Analysis vs Synthesis-based Regularization for combined Compressed Sensing and Parallel MRI Reconstruction at 7 Tesla

H. Cherkaoui, L. El Gueddari, C. Lazarus, A. Grigis, F. Poupon, A. Vignaud, S. Farrens, J.-L. Starck, Philippe Ciuciu

Analysis vs Synthesis-based Regularization for combined Compressed Sensing and Parallel MRI Reconstruction at 7 Tesla

H. Cherkaoui, L. El Gueddari, C. Lazarus, A. Grigis, F. Poupon, A. Vignaud, S. Farrens, J.-L. Starck, P. Ciuciu
CEA Saclay, Univ. Paris-Saclay, 91191 Gif-sur Yvette, France.
firstname.lastname@cea.fr

***Abstract*—Compressed Sensing (CS) has allowed a significant reduction of acquisition times in MRI, especially in the high spatial resolution (e.g., 400 μm) context. Nonlinear CS reconstruction usually relies on analysis (e.g., Total Variation) or synthesis (e.g., wavelet) based priors and ℓ_1 regularization to promote sparsity in the transform domain. Here, we compare the performance of several orthogonal wavelet transforms with those of tight frames for MR image reconstruction in the CS setting combined with parallel imaging (multiple receiver coil). We show that overcomplete dictionaries such as the fast curvelet transform provide improved image quality as compared to orthogonal transforms. For doing so, we rely on an analysis-based formulation where the underlying ℓ_1 regularized criterion is minimized using a primal dual splitting method (e.g., Condat-Vũ algorithm). Validation is performed on ex-vivo baboon brain T_2^* MRI data collected at 7 Tesla and restrospectively under-sampled using non-Cartesian schemes (radial and Sparkling). We show that multiscale analysis priors based on tight frames instead of orthogonal transforms achieve better image quality (pSNR, SSIM) in particular at low signal-to-noise ratio.**

I. Introduction

Context: Compressed-Sensing MRI. Magnetic Resonnance Imaging (MRI) is a key imaging technique to probe soft tissues (e.g., the brain) non-invasively. However, its acquisition time may be prohibitive in the context of high resolution imaging. To cope with this issue, several approaches such as parallel imaging (PI) [1], [2] or Compressed Sensing (CS) [3], [4] have been developed and successively applied to MRI [5] for more than 10 years. PI usually permits to accelerate MRI scans by a 6-8 factor but at the expense of a degraded signal-to-noise ratio (SNR) owing to deterministic under-sampling and aliasing artifacts that arise between channels of the phased-array receiver coil. Instead, CS enables to accelerate MRI exams by randomly under-sampling the data collected in the Fourier domain (k-space) while providing theoritical guarantees of exact image recovery [3]–[5]. This theory is based on (i) sparsity or compressibility of MR images in a given dictionary (e.g., wavelets), (ii) incoherence between sensing and sparsity bases and (iii) nonlinear image reconstruction for promoting sparsity in the transform domain.

Over the last decade, several approaches have been proposed for combining CS and parallel MRI reconstruction either using (possibly joint) sparsity or low-rank regularizing constraints, cf e.g. [6]–[12]. Here, in the combined CS-PI setting, we focus on sparsity promoting regularization which is usually enforced using the ℓ_1-norm of an analysis or a synthesis-based prior. The analysis-based prior operates in the image domain and usually involves some sort of derivative operators (e.g., Total Variation), whereas the synthesis prior acts in the transform domain and involves modern dictionaries for sparse multiscale image representation, such as the wavelet transform (WT) or redundant decompositions (i.e., tight frames) like curvelet and contourlet transforms [13], [14].

