COMBINED COMPRESSED SENSING AND PARALLEL MRI COMPARED FOR UNIFORM AND RANDOM CARTESIAN UNDERSAMPLING OF K-SPACE

Daniel S. Weller[1], *Jonathan R. Polimeni*[2,3], *Leo Grady*[4], *Lawrence L. Wald*[2,3], *Elfar Adalsteinsson*[1], *Vivek K Goyal*[1]

[1] Dept. of EECS, Massachusetts Institute of Technology, Cambridge, MA, USA
[2] A. A. Martinos Center, Dept. of Radiology, Massachusetts General Hospital, Charlestown, MA, USA
[3] Dept. of Radiology, Harvard Medical School, Boston, MA, USA
[4] Dept. of Image Analytics and Informatics, Siemens Corporate Research, Princeton, NJ, USA
dweller@mit.edu, jonp@nmr.mgh.harvard.edu, leo.grady@siemens.com, wald@nmr.mgh.harvard.edu, elfar@mit.edu, vgoyal@mit.edu

ABSTRACT

Both compressed sensing (CS) and parallel imaging effectively reconstruct magnetic resonance images from undersampled data. Combining both methods enables imaging with greater undersampling than accomplished previously. This paper investigates the choice of a suitable sampling pattern to accommodate both CS and parallel imaging. A combined method named SpRING is described and extended to handle random undersampling, and both GRAPPA and SpRING are evaluated for uniform and random undersampling using both simulated and real data. For the simulated data, when the undersampling factor is large, SpRING performs better with random undersampling. However, random undersampling is not as beneficial to SpRING for real data with approximate sparsity.

***Index Terms*—** Compressed sensing, magnetic resonance imaging, image reconstruction, parallel imaging, sampling patterns

1. INTRODUCTION

The time-consuming task of sampling multiple dimensions by image encoding in k-space limits the affordability, image quality, and applications of magnetic resonance imaging (MRI). In Cartesian k-space sampling, the scan time is proportional to the number of phase-encode lines of k-space required to achieve the desired field of view (FOV), voxel size and signal-to-noise ratio (SNR). Both compressed sensing (CS) and parallel imaging can reduce the number of phase-encode lines without significantly compromising the quality of the images, but both have their limitations. The combined method proposed in [1] is designed to mitigate the limitations of applying CS or parallel imaging alone.

CS, introduced in [2–4], leverages sparsity, incoherent sampling, and nonlinear reconstruction to recover a signal from undersampled data. In [5], an anatomical brain image is reconstructed from undersampled data using CS, using the sparsity of the discrete wavelet transform (DWT) of the image. Similarly, CS is used in [6] to reconstruct an axial brain image sparse in the finite-differences domain. However, when the image is not sufficiently sparse, details and contrast may be lost. Also, applying CS with uniformly undersampled observations can result in coherent aliasing.

GRAPPA, the parallel imaging method detailed in [7], consists of estimating a kernel from an additional block of non-undersampled lines in the center of k-space (called the ACS lines) and using the kernel to recover the missing lines in k-space in each coil. While analytically constrained only by the number of linearly independent coils available, GRAPPA also is limited to relatively low undersampling factors by noise amplification.

To correct these drawbacks, the L_1 SPIR-iT method described in [8] constructs a joint optimization problem that produces an image with improved SNR and contrast, while also accommodating arbitrary k-space sampling patterns. In [1], CS and GRAPPA are combined similarly, using the nullspace method described below and GRAPPA to achieve similar results with lower algorithmic complexity. Both methods can be interpreted as enabling the user to trade-off improved SNR from CS and improved contrast from parallel imaging. This work revisits the latter method, extending it to random undersampling, and comparing the reconstruction quality using two-dimensional uniform and random undersampling.

