ approprié. Nous proposons d'utiliser ici un estimateur $\hat{\sigma}^2$ usuel [6] de la variance du bruit contenue dans $\mathbf{x}$. L'instant $\hat{t}$ est alors estimé par inversion de la fonction de variance en utilisant $\hat{t} = \alpha^{-1}(\hat{\sigma})$. Plus de détails sont disponibles dans la version étendue de cet article [4].

3 PnP-SGS pour l'inpainting d'image

Pour illustrer l'application de l'algorithme PnP-SGS, cette partie considère un problème d'*inpainting* d'image. Le modèle est rappelé dans le paragraphe 3.1 et les performances attendues sont évaluées dans le paragraphe 3.2.

3.1 Formulation du problème

Une image observée $\mathbf{y} \in \mathbb{R}^M$ est issue de la dégradation d'une image inconnue $\mathbf{x} \in \mathbb{R}^N$ selon le modèle linéaire

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}. \tag{9}$$

On suppose ici que le bruit est gaussien de matrice de covariance $\sigma^2 \mathbf{I}_M$. La matrice $\mathbf{H}$ représente un opérateur de sous-échantillonnage irrégulier associé à un masque binaire ($M \ll N$). Dans ce cadre, la loi a posteriori de $\mathbf{x}$ (2) s'écrit

$$p(\mathbf{x} \mid \mathbf{y}) \propto \exp\left[-\frac{1}{2\sigma^2}\|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 - g(\mathbf{x})\right]. \tag{10}$$

Lorsque la fonction de régularisation $g(\cdot)$ est spécifiée, l'échantillonnage direct selon cette distribution peut être une tâche difficile en raison principalement de *i)* la dimension généralement élevée de l'image $\mathbf{x}$ à restaurer, *ii)* la non-conjugaison de la loi a priori choisi (TV, représentation parcimonieuse...), conduisant à une loi a posteriori non standard et *iii)* la non-différentiabilité de $g(\cdot)$ qui empêche l'utilisation de certaines techniques d'échantillonnage efficace (e.g., Hamiltonien Monte Carlo, diffusion de Langevin). Pour contourner ces problèmes, l'algorithme PnP-SGS cible la loi jointe (3) associée à un modèle AXDA. En particulier, la loi conditionnelle a posteriori (4) est la loi normale multivariée

$$p(\mathbf{x} \mid \mathbf{y}, \mathbf{z}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{11}$$

avec

$$\begin{cases} \boldsymbol{\Sigma} = \left(\frac{1}{\sigma^2}\mathbf{H}^T\mathbf{H} + \frac{1}{\rho^2}\mathbf{I}_N\right)^{-1} \\ \boldsymbol{\mu} = \boldsymbol{\Sigma}^{-1}\left(\frac{1}{\sigma^2}\mathbf{H}^T\mathbf{y} + \frac{1}{\rho^2}\mathbf{z}\right) \end{cases} \tag{12}$$

Malgré la grande dimensionalité du problème, la structure de l'opérateur $\mathbf{H}$ permet un échantillonnage efficace selon cette loi normale multivariée après application de la formule de Sherman-Morrison-Woodbury

$$\left(\frac{1}{\sigma^2}\mathbf{H}^T\mathbf{H} + \frac{1}{\rho^2}\mathbf{I}_N\right)^{-1} = \rho^2\left(\mathbf{I}_N - \frac{\rho^2}{\sigma^2 + \rho^2}\mathbf{H}^T\mathbf{H}\right).$$

Ainsi, la matrice de covariance $\boldsymbol{\Sigma}$ est diagonale et l'échantillonnage à partir de (11) peut être effectué efficacement en utilisant par exemple l'algorithme E-PO [10]. La loi (5) de la variable auxiliaire $\mathbf{z}$ est échantillonnée grâce au DDPM selon la méthode décrite en partie 2.3.

