

# ACCELERATED PARALLEL MAGNETIC RESONANCE IMAGING RECONSTRUCTION USING JOINT ESTIMATION WITH A SPARSE SIGNAL MODEL

*Daniel S. Weller<sup>1</sup>, Jonathan R. Polimeni<sup>2,3</sup>, Leo Grady<sup>4</sup>,  
Lawrence L. Wald<sup>2,3</sup>, Elfar Adalsteinsson<sup>1</sup>, Vivek K Goyal<sup>1</sup>*

<sup>1</sup> Dept. of EECS, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>2</sup> A. A. Martinos Center, Dept. of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

<sup>3</sup> Dept. of Radiology, Harvard Medical School, Boston, MA, USA

<sup>4</sup> Dept. of Image Analytics and Informatics, Siemens Corporate Research, Princeton, NJ, USA

`dwell@mit.edu, jonp@nmr.mgh.harvard.edu, leo.grady@siemens.com,  
wald@nmr.mgh.harvard.edu, elfar@mit.edu, vgoyal@mit.edu`

## ABSTRACT

Accelerating magnetic resonance imaging (MRI) by reducing the number of acquired k-space scan lines benefits conventional MR imaging significantly by decreasing the time subjects remain in the magnet. In this paper, we formulate a novel method for Joint estimation from Undersampled LinEs in Parallel MRI (JULEP) that simultaneously calibrates the GeneRalized Autocalibrating Partially Parallel Acquisitions (GRAPPA) reconstruction kernel and reconstructs the full multi-channel k-space. We employ a joint sparsity signal model for the channel images in conjunction with observation models for both the acquired data and GRAPPA reconstructed k-space. We demonstrate using real MRI data that JULEP outperforms conventional GRAPPA reconstruction at high levels of undersampling, increasing peak-signal-to-noise ratio by up to 10 dB.

**Index Terms**— Magnetic resonance imaging, image reconstruction, parallel imaging, sparsity, Bayesian estimation

## 1. INTRODUCTION

In conventional magnetic resonance imaging (MRI), the desired image or volume commonly is acquired by raster scanning the spatial Fourier transform domain called k-space, sampling these scan lines, and taking the inverse Fourier transform of these samples. MR imaging is fundamentally limited by the time required to trace all these scan lines. Greatly accelerating MRI by reducing the number of lines acquired is accompanied by reductions in image quality. Accelerated parallel MRI undersamples k-space, using fewer lines spaced farther apart, reducing the field of view (FOV), and introducing aliasing when the object is larger than the

reduced FOV; parallel imaging resolves this aliasing and produces a full-FOV image using parallel receivers.

Compressed sensing (CS) leverages the compressibility of MR images to reconstruct images from randomly undersampled data [1]. Methods like SENSE [2], GRAPPA [3], and SPIRiT [4] also are effective for reconstructing images from undersampled data, but these methods are affected by noise amplification or residual aliasing at high levels of acceleration. The synergistic combination of SENSE [5] or SPIRiT [6] with CS aims to address the shortcomings of accelerated parallel imaging and achieve acceleration beyond what is achievable with either approach individually.

The synergistic combination of GRAPPA with a sparsity model is more complicated due to characteristics of GRAPPA including the presence of a calibration step, the use of uniformly-spaced (nonrandom) undersampling, and the reconstruction of multiple channel images rather than a single combined image. Building upon the previous development of sparsity-based DESIGN denoising [7] of GRAPPA-reconstructed data and sparsity-promoting GRAPPA kernel calibration [8], we reinterpret the sparse reconstruction problem and devise a method for Joint estimation from Undersampled LinEs in Parallel MRI (JULEP) that elegantly unifies the kernel calibration and full k-space reconstruction/denoising problems with a single solution. This proposed method improves reconstruction quality at high levels of acceleration, accommodating both greater spacing between k-space samples and fewer autocalibration (ACS) lines.