2. THEORY

Consider the full-FOV Cartesian k-space $\mathbf{y}$ and observations $\mathbf{d} = \mathbf{K}\mathbf{y}$, where the $\mathbf{K}$ matrix represents undersampling k-space, and suppose the image $\mathbf{F}^{-1}\mathbf{y}$ ($\mathbf{F}$ is the DFT operator) is sparse in some domain with forward transform $\mathbf{\Psi}$. Then, assuming $\mathbf{d}$ is not noisy, the CS optimization is

$$\arg\min_{\mathbf{y}} \|\mathbf{\Psi}\mathbf{F}^{-1}\mathbf{y}\|_S \quad \text{s.t.} \quad \mathbf{d} = \mathbf{K}\mathbf{y}, \tag{1}$$

where $\|\cdot\|_S$ is an approximation to the ℓ_0 penalty function like the ℓ_1-norm or nonconvex Cauchy function $\|\mathbf{w}\|_S = \sum_n \log(1 + w_n^2)$ (the latter is used in this work). Complex data can be accommodated by enforcing sparsity on $|w_n|$. The CS optimization is extended to P coils by assuming each coil should have the same sparsity pattern, and taking the ℓ_2-norm across the sparse transforms for each coil before evaluating $\|\cdot\|_S$.

The nullspace method for solving this optimization problem consists of converting the constrained problem to an unconstrained problem operating in the nullspace of $\mathbf{K}$. Let $\overline{\mathbf{K}}$ be the subsampling matrix selecting the un-acquired data $\mathbf{x}$ from $\mathbf{y}$; then $\overline{\mathbf{K}}^T\mathbf{K} = \mathbf{0}$.

Funding acknowledgements: NSF CAREER Grant CCF-0643836, NIH R01 EB007942 and EB006847, NIH NCRR P41 RR014075, and a NSF Graduate Research Fellowship. The authors also thank Prof. Fa-Hsuan Lin of the A. A. Martinos Center for the use of his B_1 simulator.

The constrained formulation becomes unconstrained:

$$\arg\min_{\mathbf{x}} \|\mathbf{\Psi}\mathbf{F}^{-1}(\overline{\mathbf{K}}^T\mathbf{x} + \mathbf{K}^T\mathbf{d})\|_S. \quad (2)$$

This formulation guarantees that the acquired k-space data remains unaltered, while enforcing sparsity in the desired transform domain.

As a post-processing step, we linearly combine the full-FOV coil images. Motivated by the coil noise covariance $\mathbf{\Lambda}$ and multiple-coil SNR derived in [9], these coil combination weights $\mathbf{C}$ are computed from the coil sensitivity profiles so the signal gain equals one:

$$\begin{bmatrix} C_1(x,y,z) \\ \vdots \\ C_P(x,y,z) \end{bmatrix}^T = \text{pinv}\left(\mathbf{\Lambda}^{-1/2} \begin{bmatrix} S_1(x,y,z) \\ \vdots \\ S_P(x,y,z) \end{bmatrix} \right) \mathbf{\Lambda}^{-1/2}, \quad (3)$$

where $C_p(x,y,z)$ is the weight for the pixel (x,y,z) in coil p, $S_p(x,y,z)$ is the corresponding pixel in the sensitivity profile of the pth coil, and $\text{pinv}(\cdot)$ is the Moore-Penrose pseudoinverse. These coil sensitivity profiles are estimated from the low-resolution image generated by the ACS lines (the lines used to form the GRAPPA kernel), apodized with a Blackman window to reduce ringing.

Now, we add a GRAPPA fidelity term, which penalizes the solution for differing from the GRAPPA result. Rather than take the ℓ_2-norm of the k-space difference directly, we weight each pixel in the difference image by its contribution to the final combined image, since each coil's contribution to a pixel in the final image is different. Suppose $\mathbf{G}(\mathbf{d})$ is the full k-space GRAPPA reconstruction based on data $\mathbf{d}$; then, the optimization problem combining CS and GRAPPA is

$$\arg\min_{\mathbf{x}} \|\mathbf{C}\mathbf{F}^{-1}(\overline{\mathbf{K}}^T\mathbf{x} + \mathbf{K}^T\mathbf{d} - \mathbf{G}(\mathbf{d}))\|_2^2 + \lambda\|\mathbf{\Psi}\mathbf{F}^{-1}(\overline{\mathbf{K}}^T\mathbf{x} + \mathbf{K}^T\mathbf{d})\|_S. \quad (4)$$

Note that the coil combination weights are used only to weight, not to combine, the pixels of the difference image before computing the ℓ_2-norm for GRAPPA fidelity. Also, λ is a tuning parameter representing the emphasis on sparsity. Larger λ results in a solution closer to the solution of (2).

