

December 2012

A Study of Nonlinear Approaches to Parallel Magnetic Resonance Imaging

Yuchou Chang

University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>



Part of the [Electrical and Electronics Commons](#)

Recommended Citation

Chang, Yuchou, "A Study of Nonlinear Approaches to Parallel Magnetic Resonance Imaging" (2012). *Theses and Dissertations*. 195.
<https://dc.uwm.edu/etd/195>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

A STUDY OF NONLINEAR APPROACHES TO PARALLEL MAGNETIC RESONANCE IMAGING

by

Yuchou Chang

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Engineering

at

The University of Wisconsin-Milwaukee

December 2012

ABSTRACT

A STUDY OF NONLINEAR APPROACHES TO PARALLEL MAGNETIC RESONANCE IMAGING

by

Yuchou Chang

The University of Wisconsin-Milwaukee, 2012

Under the Supervision of Professor Jun Zhang

Magnetic resonance imaging (MRI) has revolutionized radiology in the past four decades by its ability to visualize not only the detailed anatomical structures, but also function and metabolism information. A major limitation with MRI is its low imaging speed, which makes it difficult to image the moving objects. Parallel MRI (pMRI) is an emerging technique to increase the speed of MRI. It acquires the MRI data from multiple coils simultaneously such that fast imaging can be achieved by reducing the amount of data acquired in each coil. Several methods have developed to reconstruct the original image using the reduced data from multiple coils

based on their distinct spatial sensitivities. Among the existing methods, Sensitivity Encoding (SENSE) and GeneRally Autocalibrating Partially Parallel Acquisition (GRAPPA) are commercially used reconstruction methods for parallel MRI. Both methods use linear approaches for image reconstruction. GRAPPA is known to outperform SENSE because no coil sensitivities are needed in reconstruction. However, GRAPPA can only accelerate the speed by a factor of 2-3. The objective of this dissertation is to develop novel techniques to significantly improve the acceleration factor upon the existing GRAPPA methods. Motivated by the success of recent study in our group which has demonstrated the benefit of nonlinear approaches for SENSE, in this dissertation, nonlinear approaches are studied for GRAPPA. Based on the fact that GRAPPA needs a calibration step before reconstruction, nonlinear models are investigated in both calibration and reconstruction using a kernel method widely used in machine learning. In addition, compressed sensing (CS), a nonlinear optimization technique will also be incorporated for even higher accelerations. In order to reduce the computation time, a nonlinear approach is proposed to reduce the effective number of coils in reconstruction. The imaging speed is expected to improve by a factor of 4-6 using the proposed nonlinear techniques. These new

techniques will find many applications in accurate brain imaging, dynamic cardiac imaging, functional imaging, and so forth.

©Copyright by Yuchou Chang, 2012

All Rights Reserved

ACKNOWLEDGEMENT

I would never have been able to complete my dissertation without the guidance of my research committee members, support from my family, and help from my friends.

I would like to express my deepest gratitude to my advisor Prof. Leslie Ying for her excellent guidance and patience on my projects and research. I would like to thank Prof. Jun Zhang who taught me the research in the field of random signal processing and practical issues. I also thank Dr. Kevin F. King for helping me to develop the technique of combining the compressed sensing and parallel imaging for fast imaging. Furthermore, I appreciate Prof. Guangwu Xu, Prof. Yi Hu , and Prof. Zeyun Yu for their precious advice on my preliminary exam and my dissertation.

I would also like to thank my parents. They were always supporting me and encouraging me with their best wishes.

I wish to express sincere appreciation to my friends Dr. Dong Liang, Dr. Yong Wang, Haifeng Wang, Yihang Zhou, Christopher Baker, Gajanan Nagarsekar, and Huajun She for their discussion and idea sharing on my research.