## 2. THEORY

Consider the  $N \times P$  matrix  $\mathbf{Y}$ , whose columns represent the full k-space we wish to estimate for each of the  $P$  coil channels. Each of the coil images is presumed to be approximately sparse in some appropriate transform domain with analysis

Funding acknowledgments: NSF CAREER Grant CCF-0643836, NIH R01 EB007942 and EB006847, NIH NCCR P41 RR014075, Siemens Corporate Research, and an NSF Graduate Research Fellowship.

(forward) transform  $\Psi$ . Let  $\mathbf{W} = \Psi \mathbf{F}^{-1} \mathbf{Y}$ , where  $\mathbf{F}$  is the discrete Fourier transform; then we can impose a sparsity model on  $\mathbf{W}$  corresponding to the  $\ell_{p,q}$  mixed norm, parameterized by  $\lambda > 0$ , which controls the level of sparsity:

$$p(\mathbf{W}) \propto e^{-\lambda \|\mathbf{W}\|_{p,q}^p} = \prod_{n=0}^{N-1} e^{-\lambda \|W_1[n], \dots, W_P[n]\|_q^p}. \quad (1)$$

To maintain convexity, we choose  $p = 1$  (one could also impose priors for  $\frac{1}{2} < p < 1$ ; see [9]). Setting  $q = 1$  favors channel-by-channel sparsity, where the support of each transformed channel image is considered independently of those of the other channel images. For joint sparsity [10], where the supports of each transformed channel are assumed to overlap sufficiently to gain by assuming all the channels share the same support, we can set  $q = 2$  (loose joint sparsity) or  $q = \infty$  (strict joint sparsity) to tie these channels together. In this work, we proceed assuming loose joint sparsity and set  $q = 2$ . Assuming  $\mathbf{W}$  is complex,

$$p(\mathbf{W}) = \prod_{n=0}^{N-1} \frac{P! \lambda^{2P}}{(2P)! \pi^P} e^{-\lambda \|W_1[n], \dots, W_P[n]\|_2}. \quad (2)$$

In accelerated MRI, the full k-space can be subdivided into  $\mathbf{Y}_a = \mathbf{K}_a \mathbf{Y}$ , the acquired k-space, and  $\mathbf{Y}_{na} = \mathbf{K}_{na} \mathbf{Y}$ , the un-acquired k-space. From the acquisition, we have direct observations of acquired k-space  $\mathbf{D}_a$ , and based on the large number of individual spins, we model our observation noise as complex Gaussian (with real and imaginary parts uncorrelated). As the noise is typically assumed to be random perturbations due to thermal and other uncorrelated variations, we assume this noise is iid across k-space frequency measurements, and we allow for correlation across channels for the same frequencies. Formally,

$$p(\mathbf{D}_a | \mathbf{Y}) = \mathcal{CN}(\text{vec}(\mathbf{D}_a); \text{vec}(\mathbf{K}_a \mathbf{Y}), \mathbf{I}_{M \times M} \otimes \Lambda), \quad (3)$$

where  $\mathcal{CN}(\cdot; \boldsymbol{\mu}, \Lambda)$  is the density function of the complex Normal distribution with mean  $\boldsymbol{\mu}$  and covariance  $\Lambda$ ,  $\text{vec}(\cdot)$  stacks the columns of a matrix into one vector, and  $\otimes$  is the Kronecker product.

Additionally, the GRAPPA reconstruction method provides observations of the un-acquired data. Given the appropriate GRAPPA kernel  $\mathbf{G}$ , the GRAPPA reconstruction ideally yields  $\mathbf{Y}_{na} = \text{GRAPPA}(\mathbf{G}, \mathbf{Y}_a)$ . Since GRAPPA is linear in  $\mathbf{Y}_a$ , substituting  $\mathbf{D}_a$  yields  $\mathbf{Y}_{na} +$  amplified noise. Calling these GRAPPA-reconstructed observations  $\mathbf{D}_{na}$ , the likelihood of  $\mathbf{D}_{na}$  given the full k-space  $\mathbf{Y}$  is

$$p(\mathbf{D}_{na} | \mathbf{Y}) = \mathcal{CN}(\text{vec}(\mathbf{D}_{na}); \text{vec}(\mathbf{K}_{na} \mathbf{Y}), \Lambda_G), \quad (4)$$

where the amplified noise has covariance  $\Lambda_G$ .