TABLE 1 : Evaluation quantitative (FID, LPIPS, PSNR, SSIM) de solutions au problème d'inpainting sur 1000 images FFHQ 256×256. **Gras** : meilleur, souligné : second

	FID ↓	LPIPS ↓	PSNR ↑	SSIM ↑
PnP-SGS	**38.36**	**0.144**	**23.59**	**0.813**
TV-SGS [15]	71.12	0.785	21.09	0.524
PnP-ADMM [2]	123.61	0.692	<u>22.41</u>	0.325
TV-ADMM	181.56	0.463	22.03	<u>0.784</u>
Score-SDE [11]	76.54	0.612	13.52	0.437
DDRM [9]	69.71	0.587	9.19	0.319
MCG [3]	<u>39.26</u>	<u>0.286</u>	21.57	0.751

3.2 Résultats expérimentaux

La méthode proposée est évaluée sur un ensemble de 1000 images de taille 256×256 de la base FFHQ [8]. Dans ces expériences, nous utilisons directement le modèle de diffusion DDPM préalablement entraîné fourni par les auteurs de [5], sans ajustement spécifique. Les 1000 images de test n'ont jamais été vues par le modèle pendant l'entraînement afin d'éviter tout biais dû à un modèle potentiellement sur-ajusté. L'opérateur $\mathbf{H}$ est construit de sorte que 80% des pixels de chaque image sont masqués aléatoirement. Un bruit blanc gaussien d'écart-type $\sigma = 0.05$ est ajouté à chaque observation. L'hyperparamètre $\rho = 0.7$ a été ajusté pour fournir les meilleurs résultats.

Nous effectuons une comparaison avec les méthodes suivantes : un algorithme SGS avec une régularisation TV (TV-SGS) [15], son équivalent déterministe TV-ADMM, un algorithme de type PnP-ADMM [2] utilisant un débruiteur convolutionnel (DnCNN) [19], DDRM [9], MCG [3] et Score-SDE [11]. En particulier, DDRM, MCG et Score-SDE sont des méthodes qui exploitent le même réseau DDRM pré-entrainé que celui utilisé par notre méthode PnP-SGS. Notons que TV-ADMM, PnP-ADMM, DDRM, MCG et Score-SDE sont des méthodes basées sur des schémas d'optimisation qui ne fournissent qu'une estimation ponctuelle de la solution. Au contraire, la méthode proposée PnP-SGS, tout comme TV-SGS, échantillonne la loi a posteriori et fournit non seulement des estimateurs ponctuels (MMSE) mais aussi des variances d'estimation et plus généralement des intervalles de crédibilité.

Quatre mesures de qualité sont utilisées pour la comparaison quantitative des méthodes : rapport signal-sur-bruit (PSNR) et indice de similarité structurelle (SSIM), qui sont des mesures standard de qualité d'image, et distance Inception de Fréchet (FID) et distance de similarité par patch perceptuel (LPIPS), qui sont deux mesures perceptuelles.

FIGURE 1 : De gauche à droite : image originale, image masquée et bruitée, résultats de PnP-ADMM, PnP-SGS et intervalles de crédibilité à 90% pour PnP-SGS

Le tableau 1 montre que la méthode proposée surpasse toutes les autres méthodes comparées avec des marges significatives sur les mesures de perception visuelle et le SSIM. La Fig. 1 compare spécifiquement les résultats obtenus par la méthode proposée à ceux obtenus par PnP-ADMM. Notre méthode est capable de fournir des reconstructions de haute qualité, nettes et réalistes, ainsi que des intervalles de crédibilité, précieux pour de nombreuses applications.

4 Conclusion

Cet article introduit un équivalent stochastique des approches de type PnP-ADMM pour résoudre des problèmes inverses. Cet équivalent, baptisé Plug-and-Play Split Gibbs Sampler (PnP-SGS), consiste à échantillonner une loi a posteriori augmentée qui s'appuie sur une stratégie de séparation de variables. Il hérite alors de l'intérêt principal de PnP-ADMM qui réside en la non-spécification explicite une régularisation ou, de manière équivalente, un loi a priori. Au contraire, une stratégie de séparation de variables permet d'identifier une sous-tâche de débruitage stochastique résolue par un modèle génératif profond. Au delà d'une seule estimation ponctuelle, telle que réalisée par PnP-ADMM, la méthode proposée fournit une description complète de la loi a posteriori, donnant accès à des intervalles de confiance sur les paramètres estimés. Des expériences approfondies montrent que la méthode proposée rivalise favorablement avec les modèles de l'état-de-l'art.