Related works: Analysis vs Synthesis penalties in MRI. This topic has been intensively covered in the literature over the last decade by comparing the theoretical properties and practical performance of analysis and synthesis priors in the context of undercomplete, complete and overcomplete transforms [15], [16]. Importantly, it has been shown that in the simpler complete (e.g., orthogonal WT) formulations, analysis and synthesis based formulations are equivalent, whereas they depart one another in their overcomplete formulation (e.g., curvelet, contourlet). Interestingly, it has been shown that the analysis-based formulation provides more accurate results in the overcomplete case, i.e. when redundant dictionaries are used for regularization, cf e.g. [16]. In image analysis and more speficically for MRI reconstruction, the use of undecimated wavelet transforms permits to guarantee the translation invariance of the image solution [8]. For improved localization properties within the space domain and excellent directional selectivity, several groups have tested discrete nonseparable shearlet transforms [17]. Although shearlets are known to be an optimal multiscale sparse representation for natural images, they suffer from high computational cost. Yet, the fast curvelet transform [18], [19] is an efficient and competitive alternative that is endowed with the same approximation rate than shearlets but that has not been tested so far for MR image reconstruction from highly under-sampled k-space *real* data.

Goals and contributions. We first recall the analysis vs synthesis-based formulation of the CS parallel MRI reconstruction problem in Section II, the analysis one being the convenient approach for dealing with overcomplete dictionaries and synthesis one for orthogonal transforms. Next, to make the paper self-contained, the minimization of the ensuing convex but nonsmooth analysis and synthesis-based criteria is addressed. In the analysis case, a primal dual splitting method, known as the Condat-Vũ algorithm [20], [21] is implemented.

Alternatively, for the synthesis-based formulation, an Nesterov accelerated proximal gradient method (.e.g, FISTA algorithm) is used [22]–[24]. Ex-vivo baboon brain high resolution 2D MRI data collected at 7 Tesla and retrospectively under-sampled using different k-space trajectories and various under-sampling factors are presented in Section IV. In Section V, we propose an quantitative comparison of the performances of various multiscale decompositions (orthogonal Mallat WT, Meyer WT, B-spline WT, undecimated bi-orthogonal WT and fast curvelets) for combined CS-PI MR image reconstruction. MR image quality are compared in terms of peak SNR (pSNR) and structural similarity (SSIM) metrics across all tested transforms and for various input SNR and under-sampling schemes. In short, we show that tight frames provide better performances in terms of image quality, especially at low input SNR. Conclusions are drawn in Section VI.

II. CS-PI MRI RECONSTRUCTION

Let L be the number of coils used to acquire the NMR signal, N be the number of pixels of the complex-valued image $\boldsymbol{x}$ to be reconstructed and M the number of samples collected per channel during acquisition. We denote by $\boldsymbol{y}_\ell \in \mathbb{C}^M$ the complex-valued data recorded by the ℓ^{th} channel, $\boldsymbol{S}_\ell \in \mathbb{C}^{N\times N}$ the corresponding diagonal sensitivity matrix. Let F be the NFFT and $\Omega \in \{1,, \ldots, N\}$ the sampling pattern in the k-space, with $|\Omega| = M \ll N$. The CS-PI acquisition model thus reads: $\forall \ell = 1 : L, \boldsymbol{y}_\ell = F_\Omega \boldsymbol{S}_\ell \boldsymbol{x} + \boldsymbol{b}_\ell$ where $\boldsymbol{b}_\ell$ is additive zero-mean Gaussian noise of variance σ_ℓ^2, which can be characterized by a separate scan (without RF pulse) considering the same bandwidth as the prospective CS acquisition.

Analysis-based formulation. The unknown MR image $\boldsymbol{x} \in \mathbb{C}^N$ is sparse in the transform domain: $\boldsymbol{\alpha} = \boldsymbol{\Psi}\boldsymbol{x}$, hence $s \triangleq \#\{i \in \{1, \ldots, N\}, \alpha_i \neq 0\} \ll N$. Dictionary $\boldsymbol{\Psi}$ may refer to an orthogonal wavelet (Daubechies or Meyer) transform or to a tight frame (undecimated bi-orthogonal, B-spline à-trous wavelet transform or curvelet). Hereafter, we will adopt the notation $\boldsymbol{\Psi}_{\text{OT}}$ and $\boldsymbol{\Psi}_{\text{TF}}$ to disentangle orthogonal transforms (OT) from tight frames (TF), with $\boldsymbol{\alpha} = \boldsymbol{\Psi}_{TF}\,\boldsymbol{x} \in \mathbb{C}^P$ and $P \gg N$. The CS-PI reconstruction problem amounts to minimizing of the ℓ_1-analysis-based regularized criterion:

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{C}^N}{\arg\min} \Big\{ \sum_{\ell=1}^{L} \frac{1}{2\sigma_\ell^2} \|F_\Omega \boldsymbol{S}_\ell \boldsymbol{x} - \boldsymbol{y}_\ell\|_2^2 + \lambda \|\boldsymbol{\Psi}\boldsymbol{x}\|_1 \Big\}, \quad (1)$$

where $\lambda > 0$ is the regularization parameter. This formulation requires the knowledge of the sensitivity maps $\{\boldsymbol{S}_\ell\}_\ell$. Here, for the sake of simplicity, we assume these maps known since they can be estimated using self-calibration techniques cf e.g. [25].

When the dictionary $\boldsymbol{\Psi}_{TF}$ is overcomplete, the analysis formulation (1) is the most appealing because the dimension of the optimization problem is smaller. In Section III, we resort to a primal-dual splitting method to minimize Eq. (1).

Synthesis-based formulation. When $\boldsymbol{\Psi}_{\text{OT}}$ is orthogonal, the synthesis-based ℓ_1-penalty, can be injected after the change of variable $\boldsymbol{\alpha} = \boldsymbol{\Psi}_{\text{OT}}\boldsymbol{x}$. The regularized criterion now reads in the transform domain:

$$\widehat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}\in\mathbb{C}^N}{\arg\min} \Big\{ \sum_{\ell=1}^{L} \frac{1}{2\sigma_\ell^2} \|F_\Omega \boldsymbol{S}_\ell \boldsymbol{\Psi}_{\text{OT}}^* \boldsymbol{\alpha} - \boldsymbol{y}_\ell\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \Big\}. \quad (2)$$

The MR image solution is then reconstructed as follows: $\widehat{\boldsymbol{x}} = \boldsymbol{\Psi}_{\text{OT}}^* \widehat{\boldsymbol{\alpha}}$. In Section III, we resort to an accelerated proximal gradient methods [22], [23] to minimize Eq. (2).

III. OPTIMIZATION ALGORITHMS

A. Primal-dual splitting algorithm

The Condat-Vũ algorithm [20], [21] allows us to solve Eq. (1). Our implementation is available in `Pysap` (i.e., Python Sparse data Analysis Package)[1]. In Eq. (1), the data consistency ℓ_2-norm term is a smooth convex function with a β-Lipschitz continuous gradient term that reads:

$$\nabla f(x) = \sum_{\ell=1}^{L} \sigma_\ell^{-2} \boldsymbol{S}_\ell^\dagger F_\Omega^\star \left(F_\Omega \boldsymbol{S}_\ell \boldsymbol{x} - \boldsymbol{y}_\ell\right) \quad (3)$$

where † denotes the Hermitian operator and $F^\star$ the conjugate of the NFFT. Importantly, because of non-Cartesian under-sampling schemes (i.e., Ω is not defined on the Cartesian grid in k-space), the Lipschitz constant β was computed using a power iteration method (eigenvalue decomposition). The Condat-Vũ algorithm is summarized as follows:

Algorithm 1: Condat-Vu algorithm

initialize k = 0, $\tau > 0$, $\gamma > 0$, $\rho_k \in (0, \delta)$, x_0, y_0;
while $k \leq K$ **do**
  $\tilde{x}_{k+1} \leftarrow x_k - \tau \left(\nabla f(x_k) + \boldsymbol{\Psi}^\star y_k\right)$;
  $z \leftarrow y_k + \gamma\, \boldsymbol{\Psi} \left(2\,\tilde{x}_{k+1} - x_k\right)$;
  $\tilde{y}_{k+1} \leftarrow z - \gamma\, \text{prox}_{\lambda\|\cdot\|_1/\gamma} \left(\frac{z}{\gamma}\right)$;
  $(x_{k+1}, y_{k+1}) \leftarrow \rho_k \left(\tilde{x}_k, \tilde{y}_k\right) + (1 - \rho_k) \left(x_k, y_k\right)$
end

where (x_{k+1}, y_{k+1}) denote the solution in the primal and dual spaces, τ and γ are proximal parameters that meet the following condition $1/\tau - \gamma\|\boldsymbol{\Psi}\|^2 \geqslant \beta/2$ and $(\rho_k)_{k\in\mathbb{N}}$ is a sequence of non-negative relaxation parameters upper bounded by $\delta = 2 - \beta/2(1/\tau - \gamma\|\boldsymbol{\Psi}\|^2)^{-1}$. The proximity operator of the ℓ_1-norm is of course the soft-thresholding operator. Weak convergence is guaranteed according to [21, Theorem 3.1].