Putting these signal and observation models together, the minimum mean squared error estimator is the posterior mean  $\mathbb{E}\{\mathbf{Y} | \mathbf{D}_a, \mathbf{D}_{na}\}$ . Due to variable mixing in both the signal and observation models, this estimator does not have a

closed form, and numeric integration methods like quadrature are computationally infeasible (due to the curse of dimensionality). One could resort to stochastic methods (e.g. importance sampling), but convergence may be rather slow due to the number of correlated variables involved. Instead, we propose finding the maximum *a posteriori* (MAP) estimate, a compressed sensing-like optimization problem:

$$\mathbf{Y} = \underset{\mathbf{Y}}{\text{minimize}} \frac{1}{2} \|\text{vec}(\mathbf{K}_a \mathbf{Y} - \mathbf{D}_a)\|_{\mathbf{I}_{M \times M} \otimes \Lambda}^2 + \frac{1}{2} \|\text{vec}(\mathbf{K}_{na} \mathbf{Y} - \mathbf{D}_{na})\|_{\Lambda_G}^2 + \lambda \|\Psi \mathbf{F}^{-1} \mathbf{Y}\|_{1,2}. \quad (5)$$

The notation  $\|\mathbf{x}\|_{\Lambda}$  is shorthand for  $\|\Lambda^{-1/2} \mathbf{x}\|_2$  (for Hermitian symmetric positive definite  $\Lambda$ ). This simple estimator can be implemented and solved using various techniques, including iteratively reweighted least squares (IRLS) [11]. The use of IRLS to solve this type of problem is carefully illustrated in [7, 8]. This formulation effectively denoises the acquired and GRAPPA reconstructed k-space, similar to the DESIGN denoising method already developed.

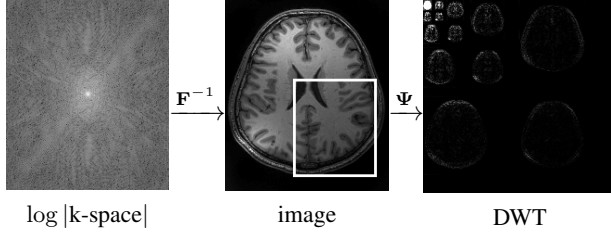
Now, suppose the GRAPPA kernel  $\mathbf{G}$  is now an unknown variable, instead of simply being computed from the ACS lines. From the matrices of source points  $\mathbf{Y}_s$  and target points  $\mathbf{Y}_t$  from our ACS lines, with observations  $\mathbf{D}_s$  and  $\mathbf{D}_t$ , we add observations of the form  $\mathbf{G} \mathbf{D}_s = \mathbf{Y}_t + \text{noise}$ . Assuming that the source points follow the same observation model as the other acquired data, the noise will be complex Gaussian and amplified by the GRAPPA kernel (call the noise covariance  $\Lambda_{\text{ACS}}$ ). The optimization problem now becomes joint over the full k-space  $\mathbf{Y}$  and the GRAPPA kernel  $\mathbf{G}$ :

$$\{\mathbf{Y}, \mathbf{G}\} = \underset{\mathbf{Y}, \mathbf{G}}{\text{minimize}} \frac{1}{2} \|\text{vec}(\mathbf{Y}_t - \mathbf{G} \mathbf{D}_s)\|_{\Lambda_{\text{ACS}}}^2 + \frac{1}{2} \|\text{vec}(\mathbf{K}_{na} \mathbf{Y} - \text{GRAPPA}(\mathbf{G}, \mathbf{D}_a))\|_{\Lambda_G}^2 + \frac{1}{2} \|\text{vec}(\mathbf{K}_a \mathbf{Y} - \mathbf{D}_a)\|_{\mathbf{I}_{M \times M} \otimes \Lambda}^2 + \lambda \|\Psi \mathbf{F}^{-1} \mathbf{Y}\|_{1,2}. \quad (6)$$