While nonconvex, (4) can be approximately solved using iteratively reweighted least squares (IRLS), which is described in [10, 11], and a conjugate gradient algorithm. To further improve the sparsity of the result, the optimization problem is repeated several times, each time tightening the Cauchy error function to more closely resemble the ℓ_0 penalty function. For a given repetition, the penalty function (parameterized by α) is

$$\|\mathbf{w}\|_{S(\alpha)} = \sum_n \frac{1}{\log(1+\alpha)} \log(1 + \alpha|w_n|^2). \quad (5)$$

As $\alpha \to \infty$, $\|\mathbf{w}\|_{S(\alpha)} \to \|\mathbf{w}\|_0$. This repetition can be viewed as a form of homotopy continuation. We name this method SpRING, for Sparse Reconstruction of Images using the Nullspace method and GRAPPA. For this work, we begin with $\alpha = 1$ and increase α by a factor of ten whenever the objective function in (4) decreases by less than 1%, up to a maximum value of $\alpha = 10^7$ for the Shepp-Logan phantom and $\alpha = 10^4$ for the real brain data.

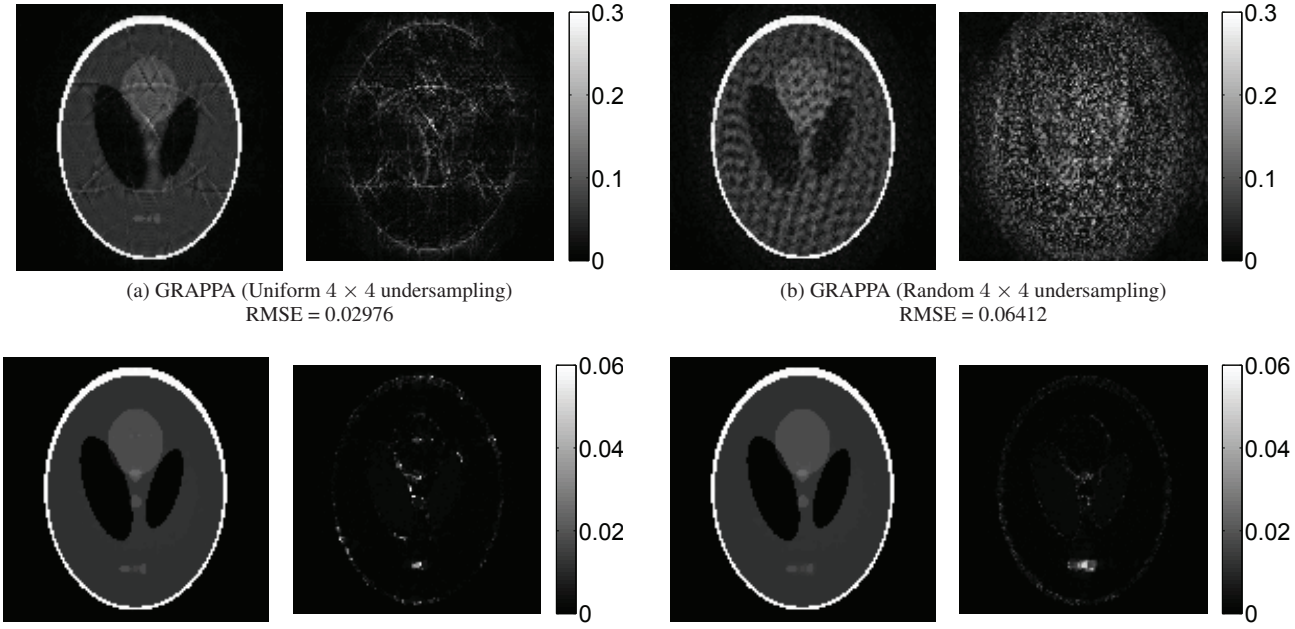
GRAPPA is well-suited for uniform Cartesian undersampling of k-space, since only one kernel needs to be estimated, and applying the kernel becomes convolution in the frequency domain. In this work, GRAPPA is extended to arbitrary subsampling on a grid by using many kernels, one for each block (of fixed size) of missing points. SpRING is extended to random sampling simply by using this GRAPPA result as $\mathbf{G}(\mathbf{d})$ in (4); the CS part already handles arbitrary Cartesian subsampling. This variant of GRAPPA is less computationally efficient because reconstruction is no longer implemented via convolution. In addition, such a large family of kernels is difficult to estimate from a small number of ACS lines, so this method is generally less accurate. Nevertheless, coherent aliasing is still greatly mitigated, so SpRING can reduce the remaining incoherent aliasing.