B. Proximal gradient methods

The original Forward Backward (FB) algorithm is a generalization of gradient descent methods to non-differentiable functions, which can be expressed as follows: $x_{k+1} = \text{prox}_{t_k}\left(x_k - t_k \nabla f(x_k)\right)$. FB accelerations read as follows:

$$x_{k+1} = \text{prox}_{t_k}\left(y_k - t_k \nabla f(y_k)\right), \quad y_{k+1} = B_k(x_k, x_{k+1})$$

[1] cf `https://github.com/neurospin/pisap`

where B_k denotes a linear combination – usually termed "momentum" rule – between the previous x_k and current solutions x_{k+1}. In FISTA [24], B_k is given by:

$$y_{k+1} = x_{k+1} + \frac{1+\tau_k}{\tau_{k+1}}(x_{k+1} - x_k), \quad \tau_{k+1} = \frac{1+\sqrt{1+4\tau_k^2}}{2}.$$

IV. MRI DATA

For validation purposes, we collected Cartesian ex-vivo baboon brain T_2^*-weighted MRI data at 7T (Magnetom Siemens scanner, Erlangen, Germany) using the 32-channel (Nova Medical Inc., Washington, MA, USA) coil. The acquisition parameters were set as follows: TR $= 550$ ms, TE $= 30$ ms and FA $= 25°$ with in-plane resolution of $400\,\mu$m corresponding to an image matrix size of $N = 512 \times 512$ and an acquisition time of 4 min 42 s for each slice. Slice thickness was set to 3 mm to maintain a high SNR during acquisition. Retrospective under-sampling was performed according to radial spokes and brand new multi-shot Sparkling trajectories [26]. Sparkling shots were generated all together using the algorithm proposed in [27] to draw samples according to a variable density with a polynomial decay: $h(k_x, k_y) = 1/|k|^2$. These two kinds of sampling patterns are shown in Fig. 1. Two undersampling factors $R = N/m$, defined as the number of image pixels over the number of measurements, have been tested: the value $R = 2.4$ and $R = 3.3$ corresponding to 40 % and 30 % of samples respectively were used to assess the robustness of MRI reconstruction under various settings. Also, zero-mean white Gaussian noise was added to the k-space data for increasing variances corresponding to input SNR (ISNR) varying from 3 dB to 40 dB (noise-free) approximately. This allows us to mimic varying SNR due to the acquisiton of thinner slices: the ISNR in MRI is actually proportional to the voxel size and thus to the slice thickness.

(a) Sparkling (b) Radial

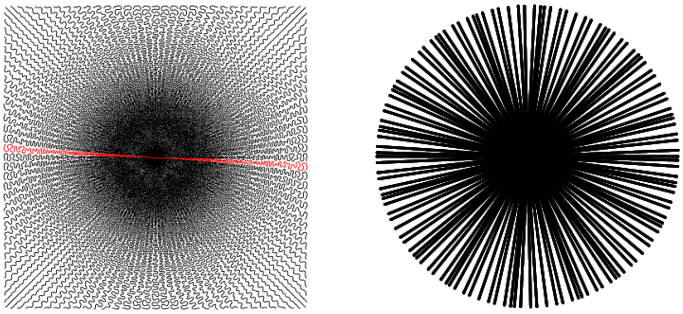

Fig. 1. Non-Cartesian multi-shot under-sampling schemes for $R = 2.4$ i.e., 40 % of full-k-space measurements.