TABLE OF CONTENT

ABSTRACT	II
COPYRIGHT	V
ACKNOWLEDGEMENT.....	VI
TABLE OF CONTENT	VII
LIST OF FIGURES	IX
LIST OF TABLES	XII
LIST OF ABBREVIATIONS.....	XIII
1. INTRODUCTION.....	1
1.1 MAGNETIC RESONANCE IMAGING	1
1.1.1 <i>General Introduction</i>	1
1.1.2 <i>K-Space and Image Reconstruction</i>	2
1.1.3 <i>Image Quality</i>	5
1.2 PARALLEL MRI	6
1.2.1 <i>Parallel Imaging</i>	6
1.2.2 <i>Image-Based Reconstruction Methods</i>	9
1.2.3 <i>K-Space-Based Reconstruction Methods</i>	10
1.3 COMPRESSED SENSING MRI.....	12
1.4 MOTIVATION AND ORGANIZATION	14
2. PROBLEM STATEMENT AND CONTRIBUTION OF THE DISSERTATION	16
2.1 PROBLEM STATEMENT.....	16
2.1.1 <i>Low SNR with High Imaging Acceleration</i>	16
2.1.2 <i>Imaging Acceleration Is Not High Enough</i>	19
2.1.3 <i>Computation Cost of Reconstruction Is Large with Many Coils</i>	20
2.2 DISSERTATION CONTRIBUTION.....	21
3. NONLINEAR GRAPPA - A KERNEL APPROACH TO PARALLEL MRI RECONSTRUCTION.....	23
3.1 ERRORS-IN-VARIABLES MODEL OF GRAPPA	25
3.2 PROPOSED NONLINEAR GRAPPA	28

3.2.1 General Formulation Using Kernel Method	29
3.2.2 Choice of nonlinear mapping $\Phi(\cdot)$	31
3.2.3 Explicit implementation of nonlinear GRAPPA	33
3.3 EXPERIMENTAL RESULTS	37
3.3.1 Experiment Settings	37
3.3.2 Results of Phantom and In Vivo Datasets	38
3.4 DISCUSSION	45
4. ACCELERATING NONLINEAR GRAPPA BY COMPRESSED SENSING	50
4.1 COMBINATION METHODS OF CS AND PMRI	50
4.2 SERIAL COMBINATION OF CS AND NLGRAPPA	52
4.2.1 Sampling Pattern	53
4.2.2 Sequential Reconstruction	55
4.3 EXPERIMENTAL RESULTS	57
4.3.1 Experiment Settings	57
4.3.2 Results of Phantom and In Vivo Datasets	58
4.4 DISCUSSION	64
5. NONLINEAR CHANNEL REDUCTION FOR EFFICIENT GRAPPA RECONSTRUCTION	67
5.1 REVIEW OF CHANNEL COMPRESSION	67
5.2 THE PROPOSED METHOD: KERNEL PCA-BASED GRAPPA	70
5.2.1 PCA-based GRAPPA	70
5.2.2 Proposed Kernel PCA-based GRAPPA	73
5.3 ANALYSIS OF COMPUTATIONAL EXPENSE	75
5.4 EXPERIMENTAL RESULTS	78
5.4.1 Experiment Settings	78
5.4.2 Results of Phantom and In Vivo Datasets	79
5.5 DISCUSSION	85
6. CONCLUSION	88
7. REFERENCE	90
8. CURRICULUM VITAE	101

LIST OF FIGURES

Page 2: Figure 1. Demonstration of a MRI scanner (picture acquired from GE MRI scanner, GE Healthcare, Waukesha, WI)

Page 3: Figure 2. A k -space data to an image requires using a 2D inverse Fourier transform. Low frequency signals and high frequency signals co-exist in k -space.

Page 8: Figure 3. (a) Fully sampled k -space provided by conventional acquisition, resulting in a full FOV image after inverse Fourier transform. (b) Undersampled acquisition ($R=2$), resulting in a reduced FOV ($\text{FOV}/2$) with aliasing artifacts.

Page 9: Figure 4. Work flow of SENSE reconstruction

Page 12: Figure 5. GRAPPA reconstruction demonstration

Page 18: Figure 6. Lower SNR as increasing outer reduction factor. The reconstruction parameters of 8-channel brain data are 40 ACS lines, coefficient size of convolution is 4×5 .

Page 28: Figure 7. Nonlinearity of the bias for GRAPPA coefficients as a function of noise in ACS data.

Page 31: Figure 8. Illustration of the calibration procedure for GRAPPA and nonlinear (NL) GRAPPA.

Page 36: Figure 9. Phantom images reconstructed from an eight-channel dataset with an ORF of 6 and 38 ACS lines (denoted as 6-38 on the right corner of each image). With the SOS reconstruction as the reference, the proposed nonlinear GRAPPA (NL-GRAPPA) method is compared with conventional GRAPPA, regularized GRAPPA, and IRLS methods. The corresponding difference images with the reference ($7\times$ amplification) and g-factor maps are also shown on the right two columns respectively.

Page 39: Figure 10. Axial brain images reconstructed from a set of eight-channel data with an ORF of 5 and 48 ACS lines using GRAPPA, regularized GRAPPA, IRLS, and the proposed nonlinear method. The corresponding difference images with the reference ($5\times$ amplification) are shown on the middle column and g-factor maps on the right column.

