

CHAPITRE 3

Parcimonie, problèmes inverses et acquisition comprimée

9 septembre 2026

Comment décrire un signal complexe à partir d'un petit nombre de mesures ? Les représentations parcimonieuses relient la compression à la reconstruction de données incomplètes ou bruitées. Ce chapitre développe ce lien, explique le rôle de l'optimisation convexe et précise les hypothèses qui permettent de garantir une reconstruction par acquisition comprimée.

3.1 Du signal continu aux échantillons

Un dispositif d'acquisition transforme un signal continu $\tilde{f}$ en un vecteur fini $f = (f_q)_{q=1}^Q \in \mathbb{R}^Q$. Pour un son, ces valeurs correspondent à des mesures effectuées à des instants successifs. Pour une image en niveaux de gris, elles représentent la lumière reçue par de petites surfaces du capteur. Une mesure de pixel idéalisée s'écrit

$$f_q = \int_{c_q} \tilde{f}(s) ds,$$

où c_q est la surface de l'élément sensible. Diviser par son aire donne une intensité moyenne ; le temps d'exposition et l'étalonnage introduisent d'autres facteurs d'échelle. Les signaux en couleur ou multicanaux se décrivent en enregistrant plusieurs valeurs à chaque position.

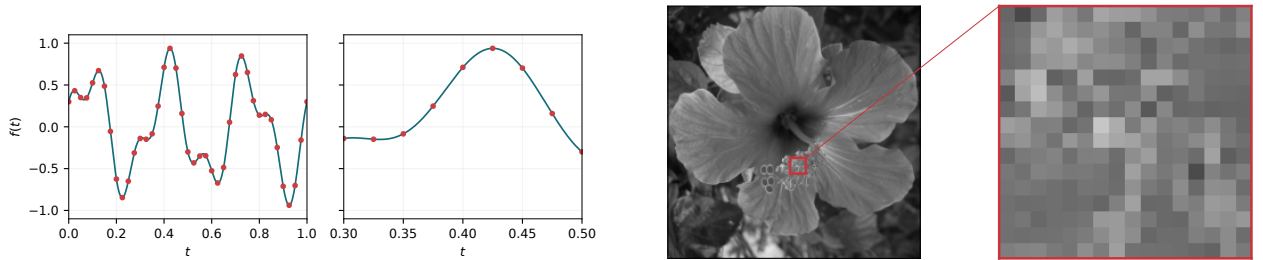


FIGURE 3.1. Échantillonnage d'un signal unidimensionnel et d'une image. Les points indiquent les échantillons temporels ; les cadres rouges repèrent la région de l'image agrandie à droite.

L'échantillonnage préserve l'information sous des hypothèses appropriées. Supposons par exemple que $\tilde{f} \in L^2(\mathbb{R})$ ait une transformée de Fourier à support dans $[-B, B]$, la fréquence étant mesurée en cycles par unité de temps. Si le pas vérifie $0 < T < 1/(2B)$, le théorème d'échantillonnage [15] donne

$$\tilde{f}(t) = \sum_{k \in \mathbb{Z}} \tilde{f}(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right), \quad \operatorname{sinc}(u) = \begin{cases} \sin(\pi u)/(\pi u), & u \neq 0, \\ 1, & u = 0. \end{cases}$$

Pour le représentant continu de cette fonction à bande limitée, la série converge dans L^2 et uniformément. Le théorème utilise les échantillons sur tout l'axe temporel. Un enregistrement fini, la quantification et le bruit de

mesure introduisent des erreurs supplémentaires. Les fréquences dépassant la bande d'acquisition doivent être atténuées avant l'échantillonnage pour éviter le repliement spectral; augmenter ensuite la précision numérique ne supprime pas ce repliement.

L'échantillonnage concerne la représentation d'un signal continu par des mesures, tandis que le codage sans perte concerne la représentation binaire de ces mesures. La parcimonie fournit une autre hypothèse utile sur le signal et conduit à une autre stratégie d'acquisition.