When compared to L_1 SPIR-iT, this method possesses two distinct computational advantages. By assuming the data is relatively noiseless, a tuning parameter for data fidelity, which must be determined empirically from the quality of the observations, is eliminated. In addition, whereas a single conjugate gradient step of SpRING requires for each coil two FFT's, one $\mathbf{\Psi}$ times a vector, one $\mathbf{\Psi}^H$ times a vector, and two element-wise vector multiplications, each conjugate gradient step of the L_1 SPIR-iT algorithm additionally requires convolution with both the SPIR-iT kernel (similar to the GRAPPA kernel) and its conjugate-transpose.

3. METHODS

Two sets of data are used in the experiments that follow: a simulated multiple-coil Shepp-Logan phantom and a real axial slice of an anonymous brain previously acquired from a consented subject. Simulated data for a 128×128 pixel Shepp-Logan phantom is generated for an 8-channel coil with the aid of the Biot-Savart B_1 simulator available online at `http://www.nmr.mgh.harvard.edu/~fhlin/`, and complex noise is added to each coil image, according to the coil noise covariance matrix derived in [9]. The real data was acquired using an eight-minute T_1-weighted MPRAGE ($256 \times 256 \times 176$ sagittal slices, 1.0 mm isotropic) sequence (no acceleration) on a 3 T Siemens system with a 32-channel Siemens head coil. An axial slice is extracted from this full-FOV 3-D dataset and cropped to 148×236 lines (still full-FOV). The fully sampled (readout) direction is assumed perpendicular to the axial slice plane, which leaves two directions in the plane for acceleration. The noise covariance $\mathbf{\Lambda}$ was measured for the real data from a separate SNR map acquired before collecting the anatomical data. Since we are only interested in an axial slice of the brain, which is near only a small fraction of the 32 coils, we keep k-space data for only the ten coils with significant signal in them. This way, the coherent aliasing present in the GRAPPA result's difference image will be more visible over the amplified noise. When processing the Shepp-Logan phantom, a finite-differences sparsifying transform is used; a 4-level '9-7' DWT is used for the real brain data.

The full-FOV slice is then undersampled in both directions in k-space, and a small set of ACS lines are also extracted from the reference k-space data. The coil combination weights are then computed from apodized ACS lines. When choosing a random sampling pattern, we follow the Monte Carlo method described in [5] to ensure the pattern is sufficiently incoherent. Finally, GRAPPA is performed using the ACS lines and the undersampled data, and SpRING is performed using the GRAPPA full-FOV k-space result, as well as the undersampled k-space data and the coil combination weights. The ACS lines are treated as additional data in both GRAPPA and SpRING reconstructions. SpRING is repeated for λ logarithmically spaced between 10^{-5} and 10^6. The difference image of the magnitude of the result and the magnitude of the reference image is used for qualitative evaluation, and the root mean squared error (RMSE) of the magnitude image is used for quantitative comparison.



(a) GRAPPA (Uniform 4×4 undersampling)
RMSE = 0.02976

(b) GRAPPA (Random 4×4 undersampling)
RMSE = 0.06412

(c) SpRING (Uniform 4×4 undersampling; $\lambda = 10^{1.4}$)
RMSE = 0.002949

(d) SpRING (Random 4×4 undersampling; $\lambda = 10^{3}$)
RMSE = 0.002386

Fig. 1. Shepp-Logan phantom reconstructions and difference images for uniform and random undersampling.

4. RESULTS

In this first experiment, the Shepp-Logan phantom is 4×4 (factor of 4 in both directions) uniform and random Cartesian undersampled, and a 30×30 block at the center of the image is used as ACS lines for estimating the coil combination weights and GRAPPA kernel. In Figure 1, GRAPPA and SpRING are applied to both uniformly and randomly undersampled 8-channel Shepp-Logan phantom data. In both cases, the difference images suggest that the SpRING method is able to mitigate much of the error in the GRAPPA results, since the residual is not sparse in the finite differences domain.