Page 40: Figure 11. Sagittal brain images reconstructed from a set of eight-channel data with an ORF 5 and 48 ACS lines and their corresponding difference images on the right. The proposed nonlinear GRAPPA suppresses most noise without aliasing artifacts.

Page 41: Figure 12. Results from the four-channel cardiac dataset with an ORF 5 and 48 ACS lines. The reconstructed images, zoomed ROI, difference images, and g-factor maps are shown from left to right respectively. They show that the proposed method can remove more noise than other methods while still preserving the resolution.

Page 43: Figure 13. NMSE curves of the proposed method as a function of the number of the second-order terms using the long-axis cardiac dataset. The “U” shape of the curve suggests that some intermediate number should be chosen.

Page 44: Figure 14. The nonlinear GRAPPA reconstructions with an increasing number of the second-order terms show that the noise is gradually removed but artifacts gradually increase. In the extreme case when all second-order terms are included, both blurring and aliasing artifacts are serious.

Page 46: Figure 15. Comparison between GRAPPA and nonlinear GRAPPA when ORF increases with fixed ACS lines. Contrary to GRAPPA, noise in the proposed method does not increase with the ORF.

Page 47: Figure 16. Comparison of GRAPPA and nonlinear GRAPPA reconstructions when ACS increases with fixed ORF. It shows nonlinear GRAPPA needs more ACS lines than GRAPPA to avoid aliasing artifacts, but GRAPPA has more noise than nonlinear GRAPPA.

Page 48: Figure 17. Comparison between GRAPPA and nonlinear GRAPPA when different numbers of columns are chosen for the coefficients. Contrary to GRAPPA, the use of more columns in nonlinear GRAPPA can suppress more noise in reconstruction.

Page 54: Figure 18. Sampling pattern and CS reconstruction demonstration of the proposed method.

Page 59: Figure 19. With (a) the SoS reconstruction of the 8-channel phantom image as the reference, we compare (b) CS-GRAPPA and (c) the proposed method at a net reduction factor of 4. Subfigures (d) – (f) show zoomed-in regions of (a) – (c) respectively.

Page 60: Figure 20. With (a) the SoS reconstruction of the 8-channel brain image as the reference, we compare (b) CS-GRAPPA and (c) the proposed method at a net reduction factor of 3.93.

Page 60: Figure 21. Comparison of (a) SoS, (b) CS-GRAPPA and (c) the proposed method in reconstructing the four-channel cardiac image at a net reduction of 4.12.

Page 63: Figure 22. The performance comparison between SPIRiT and the proposed CS-NLGRAPPA method. Reconstructions of the phantom and the brain data sets are presented at net reduction factor 3.0 and 4.0, respectively.

Page 65: Figure 23. CS-NLGRAPPA reconstructions with three different combinations of ACS and R_{CS} with fixe ORF of 4 and a net reduction factor of 3.93.

Page 66: Figure 24. CS-NLGRAPPA reconstructions with six different combinations of RCS with various ORF used the fixed ACS.

Page 73: Figure 25. Demonstration of PCA and KPCA roles in channel reduction for GRAPPA reconstruction.

Page 79: Figure 26. For the 32-channel phantom dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively. Their difference maps are also given for performance comparison.

Page 80: Figure 27. For the 32-channel brain (axial) dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively. Their difference maps are also given for performance comparison.

Page 82: Figure 28. For the 32-channel brain (coronary) dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively.

Page 83: Figure 29. Quantitative comparison among GRAPPA, PCA-GRAPPA, and KPCA-GRAPPA at ourter reduction factors 4, 5, and 6.

Page 87: Figure 30. Comparison of sensitivity maps among original channels, reduced channels by PCA, reduced channels by KPCA with different parameters.

LIST OF TABLES

Page 62: Table 1. Comparison of computational time between CS-GRAPPA and CS-NLGRAPPA.

Page 84: Table 2. (a) reconstruction times for 32-channel phantom dataset using different combinations of source and target channels, (b) reconstruction times for 32-channel brain dataset using different combinations of source and target channels, (c) reconstruction times for 32-channel brain dataset using different convolution columns.

Page 85: Table 3. The computational costs via CPU time of PCA and KPCA at reduced 8, 12, 16, 20, and 24 channels.