At first glance, this optimization problem has the same form as Eq. (5); however, the covariance matrices  $\Lambda_G$  and  $\Lambda_{\text{ACS}}$  depend on  $\mathbf{G}$ , so the first two parts of the objective are not strictly least-squares terms, and the overall problem is not convex. We approach the problem by fixing these covariance matrices, applying IRLS to form an iterative algorithm, and update the covariance matrices at the same time we update the diagonal IRLS re-weighting matrix  $\Delta$ .

$$\{\mathbf{Y}, \mathbf{G}\} = \underset{\mathbf{Y}, \mathbf{G}}{\text{minimize}} \frac{1}{2} \|\text{vec}(\mathbf{Y}_t - \mathbf{G} \mathbf{D}_s)\|_{\Lambda_{\text{ACS}}}^2 + \frac{1}{2} \|\text{vec}(\mathbf{K}_{na} \mathbf{Y} - \text{GRAPPA}(\mathbf{G}, \mathbf{D}_a))\|_{\Lambda_G}^2 + \frac{1}{2} \|\text{vec}(\mathbf{K}_a \mathbf{Y} - \mathbf{D}_a)\|_{\mathbf{I}_{M \times M} \otimes \Lambda}^2 + \frac{\lambda}{2} \|\Delta^{\frac{1}{2}} \Psi \mathbf{F}^{-1} \mathbf{Y}\|_F^2, \quad (7)$$

where  $\Delta_{n,n} = \frac{1}{\|W_1[n], \dots, W_P[n], \epsilon\|_2}$ , with  $\mathbf{W} = \Psi \mathbf{F}^{-1} \mathbf{Y}$  from the previous iteration, and small  $\epsilon > 0$ .



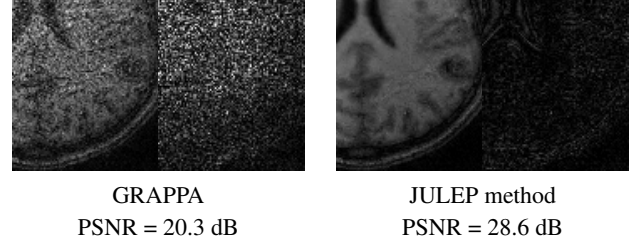
**Fig. 1.** This  $T_1$ -weighted image, acquired in the k-space domain, has a sparse four-level ‘9-7’ DWT representation.

Since Eq. (6) is not convex, convergence to a global minimum is not guaranteed, and initialization affects the solution. The initial full k-space is zero-filled, containing only the acquired data, and the initial GRAPPA kernel is the least-squares or minimum energy solution to the system of ACS fit equations ( $\mathbf{G} = \mathbf{D}_t \mathbf{D}_s^\dagger$ , where  $\mathbf{D}_s^\dagger$  is the left or right pseudo-inverse of  $\mathbf{D}_s$ ). We anticipate convergence when the objective decreases by less than  $tol$  percent. While this criterion is susceptible to the possibility that updating  $\mathbf{\Lambda}_{ACS}$  and  $\mathbf{\Lambda}_G$  can increase the objective, such behavior did not hamper our simulations. Although the covariance matrices have lots of nonzero cross terms, because the GRAPPA kernel introduces correlations among k-space frequencies, we focus on the noise amplification aspect of GRAPPA by using a block diagonal matrix with correlations only across channels (blocks are  $P \times P$ ). The JULEP method is summarized below:

1. Set  $\mathbf{G}$  and  $\mathbf{Y}$ , and initialize  $\mathbf{\Lambda}_G$  and  $\mathbf{\Lambda}_{ACS}$ .
2. Compute re-weighting matrix  $\mathbf{\Delta}$  for the  $\ell_{1,2}$  term using the current estimate of  $\mathbf{Y}$ .
3. Update  $\mathbf{G}$  and  $\mathbf{Y}$  using an iterative solver for Eq. (7).
4. Update  $\mathbf{\Lambda}_G$  and  $\mathbf{\Lambda}_{ACS}$  based on the new value of  $\mathbf{G}$ .
5. Repeat steps (2)–(4) until the objective decreases by less than  $tol$  percent.