To compare uniform and random undersampling, we first consider GRAPPA alone. When the undersampling factor exceeds the number of linearly independent coils, the GRAPPA kernel cannot undo all the aliasing. In the uniform case, the error is consistent across all frequencies, yielding the coherent aliasing visible in the difference image in Figure 1a. In the random undersampling case, the error differs for different kernels, so each frequency block is aliased differently, and Figure 1b features incoherent aliasing. When evaluating the combined method, SpRING achieves a lower RMSE with random Cartesian sampling than with uniform sampling. Thus, even though the GRAPPA result may look less useful in the random sampling case, CS is better suited for mitigating the incoherent noise in Figure 1b than the coherent aliasing in Figure 1a.

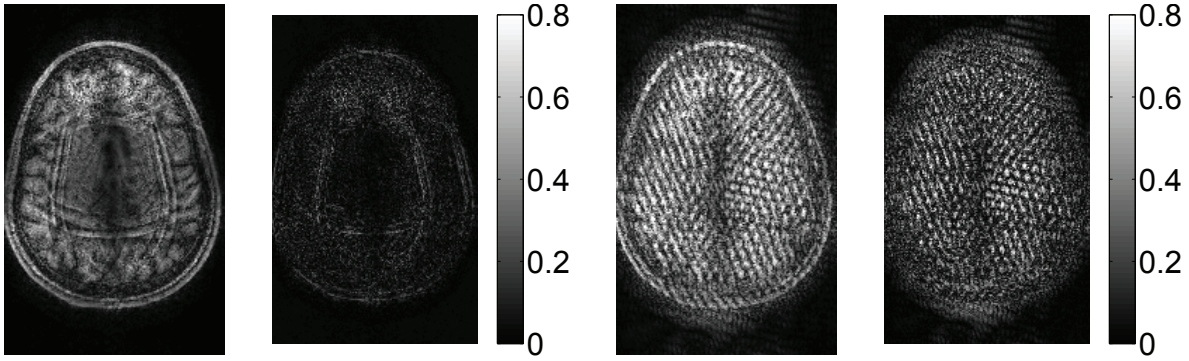
The 8-channel Shepp-Logan phantom is very sparse in the finite-differences domain (90% of the energy in 4% of the coefficients), so while it illustrates the advantages of including sparsity on the reconstruction performance, conclusions about the relative efficacy of the algorithms under uniform and random undersampling may not continue to hold for real images, which are only approximately sparse. In the next experiment, the real axial brain data (10 of 32 channels) is 5×5 uniform and random Cartesian undersampled, and a 36×36 fully sampled block at the center functions as ACS lines for estimating the coil combination weights and GRAPPA kernel. The results for uniform and random subsampling are shown in Figure 2.

In the uniform undersampling case, GRAPPA amplifies the noise in the data and introduces coherent aliasing due to the high oversampling factor. While the former is easily mitigated by the SpRING method, the latter is not; the coherent aliasing clearly remains in both the result and difference images in Figure 2c. For the random undersampling case, the same behavior is observed as in the Shepp-Logan phantom case; incoherent aliasing is introduced by random undersampling affecting different k-space frequencies differently. SpRING mitigates both the incoherent artifacts introduced by GRAPPA as well as the amplified noise, leaving behind only the error due to approximate sparsity. In applications where this error is unsuitable, employing a more appropriate sparsifying transform may reduce the error substantially.