Références

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.* : Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine learning*, 3(1):1–122, 2011.

[2] S. H. Chan, X. Wang et O. A. Elgendy : Plug-and-play ADMM for image restoration : Fixed-point convergence and applications. *IEEE Trans. Comput. Imag.*, 3(1):84–98, 2016.

[3] H. Chung, B. Sim, D. Ryu et J. C. Ye : Improving Diffusion Models for Inverse Problems using Manifold Constraints. *In* A. H. Oh, A. Agarwal, D. Belgrave et K. Cho, éditeurs : *Adv. in Neural Information Process. Systems (NIPS)*, 2022.

[4] F. Coeurdoux, N. Dobigeon et P. Chainais : Plug-and-play split Gibbs sampler : embedding deep generative priors in Bayesian inference. *arXiv preprint arXiv :2304.11134*, 2023.

[5] P. Dhariwal et A. Nichol : Diffusion models beat gans on image synthesis. *In Adv. in Neural Information Process. Systems (NIPS)*, volume 34, pages 8780–8794, 2021.

[6] D. L. Donoho et J. M. Johnstone : Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[7] A. Durmus, E. Moulines et M. Pereyra : Efficient Bayesian computation by proximal markov chain monte carlo : when langevin meets moreau. *SIAM J. Imag. Sci.*, 11(1):473–506, 2018.

[8] T. Karras, S. Laine et T. Aila : A style-based generator architecture for generative adversarial networks. *In Proc. Int. Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 4401–4410, 2019.

[9] B. Kawar, M. Elad, S. Ermon et J. Song : Denoising diffusion restoration models. *arXiv preprint arXiv :2201.11793*, 2022.

[10] Y. Marnissi, E. Chouzenoux, A. Benazza-Benyahia et J.-C. Pesquet : An Auxiliary Variable Method for Markov Chain Monte Carlo Algorithms in High Dimension. *Entropy*, 20(2), 2018.

[11] J. Song, C. Meng et S. Ermon : Denoising diffusion implicit models. *arXiv preprint arXiv :2010.02502*, 2020.

[12] Y. Song et S. Ermon : Generative modeling by estimating gradients of the data distribution. *In Adv. in Neural Information Process. Systems (NIPS)*, volume 32, 2019.

[13] L. D. Tran, S. M. Nguyen et M. Arai : GAN-based noise model for denoising real images. *In Proc. Asian Conf. Computer Vision*, 2020.

[14] S. V. Venkatakrishnan, C. A. Bouman et B. Wohlberg : Plug-and-play priors for model based reconstruction. *In 2013 IEEE Global Conference on Signal and Information Processing*, pages 945–948. IEEE, 2013.

[15] M. Vono, N. Dobigeon et P. Chainais : Split-and-augmented gibbs sampler—application to large-scale inference problems. *IEEE Trans. Signal Process.*, 67(6):1648–1661, 2019.

[16] M. Vono, N. Dobigeon et P. Chainais : Asymptotically exact data augmentation : Models, properties, and algorithms. *Journal of Computational and Graphical Statistics*, 30(2):335–348, 2020.

[17] M. Vono, N. Dobigeon et P. Chainais : High-dimensional Gaussian sampling : a review and a unifying approach based on a stochastic proximal point algorithm. *SIAM Review*, 64(1):3–56, 2022.

[18] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool et R. Timofte : Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.

[19] K. Zhang, W. Zuo, Y. Chen, D. Meng et L. Zhang : Beyond a Gaussian denoiser : Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.*, 26(7):3142–3155, 2017.