### 3. METHODS

A  $T_1$ -weighted full-FOV image ( $256 \times 256 \times 176$  mm FOV; 1.0 mm isotropic resolution) is acquired on a Siemens 3 T scanner using a Siemens 32-channel receive array head coil. An axial slice is extracted, cropped, and normalized; the combined magnitude reference image and its k-space are shown in Fig. 1 along with its sparse transform using a four-level ‘9-7’ 2-D DWT. The k-space is undersampled in both directions in the axial plane, a  $36 \times 36$  ACS block is retained in the center of k-space, and full-FOV images are reconstructed using GRAPPA and the proposed method.



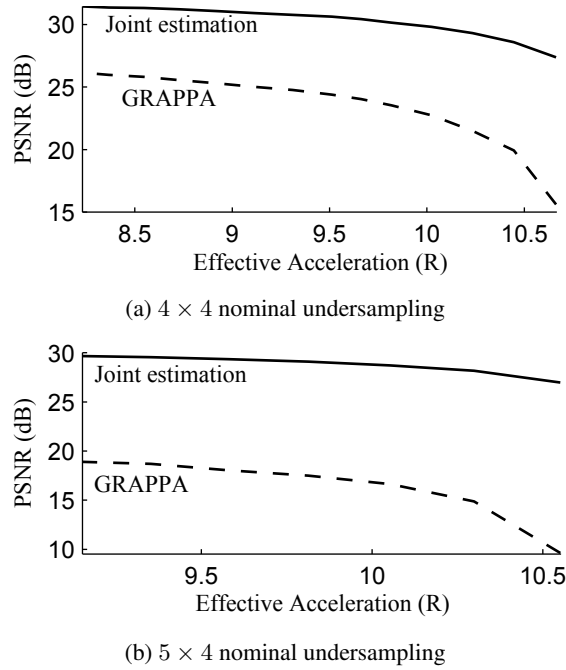
**Fig. 2.** The inset images and difference images above suggest that while the GRAPPA reconstruction is very noisy, the JULEP joint estimated image has much less noise.

The channel noise covariance matrix  $\mathbf{\Lambda}$  is measured from a separate noise-only (no RF excitation) acquisition. The reconstructed full k-space is combined to form a single combined magnitude image using the SNR-optimal coil combination based on sensitivities estimated from apodized ACS lines and  $\mathbf{\Lambda}$ . The magnitude images are compared using difference images and peak-signal-to-noise ratio (PSNR). Although PSNR, as a non-specific measure, does not accurately reflect image quality as perceived by a clinician, it facilitates study of quantitative trends in image quality. A convergence threshold of  $tol = 1\%$  is used for all experiments.

### 4. RESULTS

In the first experiment, the  $T_1$ -weighted k-space shown in Fig. 1 is nominally undersampled by a factor of 5 along the coronal axis (vertical direction) and a factor of 4 along the sagittal axis (horizontal direction), and a  $36 \times 36$  block of ACS data is retained, for an effective acceleration of  $R = 12.1$ . For both the GRAPPA and JULEP algorithms, a GRAPPA kernel for  $5 \times 4$  source points is calibrated for each of the target points. In the case of un-regularized GRAPPA kernel calibration, the calibrated kernel yields a GRAPPA-reconstructed image with significant noise amplification in about 1 second. The joint estimation of the kernel and full k-space produces a reconstruction in under 80 minutes (4 IRLS steps with a total of 555 inner iterations of the least-squares solver LSMR [12]) with far less noise than before while retaining important structural information. The inset region (outlined by the rectangle in Fig. 1) of the reconstructed images alongside corresponding inset regions of the difference images in Fig. 2 portray a significant difference in image quality.