3.2 Approximation non linéaire et compression

3.2.1 Représentations parcimonieuses

Un dictionnaire $\Psi = (\psi_n)_{n=1}^N$ est une famille de vecteurs $\psi_n \in \mathbb{R}^Q$. On cherche une approximation

$$f \approx \Psi x = \sum_{n=1}^N x_n \psi_n$$

comportant peu de coefficients non nuls. Les sinusoides, les cosinus locaux et les ondelettes sont des exemples d'atomes. Le codage par transformée utilise ces représentations pour concentrer une grande partie de l'énergie du signal dans un petit nombre de coefficients.

Pour un budget entier $0 \leq M \leq N$, une meilleure approximation à M termes résout

$$x^* \in \arg \min_{x \in \mathbb{R}^N, \|x\|_0 \leq M} \|f - \Psi x\|_2^2, \quad \|x\|_0 = \#\{n : x_n \neq 0\}. \quad (3.1)$$

On note ici $\|v\|_2^2 = \sum_n |v_n|^2$. La notation $\|x\|_0$ est conventionnelle : ce comptage n'est pas une norme, car multiplier un vecteur non nul par un scalaire non nul ne change pas le nombre de ses composantes non nulles. Pour un dictionnaire général, l'approximation parcimonieuse est un problème NP-difficile [12]. Tester tous les supports possibles est généralement prohibitif; cette difficulté dans le pire cas n'exclut pas des algorithmes efficaces pour certains dictionnaires ou certaines données.

3.2.2 Le cas orthonormal

Supposons $N = Q$ et Ψ base orthonormale, c'est-à-dire $\langle \psi_n, \psi_m \rangle = \delta_{nm}$. En posant $a_n = \langle f, \psi_n \rangle$, l'identité de Parseval donne

$$f = \sum_{n=1}^N a_n \psi_n, \quad \|f - \Psi x\|_2^2 = \sum_{n=1}^N |a_n - x_n|^2. \quad (3.2)$$

Une fois le support choisi, les meilleurs coefficients sont les a_n correspondants. Pour minimiser l'énergie éliminée, on conserve les M plus grandes amplitudes. Plus précisément, choisissons un ordre $|a_{n_1}| \geq \dots \geq |a_{n_N}|$. Alors

$$x_n^* = \begin{cases} a_n, & n \in \{n_1, \dots, n_M\}, \\ 0, & \text{sinon.} \end{cases} \quad (3.3)$$

Des égalités d'amplitude au niveau du seuil peuvent donner plusieurs minimiseurs. S'il y a moins de M coefficients non nuls, l'approximation utilise moins de M termes.

La reconstruction $x \mapsto \Psi x$ est linéaire, mais la sélection des plus grands coefficients dépend des données. L'approximation $f \mapsto \Psi x^*$ est donc non linéaire. Pour de nombreux signaux réguliers par morceaux, les ondelettes donnent de petites erreurs avec relativement peu de coefficients; la régularité du signal contrôle la décroissance de cette erreur [10, chapitre 9]. Les textures et les oscillations fines peuvent demander bien davantage de termes.

Le rapport M/N compte les coefficients conservés; ce n'est pas un débit binaire. Un compresseur complet doit également quantifier leurs valeurs, décrire leurs positions et coder les symboles obtenus. Le codage entropique, introduit au chapitre consacré à la théorie de l'information, complète cette étape. Les transformées en ondelettes disposent en outre d'algorithmes rapides dont le coût est proportionnel au nombre d'échantillons [10, chapitre 7].



FIGURE 3.2. Meilleures approximations à M termes d'une image de fleur de $N = 256^2$ pixels. La transformée orthonormale utilise des ondelettes de Daubechies à quatre moments nuls, cinq niveaux et des conditions périodiques au bord. Les intensités affichées sont tronquées à $[0, 1]$.

3.3 Problèmes inverses et régularisation

3.3.1 Le modèle d'observation

Un problème inverse consiste à retrouver un signal inconnu $f \in \mathbb{R}^Q$ à partir de mesures

$$y = \Phi f + w \in \mathbb{R}^P. \quad (3.4)$$

L'opérateur linéaire connu $\Phi \in \mathbb{R}^{P \times Q}$ décrit l'acquisition, et w représente les erreurs inconnues. On applique généralement Φ par un algorithme plutôt que de stocker cette matrice potentiellement immense.