V. MR IMAGE RECONSTRUCTION RESULTS

We performed MR image reconstruction from the retrospectively undersampled k-space data collected using either the radial or Sparkling schemes over the 32 channel coil. In the analysis-based formulation (1), we run the Condat-Vũ algorithm for 200 iterations to guarantee that convergence was achieved. In the synthesis-based formulation (2), we run FISTA

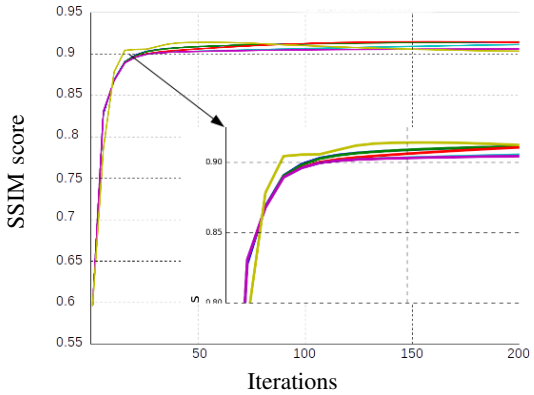

Fig. 2. **Left:** Evolution of the SSIM score over FISTA (yellow curve) and Condat-Vũ (blue, red and green curves) iterations for an orthogonal Mallat WT used as $\mathbf{\Psi}_{\mathrm{OT}}$. **Insert**: zoom showing that FISTA reaches the SSIM plateau slightly earlier than Condat-Vũ.

algorithm for 200 iterations too. We swept several user-defined parameters such as the number of scales ($J \in \{3, 4, 5\}$) in WT and the regularization parameter $\lambda \in \{10^{-1}, \ldots, 10^{-8}\}$ involved in Eqs. (1)-(2). For each ISNR value, a grid search over 20 values of λ and all above mentioned J to retain the pair that gave the best pSNR and SSIM scores (e.g., $J = 5$ and $\lambda = 10^{-5}$ for ISNR $\simeq 30$ dB).

Importantly, we checked that both algorithms converged to the same MR image for orthogonal transforms ($\mathbf{\Psi}_{\mathrm{OT}}$), and we observed a slightly faster convergence for FISTA (cf Fig. 2) in this context. However, we clearly observed that the analysis-based formulation provides much better MR image quality when a tight frame was involved in multiscale analysis as compared to the synthesis-based approach (results not shown). Hence, in what follows, the results obtained using overcomplete transforms are systematically yielded by the analysis formulation and the Condat-Vũ algorithm.

A. Comparison of multiscale decompositions

Due to the lack of space, in Fig. 3 we report the image quality scores (pSNR and SSIM) only for the Sparkling under-sampling scheme with $R = 2.4$ and $R = 3.3$ when sweeping the different multiscale image decompositions. A similar trend was observed for the Radial scheme even though the overall image quality was lower (results not shown). Interestingly, tight frames such as fast curvelets or undecimated bi-orthogonal WT (green and blue traces, respectively) yield a much better image quality score (both in terms of pSNR and SSIM) at low ISNR than orthogonal transforms depicted by yellow, red and purple curves. The worse decompositions are clearly the Mallat and B-spline WT whereas the Meyer WT provides an intermediate solution, whose computational cost is quite fast,for analysing and synthetizing an image as compared to overcomplete dictionaries. Indeed, actual implementation of the Meyer WT is performed in the Fourier domain given its compact support in this space. At high ISNR, all transforms more or less converge towards the same image quality scores. As expected, we observed slightly better image quality scores