5. CONCLUSION

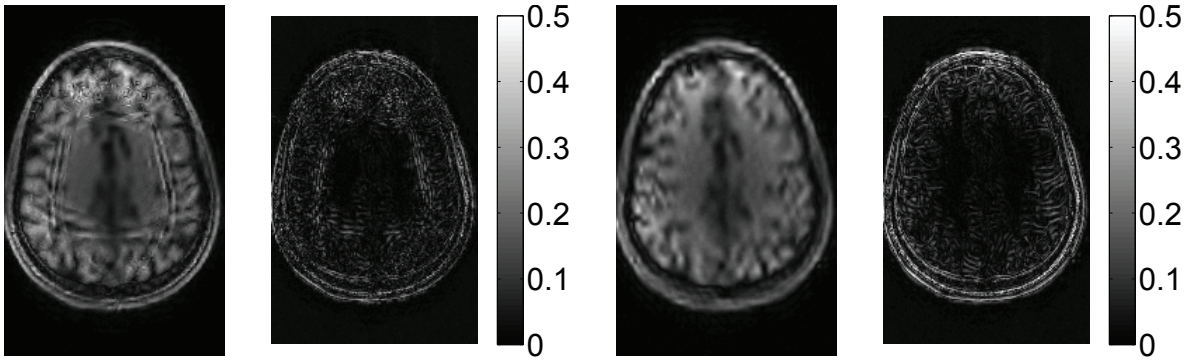
Clearly, the SpRING method greatly improves upon the GRAPPA result when the undersampling is high enough that parallel imaging alone fails. From the results, two valuable conclusions can be drawn about choosing a sampling pattern when combining CS and GRAPPA. The benefits of random undersampling to CS far outweigh the benefits of uniform undersampling to GRAPPA, especially at high undersampling factors, when GRAPPA results in residual coherent aliasing. However, when the image is only approximately sparse, the level of sparsity limits the effectiveness of SpRING. Thus, the sparsifying transform is just as important to the performance of SpRING as the sampling pattern. Repeating the analysis of the real data with a more effective sparsifying transform may yield additional insights.

The same analysis can be repeated for variable-density random



(a) GRAPPA (Uniform 5×5 undersampling)
RMSE = 0.08507

(b) GRAPPA (Random 5×5 undersampling)
RMSE = 0.2162



(c) SpRING (Uniform 5×5 undersampling; $\lambda = 10^{0.6}$)
RMSE = 0.06551

(d) SpRING (Random 5×5 undersampling; $\lambda = 10^{2.2}$)
RMSE = 0.07977

Fig. 2. T_1-weighted axial brain image reconstructions and difference images for uniform and random undersampling.

or Poisson-disc undersampling, as used in [5, 8]. Also, computing g-factors for GRAPPA and SpRING for different sampling strategies may provide insights into how sampling patterns affect noise amplification more generally. In addition, measuring the prevalence of artifacts in the GRAPPA results may yield a novel approach towards optimizing the choice of tuning parameter λ for a given sampling pattern; further study connecting the sampling pattern with the type of noise in the GRAPPA result is warranted.

6. REFERENCES

[1] D. S. Weller, J. R. Polimeni, L. J. Grady, L. L. Wald, E. Adalsteinsson, and V. K. Goyal, "Combining nonconvex compressed sensing and GRAPPA using the nullspace method," in *18th Annual Meeting of ISMRM.*, May 2010, p. 4880.

[2] 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, Feb. 2006.

[3] E. J. Candès and J. Romberg, "Quantitative robust uncertainty principles and optimally sparse decompositions," *Found. Comput. Math.*, vol. 6, no. 2, pp. 227–254, Apr. 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 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, Dec. 2007.

[6] L. Grady and J. R. Polimeni, "Nullspace compressed sensing for accelerated imaging," in *17th Annual Meeting of ISMRM.*, Apr. 2009, p. 2820.

[7] 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, Jun. 2002.

[8] M. Lustig, M. Alley, S. S. Vasanawala, D. L. Donoho, and J. M. Pauly, "L_1 SPIR-iT: Autocalibrating parallel imaging compressed sensing," in *17th Annual Meeting of ISMRM.*, Apr. 2009, p. 379.

[9] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR phased array," *Magn. Reson. Med.*, vol. 16, no. 2, pp. 192–225, Nov. 1990.

[10] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. in Statist. - Theory and Methods.*, vol. 6, no. 9, pp. 813–827, 1977.

[11] A. C. Zelinski, V. K. Goyal, and E. Adalsteinsson, "Simultaneously sparse solutions to linear inverse problems with multiple system matrices and a single observation vector," *SIAM J. Sci. Comput.*, vol. 31, no. 6, pp. 4533–4579, Jan. 2010.