FIGURE 3.3. Deux opérateurs d'observation sans bruit appliqués à la même image de fleur : un flou gaussien et un masque conservant environ 10% des pixels. Les valeurs manquantes sont affichées en noir.

Plusieurs problèmes classiques entrent dans ce cadre :

- Débruitage : $\Phi = \text{Id}$ et $P = Q$.
- Déconvolution : $\Phi f = \varphi \star f$, où le filtre φ décrit le flou et où les conditions au bord sont précisées.
- Données manquantes : Φ sélectionne les coordonnées observées. On peut aussi conserver le tableau complet et utiliser le masque diagonal $\Phi = \text{diag}(\mu_q)$, avec $\mu_q \in \{0, 1\}$.
- Tomographie : Φ modélise des intégrales de l'objet le long de droites, comme une transformée de Radon discrétisée.

Si Φ a un noyau non trivial, les seules mesures ne déterminent pas f . Même lorsqu'une matrice carrée Φ est inversible, ses petites valeurs singulières peuvent amplifier les erreurs dans $\Phi^{-1}y = f + \Phi^{-1}w$. La régularisation apporte une information supplémentaire pour sélectionner une reconstruction utile et stable.

3.3.2 Régularisation parcimonieuse

L'hypothèse $f \approx \Psi x$ avec peu de coefficients actifs conduit à

$$x^* \in \arg \min_{\|x\|_0 \leq M} \|y - \Phi \Psi x\|_2^2, \quad f^* = \Psi x^*. \quad (3.5)$$

Pour le débruitage dans une base orthonormale, on obtient un seuillage dur des coefficients du *signal observé* y . Il fournit le meilleur ajustement parcimonieux à y pour le budget choisi ; il ne garantit pas, à lui seul, la plus petite erreur par rapport au signal propre inconnu f . Pour un opérateur d'acquisition général, $\Phi \Psi$ n'est pas orthogonal et la formule de seuillage ne résout plus le problème.

Les algorithmes gloutons et la relaxation convexe proposent deux façons d'aborder cette difficulté [11, 17]. Pour motiver la relaxation convexe, considérons

$$\|x\|_\alpha^\alpha = \sum_{n=1}^N |x_n|^\alpha, \quad \alpha > 0.$$

Pour chaque vecteur x fixé, cette quantité tend vers $\|x\|_0$ lorsque $\alpha \downarrow 0$. Les ensembles $B_\alpha = \{x : \sum_n |x_n|^\alpha \leq 1\}$ sont contenus dans $[-1, 1]^N$ et convergent, pour la distance de Hausdorff, vers

$$B_0 = \{x \in [-1, 1]^N : \|x\|_0 \leq 1\}.$$

Cet ensemble limite est constitué de segments portés par les axes de coordonnées, y compris l'origine. L'ensemble $\{x : \|x\|_0 \leq 1\}$ sans restriction au cube contient, lui, les axes entiers et n'est pas borné.

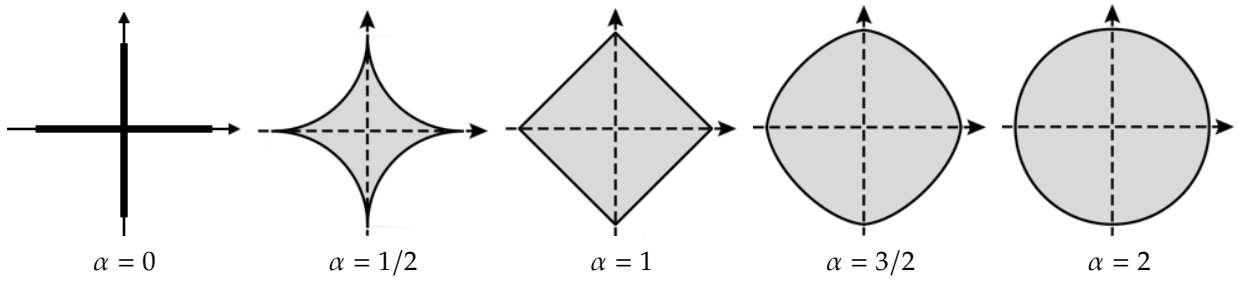


FIGURE 3.4. Les ensembles B_α en dimension deux. Le premier panneau représente l'ensemble limite borné B_0 .