LIST OF ABBREVIATIONS

NMR: Nuclear Magnetic Resonance

MRI: Magnetic Resonance Imaging

pMRI: Parallel Magnetic Resonance Imaging

PI: Parallel Imaging

RF: Radio Frequency

TR: Repetition Time

FFT: Fast Fourier Transform

GRAPPA: GeneRalized Autocalibrating Partially Parallel Acquisitions

SENSE: SENSitivity Encoding

PE: Phase Encoding

FE: Frequency Encoding

SNR: Signal-to-Noise Ratio

PCA: Principle Component Analysis

KPCA: Kernel Principle Component Analysis

TV: Total Variation

CS: Compressed Sensing

NLGRAPPA: NonLinear GRAPPA

IRLS: Iterative Reweighted Least Squares

CT: Computed Tomography

1. Introduction

1.1 Magnetic Resonance Imaging

1.1.1 General Introduction

Since the first successful nuclear magnetic resonance (NMR) experiments were performed in 1946, many scientists over the next more than a half-century developed MR Imaging into the technology that we now know today. Magnetic Resonance Imaging (MRI) has increasingly become one of the preferred imaging modalities in modern times to examine the structures inside the body, which enables physicians to visualize differences among soft tissue with incredible acuity [1].

MRI provides an imaging technique which is practically non-invasive and without any ionizing radiation. In comparison with other imaging modalities including X-ray, ultrasound, or computed tomography (CT) scan, MRI provides more information about structures inside the body in many cases. Unlike other imaging techniques, MRI uses a strong magnetic field and radio waves to create computerized images of the body that include head, chest, blood vessels, abdomen, pelvis, bones, joints, spine, and so forth. A commercial MRI scanner is demonstrated in Figure 1. It is used to identify problems such as tumors, bleeding, injury, blood vessel diseases, or infection. Furthermore, contrast agents may be used during MRI to show abnormal tissue more clearly.



Figure 1. Demonstration of a MRI scanner (picture acquired from GE MRI scanner, GE Healthcare, Waukesha, Wisconsin)

1.1.2 K-Space and Image Reconstruction

The readout MR signals that include radio frequency (RF) waves with different amplitudes, frequencies, and phases contain spatial information of the structures inside the body. These MR signals are digitized, and then written into a data matrix called k -space, which is the 2D or 3D Fourier transform of the MR image measured. An inverse Fourier transform is applied on k -space to produce an image [2, 3], as shown in Figure 2. In k -space, k_x -coordinates represent frequency-encoding (FE) direction (horizontal direction), and k_y -coordinates denote phase-encoding (PE) direction (vertical direction). Low frequency signals locate near the center of k -space, while high frequency signals locate near the periphery of k -space. As shown in Figure 2, after two square windows

with different sizes are applied on k -space to truncate the data, the corresponding reconstructed images display different spatial resolutions. The reconstructed image with a smaller square window for data truncation is blurred due to loss of high frequency signals near the periphery of k -space. After incorporating more high frequency signals with a larger square window, resolution of the reconstructed image is improved.