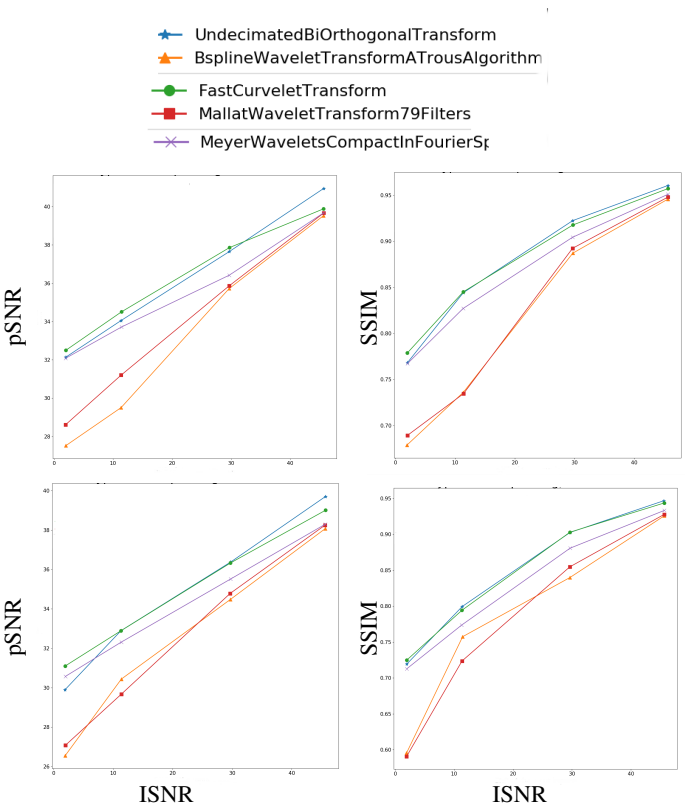

Fig. 3. pSNR (**left**) and SSIM (**right**) scores after CS-PI MR image reconstruction from retrospectively under-sampled k-space data using Sparkling undersampling schemes (**top:** $R = 2.4$, **bottom:** $R = 3.3$). Each line codes for a specific multiscale decomposition, see the legend within each panel.

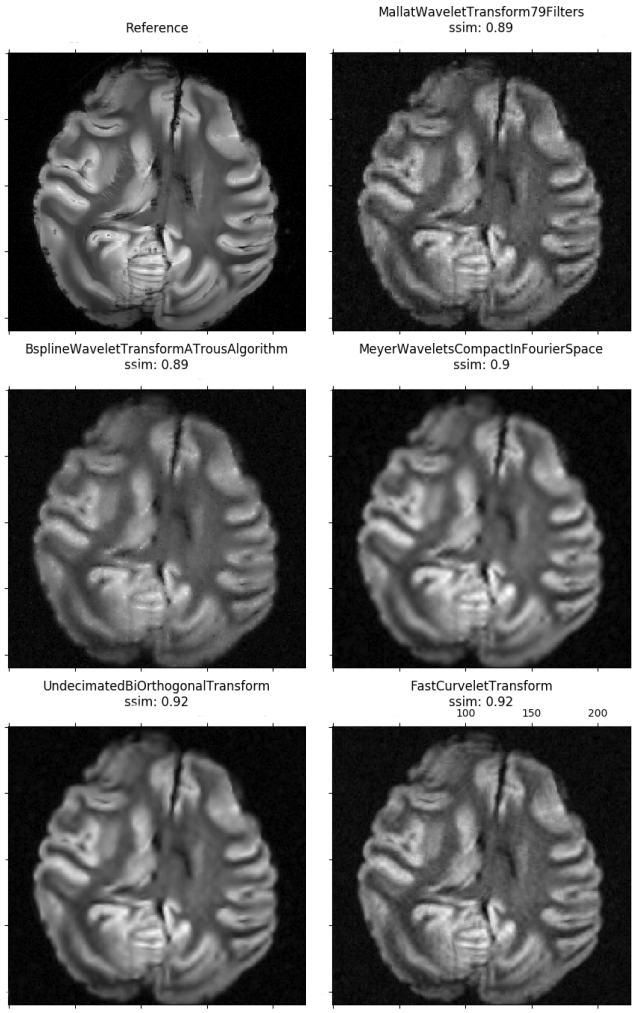

Fig. 4. **Comparisons of multiscale transforms:** Reconstructed MR images from retrospectively under-sampled k-space noisy data using Sparkling sampling pattern ($R = 2.4$ and ISNR $\simeq$ 30 dB). **Top**: Reference MR baboon image (left) and reconstructed image using Eq. 2 and Mallat WT. **Center**: MR images reconstructed image B-spline and Meyer WT. **Bottom**: MR images solutions computed using the analysis-based formulation and tight frames. The SSIM scores appear on top of each panel.

for $R = 2.4$ as compared to $R = 3.3$, but the Sparkling undersampling scheme was quite robust. In comparison, the image quality scores significantly drop off for the Radial scheme when moving from $R = 2.4$ to $R = 3.3$. This point is illustrated in the noise-free case in Fig. 5.