Dans cette famille, $\|\cdot\|_\alpha$ est une norme et B_α est convexe exactement lorsque $\alpha \geq 1$. Le choix $\alpha = 1$ associe la convexité à une géométrie qui favorise les coefficients nuls. On obtient ainsi le problème du Lasso sous contrainte

$$x^* \in \arg \min_{\|x\|_1 \leq \tau} \frac{1}{2} \|Ax - y\|_2^2, \quad A = \Phi \Psi. \quad (3.6)$$

Augmenter τ permet de mieux ajuster les mesures, mais peut aussi conduire à reproduire leurs erreurs. Ce paramètre se choisit en fonction du bruit, de l'opérateur et du modèle de signal. Faire tendre τ vers l'infini supprime la régularisation et laisse généralement plusieurs reconstructions possibles lorsque $P < N$.

Deux formulations proches sont utiles :

$$\min_x \frac{1}{2} \|Ax - y\|_2^2 + \lambda \|x\|_1, \quad \lambda > 0, \quad (3.7)$$

$$\min_x \|x\|_1 \quad \text{sous la contrainte } \|Ax - y\|_2 \leq \varepsilon, \quad \varepsilon \geq 0. \quad (3.8)$$

La formulation pénalisée équilibre fidélité aux données et parcimonie. La seconde, appelée *basis pursuit denoising*, est commode lorsqu'une borne d'erreur ε est connue. Pour $\varepsilon = 0$, elle devient le problème de *basis pursuit*, $\min_{Ax=y} \|x\|_1$. Ces formulations sont reliées par des multiplicateurs de Lagrange appropriés ; leurs paramètres ne sont pas des valeurs numériques interchangeables. Leurs premiers développements concernent notamment l'inversion sismique, la décomposition par *basis pursuit* et le Lasso [5, 14, 16].

La convexité rend possible l'optimisation globale, mais la grande dimension et la non-différentiabilité de $\|x\|_1$ demandent des algorithmes adaptés. Les méthodes proximales exploitent le seuillage doux explicite associé à cette pénalité [6]. Un solveur numérique atteint une précision prescrite, plutôt qu'une solution réelle exacte après un petit nombre d'opérations garanti.

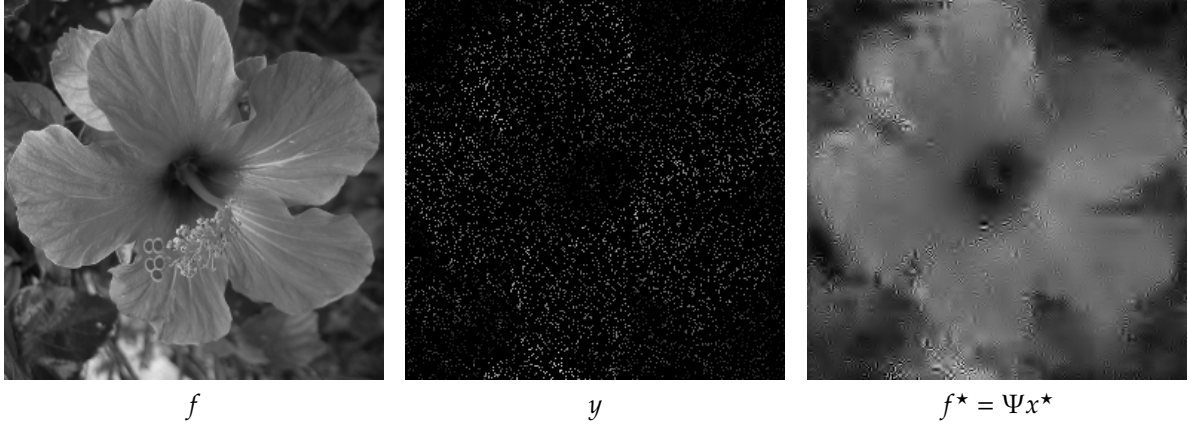


FIGURE 3.5. Reconstruction à partir de 6 554 pixels observés sur 256^2 , soit environ 10%. Le problème pénalisé (3.7) est résolu numériquement avec $\lambda = 0,005$, des intensités dans $[0, 1]$ et les mêmes ondelettes orthonormales qu'à la figure 3.2. Les erreurs de texture restantes illustrent les limites du modèle.