This experiment is repeated on the same image for both  $4 \times 4$  and  $5 \times 4$  nominal undersampling, changing the effective acceleration by varying the number of ACS lines, and comparing the magnitude image PSNRs for both GRAPPA and joint estimation reconstructions. The trend lines portrayed in Fig. 3 confirm that as the effective acceleration increases due to having fewer ACS lines, the JULEP method consistently produces significantly higher quality images (as measured crudely by PSNR). The ACS block grows from  $35 \times 35$



**Fig. 3.** As the number of ACS lines varies, changing the total acceleration, the JULEP joint estimation method consistently outperforms the GRAPPA reconstruction in terms of PSNR.

to  $48 \times 48$  for the  $4 \times 4$  nominally undersampled data, and from  $42 \times 42$  to  $48 \times 48$  for the  $5 \times 4$  undersampled data. For fewer ACS lines, the number of fit equations is insufficient to perform a least-squares fit for the GRAPPA kernel calibration. Similar improvement in image quality is observed in the first experiment, which uses underdetermined calibration.

## 5. DISCUSSION

The significant improvement in image quality achievable by the proposed JULEP joint estimation algorithm is evident in both experiments, yielding 5-10 dB improvement over GRAPPA. When considering the number of ACS lines needed for a given nominal undersampling factor to achieve a minimum PSNR, this novel estimation method can achieve that PSNR with far fewer ACS lines. Shifting this trade-off has implications for MRI acquisitions where maximum acceleration is desired; i.e. ACS lines are expensive. While we cannot validate these methods clinically without extensive evaluation of images featuring abnormalities like lesions, the proposed method enables greater acceleration of MRI.

## 6. REFERENCES

- [1] 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. 1182–95, Dec. 2007.

- [2] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.
- [3] 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. 1202–10, June 2002.
- [4] M. Lustig and J. M. Pauly, "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space," *Magn. Reson. Med.*, vol. 64, no. 2, pp. 457–471, Aug. 2010.
- [5] D. Liang, B. Liu, J. Wang, and L. Ying, "Accelerating SENSE using compressed sensing," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1574–1584, Dec. 2009.
- [6] M. Lustig, M. Alley, S. Vasanawala, D. L. Donoho, and J. M. Pauly, "L<sub>1</sub> SPIRiT: Autocalibrating parallel imaging compressed sensing," in *Proc. ISMRM 17th Scientific Meeting*, Apr. 2009, p. 379.
- [7] D. S. Weller, J. R. Polimeni, L. Grady, L. L. Wald, E. Adalsteinsson, and V. K. Goyal, "Denoising sparse images from GRAPPA using the nullspace method (DESIGN)," *Magn. Reson. Med.*, to appear.
- [8] D. S. Weller, J. R. Polimeni, L. Grady, L. L. Wald, E. Adalsteinsson, and V. K. Goyal, "Regularizing grappa using simultaneous sparsity to recover de-noised images," in *Proc. SPIE Wavelets and Sparsity XIV*, Aug. 2011, vol. 8138, pp. 81381M–1–9.
- [9] D. S. Weller, J. R. Polimeni, L. Grady, L. L. Wald, E. Adalsteinsson, and V. K. Goyal, "Evaluating sparsity penalty functions for combined compressed sensing and parallel MRI," in *Proc. IEEE Int. Symp. on Biomedical Imaging*, March–April 2011, pp. 1589–92.
- [10] S.F. Cotter, B.D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–88, July 2005.
- [11] P. W. Holland and R. E. Welsch, "Robust regression using iteratively re-weighted least-squares," *Commun. Statist. A: Theory Meth.*, vol. 6, no. 9, pp. 813–827, 1977.
- [12] D. C.-L. Fong and M. A. Saunders, "LSMR: An iterative algorithm for sparse least-squares problems," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2950–2971, 2011.