For illustration purposes, in Fig. 4, we show the reference baboon brain image as well as all reconstruced MR images using the five tested multiscale decompositions in the scenario corresponding to an ISNR of about 30 dB. As expected given Fig. 3, we retrieved better image quality using tight frames. It is worth noting that even though different decompositions give the same SSIM score, visually the images appear quite different. The MR image reconstructed using the Mallat WT looks noisier and the image solution reconstructed from the Meyer WT looks smoother. The same remark holds for tight frames which both achieve an SSIM score of 0.92. This basically means that any single quantitative index is *not* able to capture all image properties.

B. Comparison of sampling schemes

Beyond the comparison of multiscale decompositions, the under-sampling scheme may significantly impact the image quality, both in terms of shape of k-space trajectories (radial spokes vs Sparkling shots) and number of measurements (R factor). The current study was performed retrospectively and thus neglects potential discrepancies between prescribed sampling trajectories and the actual ones that may occur in prospective accelerated acquisition scenarios. For reconstruction, we used the undecimated bi-orthogonal WT and considered the noise-free scenario (ISNR $= 40$ dB) to reach the asymptotic MR image quality. As shown in Fig. 5, the Sparkling pattern yielded improved MR image quality as compared to Radial lines for a given R factor. Also, Sparkling is much more robust than Radial lines since it provides the same SSIM score for $R = 2.4$ and $R = 3.3$, whereas the image quality was signigicantly degraded when Radial lines were more undersampled (SSIM $= 0.83$ for $R = 3.3$ compared to SSIM $= 0.91$ for $R = 2.4$).

VI. Conclusions

In this paper, we have compared analysis and synthesis-based formulations for compressed sensing MR image recon-

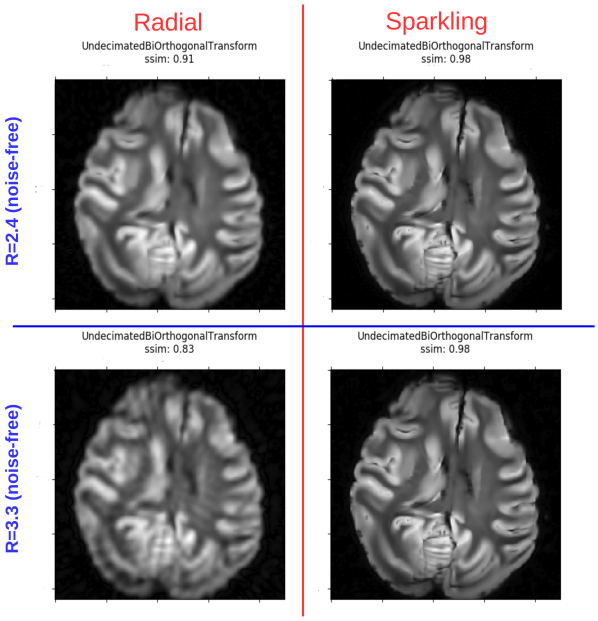

Fig. 5. **Comparisons of under-sampling schemes (Radial, Sparkling) in the noise-free case.** MR reconstructed images from retrospectively undersampled data ($R = 2.4$ in top row, $R = 3.3$ in bottom row) using the analysis-based formulation (1) with the undecimated bi-orthogonal transform.

struction in the parallel imaging context, i.e. when multiple receivers are combined within the same phased array coil. As already known in the literature, we have shown that translation invariant overcomplete decomposition outperform orthogonal wavelet transforms especially at low input SNR. Well known optimization algorithms have been respectively implemented to minimize the cost functions associated with the analysis and synthesis formulations. We have also pointed out the superiority of the Sparkling under-sampling scheme over the Radial one in terms of quality assessed by SSIM scores. We also provide a Python package, called `Pysap`, for CS image reconstruction in MRI and astrophysics, interfaced with the `pynfft` package to deal with non-Cartesian Fourier sampling.