3.3.3 Le rôle de la géométrie

En dimension deux, la boule ℓ^1 possède des sommets sur les axes de coordonnées. Un ensemble de niveau du terme d'attache aux données peut rencontrer la boule pour la première fois en l'un de ces sommets, ce qui produit un coefficient nul. Dans la même configuration, la boule ronde ℓ^2 donne généralement deux coefficients non nuls (figure 3.6). Cette géométrie explique une tendance, sans garantir le résultat pour tout opérateur et toute donnée : un contact le long d'une face peut donner plusieurs minimiseurs, et une solution ℓ^1 n'est pas nécessairement assez parcimonieuse pour retrouver le signal recherché.

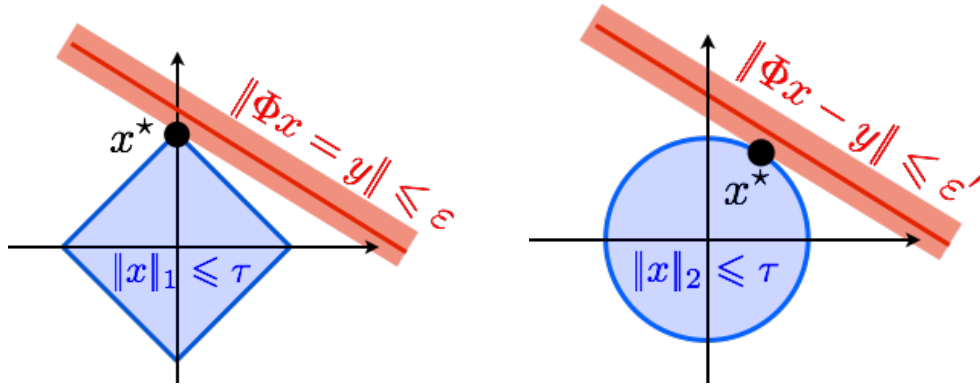


FIGURE 3.6. Ajustement aux données sous une contrainte ℓ^1 (à gauche) ou ℓ^2 (à droite). À partir de $\varepsilon = 0$, on élargit la bande $\{x : \|Ax - y\|_2 \leq \varepsilon\}$ jusqu'à son premier contact avec la boule de contrainte. Le point de contact minimise le résidu sur cette boule.

La reconstruction dépend de l'action de l'opérateur de mesure sur les vecteurs parcimonieux. Augmenter la dimension ambiante ne suffit pas à l'améliorer. Des conditions précises existent pour de nombreux opérateurs aléatoires et structurés ; la convolution et l'acquisition de Fourier demandent des hypothèses adaptées à leur géométrie [1, 3, 4, 9].

3.4 Acquisition comprimée

3.4.1 Mesurer des combinaisons de pixels

Le codage par transformée commence par acquérir Q échantillons, puis en conserve une représentation beaucoup plus petite. L'acquisition comprimée cherche à mesurer directement $P \ll Q$ combinaisons linéaires bien choisies et à reconstruire un signal parcimonieux, ou approximativement parcimonieux, dans une base connue.

La caméra à pixel unique développée à l'université Rice illustre cette idée [8]. Une grille de Q micromiroirs dirige certaines portions de la lumière vers un détecteur unique ; les autres miroirs les dirigent ailleurs. Pour

chaque configuration p , le détecteur mesure la somme

$$y_p = \sum_{q=1}^Q \Phi_{p,q} \int_{c_q} \tilde{f}(s) ds + w_p = \sum_{q=1}^Q \Phi_{p,q} f_q + w_p, \quad \Phi_{p,q} \in \{0, 1\}.$$

Les miroirs n'enregistrent pas les pixels individuels. Changer leur configuration produit directement le vecteur $y \in \mathbb{R}^P$. La scène doit rester suffisamment stable pendant ces mesures. Des configurations complémentaires, ou une mesure étalonnée de la lumière totale, permettent de convertir les mesures 0/1 en mesures signées centrées.

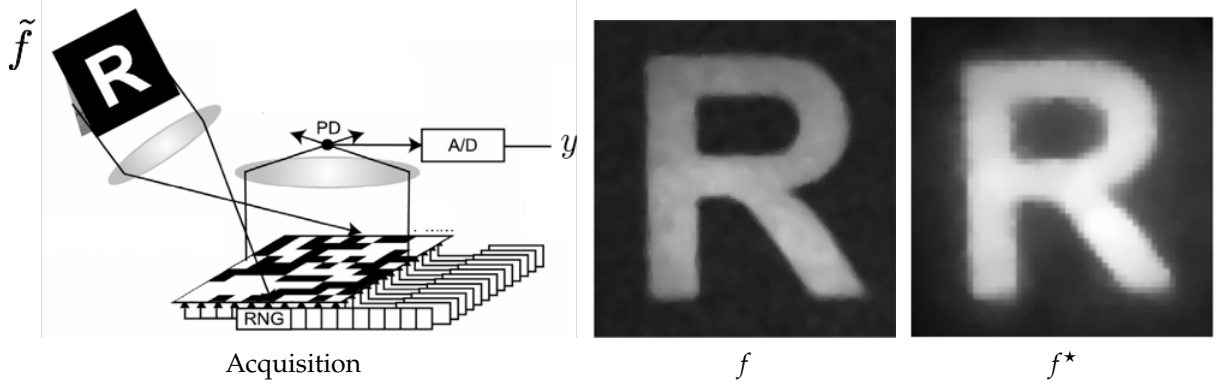


FIGURE 3.7. Acquisition à pixel unique et exemple de reconstruction original, repris des documents associés à [8]. Le facteur annoncé de réduction du nombre de mesures est $Q/P = 6$, donc $P/Q \approx 1/6$.

3.4.2 Une garantie de reconstruction

Voici une forme classique de la garantie d'acquisition comprimée [2, 7]. Supposons que Ψ soit une base orthonormale de $\mathbb{R}^N$, que $f = \Psi x_0$ et que $A = \Phi \Psi$ ait des entrées gaussiennes indépendantes, de moyenne nulle et de variance $1/P$. On peut, de façon équivalente, choisir Φ suivant cette loi gaussienne, qui est invariante par multiplication par une matrice orthogonale fixée.

Il existe des constantes universelles positives C, c, C_0, C_1 telles que, pour $1 \leq M \leq N$ et

$$P \geq CM \log(eN/M), \quad (3.9)$$

avec une probabilité au moins égale à $1 - 2e^{-cP}$ sur le choix de A , la borne suivante soit vraie simultanément pour tous les x_0 et toutes les erreurs vérifiant $\|w\|_2 \leq \varepsilon$. Toute solution x^* de (3.8), avec $y = Ax_0 + w$, satisfait

$$\|x^* - x_0\|_2 \leq C_0 \frac{\sigma_M(x_0)_1}{\sqrt{M}} + C_1 \varepsilon, \quad \sigma_M(x_0)_1 = \min_{\|z\|_0 \leq M} \|x_0 - z\|_1. \quad (3.10)$$

Le logarithme de (3.9) est naturel; changer sa base modifie seulement la constante. Le régime comprimé intéressant est celui où le nombre requis P reste inférieur à N .

La quantité $\sigma_M(x_0)_1$ est la somme des amplitudes des coefficients éliminés après conservation des M plus grands. Si x_0 est exactement M -parcimonieux et si l'on choisit $\varepsilon = 0$ dans le problème sans bruit (donc $w = 0$), les deux termes de (3.10) s'annulent : le *basis pursuit* retrouve exactement x_0 . Pour les signaux approximativement parcimonieux et les mesures bruitées, ces termes séparent l'erreur d'approximation de l'erreur de mesure. L'orthonormalité donne la même erreur euclidienne sur le signal reconstruit : $\|\Psi x^* - f\|_2 = \|x^* - x_0\|_2$.