Future work will be devoted to extend this analysis to 3D multiscale decompositions and especially to identify which of the Beam-curvelet or ridge-curvelet are more accurate to sparsify isotropic high resolution 3D MR images, which are collected using dedicated 3D Sparkling trajectories.

REFERENCES

[1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger *et al.*, "SENSE: sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.

[2] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, 2002.

[3] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[4] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[5] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[6] M. Lustig, M. Alley, S. Vasanawala, D. Donoho, and J. Pauly, "ℓ_1 SPIRiT: Autocalibrating parallel imaging compressed sensing," in *Proc Intl Soc Mag Reson Med*, vol. 17, 2009, p. 379.

[7] F. Huang, Y. Chen, W. Yin, W. Lin, X. Ye, W. Guo, and A. Reykowski, "A rapid and robust numerical algorithm for sensitivity encoding with sparsity constraints: Self-feeding sparse SENSE," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1078–1088, 2010.

[8] M. Guerquin-Kern, M. Haberlin, K. Pruessmann, and M. Unser, "A fast wavelet-based reconstruction method for magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1649–1660, 2011.

[9] L. Chaari, J.-C. Pesquet, A. Benazza-Benyahia, and P. Ciuciu, "A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging," *MedIA*, vol. 15, no. 2, pp. 185–201, 2011.

[10] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z.-P. Liang, "Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1809–1820, 2012.

[11] I. Y. Chun, B. Adcock, and T. M. Talavage, "Efficient compressed sensing SENSE pMRI reconstruction with joint sparsity promotion," *IEEE Trans. Imag. Proc.*, vol. 35, no. 1, pp. 354–368, 2016.

[12] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel mri using annihilating filter based low-rank hankel matrix," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 480–495, 2016.

[13] E. J. Candes and D. L. Donoho, "Curvelets: A surprisingly effective nonadaptive representation for objects with edges," Stanford Univ Ca Dept of Statistics, Tech. Rep., 2000.

[14] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Imag. Proc.*, vol. 14, no. 12, pp. 2091–2106, 2005.

[15] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse problems*, vol. 23, no. 3, p. 947, 2007.

[16] I. W. Selesnick and M. A. Figueiredo, "Signal restoration with overcomplete wavelet transforms: Comparison of analysis and synthesis priors," in *Wavelets XIII*, vol. 7446. International Society for Optics and Photonics, 2009, p. 74460D.

[17] S. Pejoski, V. Kafedziski, and D. Gleich, "Compressed sensing mri using discrete nonseparable shearlet transform and fista," *IEEE Sig. Proc. Let.*, vol. 22, no. 10, pp. 1566–1570, 2015.

[18] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Imag. Proc.*, vol. 11, no. 6, pp. 670–684, 2002.

[19] E. Candes, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling & Simulation*, vol. 5, no. 3, pp. 861–899, 2006.

[20] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Adv. in Comput. Math.*, vol. 38, no. 3, pp. 667–681, 2013.

[21] L. Condat, "A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Th. App.*, vol. 158, no. 2, pp. 460–479, 2013.

[22] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*. Springer, 2011, pp. 185–212.

[23] Y. Nesterov *et al.*, "Gradient methods for minimizing composite objective function," 2007.

[24] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[25] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT—An Eigenvalue Approach to Autocalibrating Parallel MRI: Where SENSE Meets GRAPPA," *Magn. Reson. Med.*, vol. 71, pp. 990–1001, 2014.

[26] C. Lazarus, P. Weiss, N. Chauffert, F. Mauconduit, M. Bottlaender, A. Vignaud, and P. Ciuciu, "SPARKLING: Novel Non-Cartesian Sampling Schemes for Accelerated 2D Anatomical Imaging at 7T Using Compressed Sensing," in *25th ISMRM conf.*, Honolulu, USA, 2017.

[27] C. Boyer, N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss, "On the generation of sampling schemes for magnetic resonance imaging," *SIAM J. Imaging Sci.*, vol. 9, no. 4, pp. 2039–2072, 2016.