Une démonstration de ce théorème passe par la propriété d'isométrie restreinte : A préserve approximativement les longueurs euclidiennes des vecteurs portés par des ensembles suffisamment petits. Les matrices aléatoires vérifient cette propriété lorsque le nombre de mesures est de l'ordre indiqué dans (3.9). Des modèles sous-gaussiens centrés appropriés admettent aussi des garanties, avec des constantes dépendant de leur loi. Pour une matrice brute à entrées 0/1, il faut tenir compte du centrage et de la normalisation; ce n'est pas le modèle gaussien énoncé ci-dessus.

Le théorème porte sur un nombre de mesures réelles, et non sur un nombre de bits. Leur précision, leur quantification et leur stockage restent à prendre en compte. L'algorithme de reconstruction doit connaître la matrice de mesure, par exemple grâce à une graine aléatoire partagée. Garder cette matrice secrète ne constitue

pas, à lui seul, une garantie de sécurité cryptographique. De même, les économies de temps ou d'énergie dépendent du dispositif physique et du coût de reconstruction.

Les acquisitions structurées, notamment en imagerie par résonance magnétique, imposent des contraintes supplémentaires : on ne peut généralement pas choisir des mesures gaussiennes indépendantes arbitraires. Les trajectoires d'échantillonnage et les modèles de parcimonie doivent donc être étudiés ensemble [1, 4]. Le principe directeur reste de concevoir conjointement l'acquisition et la reconstruction, en s'appuyant sur un modèle adapté aux données.

Pour aller plus loin. Les liens entre approximation, problèmes inverses et optimisation dépassent la parcimonie. D'autres pénalités décrivent la régularité, le faible rang ou la structure géométrique. L'article de synthèse [17] développe ces idées, et les Numerical Tours [13] proposent des expériences numériques.

Remerciements. Je remercie Charles Dossal, Jalal Fadili, Samuel Vaiter, Stéphane Seuret et le rapporteur anonyme du texte original pour leur aide.

Références

- [1] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier : asymptotic incoherence and asymptotic sparsity in compressed sensing. *CoRR*, abs/1302.0561, 2013. URL <http://arxiv.org/abs/1302.0561>.
- [2] E. J. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8) :1207–1223, 2006.
- [3] E.J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6) :906–956, 2014. ISSN 1097-0312. doi : 10.1002/cpa.21455. URL <http://dx.doi.org/10.1002/cpa.21455>.
- [4] N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss. Variable density sampling with continuous trajectories. *SIAM Journal on Imaging Sciences*, 7(4) :1962–1992, 2014.
- [5] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1) :33–61, 1998. doi : 10.1137/S1064827596304010.
- [6] Patrick L. Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer, 2011. doi : 10.1007/978-1-4419-9569-8_10.
- [7] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4) :1289–1306, 2006.
- [8] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2) :83–91, March 2008. ISSN 1053-5888.
- [9] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5) :1315–1355, 2015. doi : 10.1007/s10208-014-9228-6. URL <http://hal.archives-ouvertes.fr/hal-00839635/>.
- [10] S. G. Mallat. *A Wavelet Tour of Signal Processing : The Sparse Way*. Elsevier/Academic Press, Amsterdam, third edition, 2009.
- [11] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12) :3397–3415, 1993.
- [12] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2) : 227–234, 1995. doi : 10.1137/S0097539792240406.
- [13] Gabriel Peyré. The numerical tours of signal processing. *Computing in Science & Engineering*, 13(4) :94–97, 2011. doi : 10.1109/MCSE.2011.71. URL <https://www.numerical-tours.com/>.
- [14] F. Santosa and W.W. Symes. Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, 7(4) :1307–1330, 1986. doi : 10.1137/0907087.
- [15] C. E. Shannon. Communication in the presence of noise. *Proc. Institute of Radio Engineers*, 37(1) :10–21, 1949.
- [16] R. Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B. Methodological*, 58(1) :267–288, 1996.
- [17] Samuel Vaiter, Gabriel Peyré, and Jalal Fadili. Low complexity regularization of linear inverse problems. In Götz E. Pfander, editor, *Sampling Theory, a Renaissance*, pages 103–153. Birkhäuser, 2015. URL <https://hal.science/hal-01018927>.