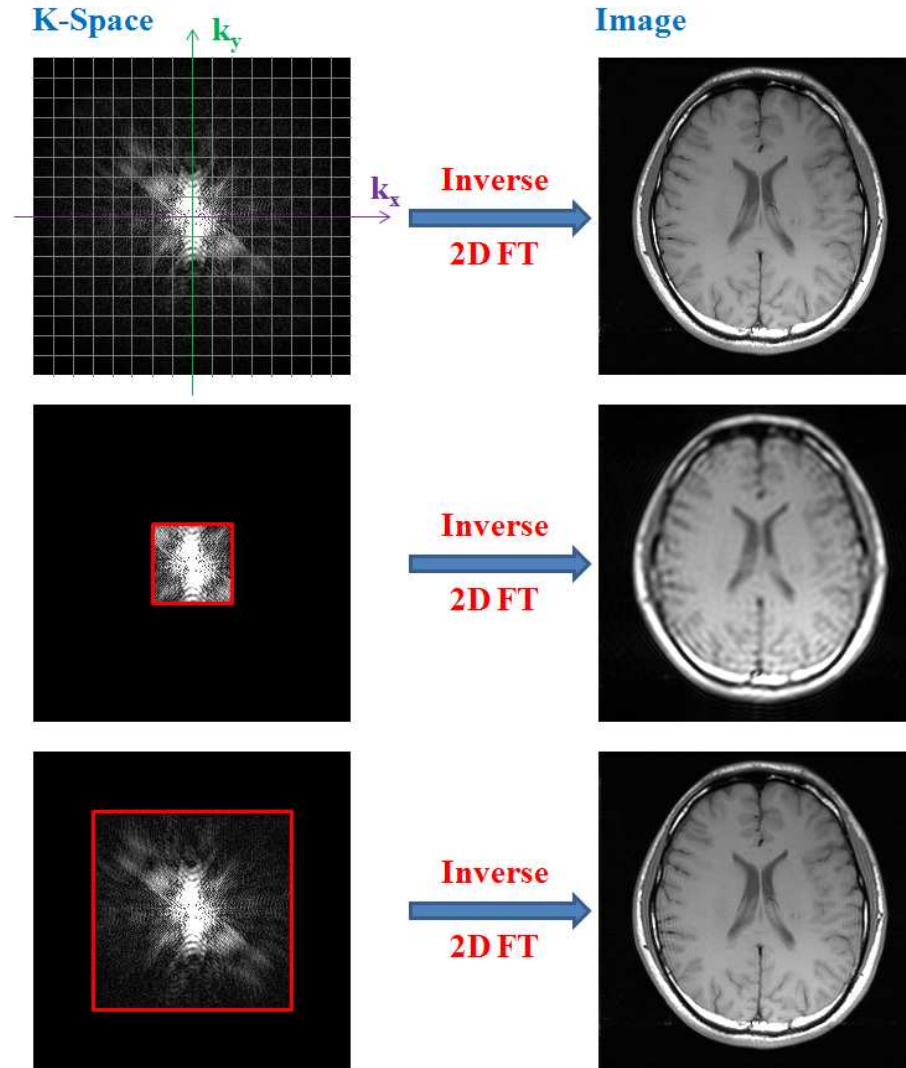


Figure 2. A k -space data to an image requires using a 2D inverse Fourier transform. k -space contains both low frequency signals and high frequency signals.

Furthermore, the differences in the spatial, temporal, and resolution of the MR images depend on how MR signals are mapped into the k -space. Three gradients: slice selection (G_z), frequency encoding (G_x), and phase encoding (G_y) are used to spatially encode the signals in MRI. Each gradient is characterized by its strength (magnetic field variation for the same unit of distance), duration of time, and direction. The net strength and duration of the frequency-encoding gradient G_x and phase-encoding gradient G_y decide the locations (k_x and k_y coordinates) of data on k -space:

- 1) A low-amplitude or short-duration gradient encodes low frequency information; this kind of information is generally mapped at round the center of k -space.
- 2) A high-amplitude of long-duration gradient encodes high frequency information; this kind of information is mapped to the periphery of k -space.

As shown in Figure 2, the greater the net strength of the phase-encoding gradient G_y (or the longer the gradient is on), the farther from the k -space center the data belong, in the upper direction if the gradient is positive or in the lower direction if the gradient is negative. As the duration of phase-encoding gradient is most often constant, the strength of the phase-encoding gradient governs the location on the vertical axis (k_y -coordinate). On the other hand, the longer the frequency-encoding gradient G_x is on (or the greater the net strength of the gradient is), the farther from the k -space center the data belong, in the right direction if the gradient is positive or in the left direction if the gradient is negative. As the strength of the frequency-encoding gradient is most often constant, the duration of the frequency-encoding gradient governs the location on the horizontal axis (k_x -coordinate).

As we can see, after the entire k -space is collected, inverse Fourier transform can be used to reconstruct the image. Typically, k -space is filled with raw data during the scan, usually one phase-encoding line per TR (repetition time). Between each repetition, there is a change in phase-encoding-gradient strength, corresponding to a change in k_y -coordinate. This allows filling of all the lines of k -space from top to bottom. Therefore, more PE lines are filled, more scanning time is consumed. Full k -space reconstruction can achieve the best quality due to all signals have been sampled and filled on the k -space. Considering the scanning time consumed, some researchers applied partial k -space reconstruction that only undersamples a part of k -space signals to reconstruct the image. Furthermore, MR image reconstruction is ill-posed as an inverse problem [71], so that quality of reconstructed images would be deteriorated. In order to evaluate reconstruction performance, image quality can be measured qualitatively and quantitatively, which will be presented in the next sub-section.

1.1.3 Image Quality

In MRI, there are some issues on image quality. In this section, we will present the concepts of spatial resolution, noise, and artifacts. These measures will be used to evaluate the reconstruction quality in comparison with the proposed methods and existing methods in the dissertation.

Spatial resolution corresponds to the size of the smallest detectable detail. The smaller the voxels are, the higher the spatial resolution will be. Spatial resolution that depends on the properties of the imaging system generally quantifies how close lines can be to each other and still be visibly resolved in content of the MR image. As shown in

Figure 2, different resolutions of the reconstructed images are compared and details are difficult to be resolved in the low resolution image.

Noise presents irregular patterns on the MR image, whose random variations in signal degrade image quality. The sources of MR image noise are very complex that depend on some factors including MR scanner specifications, tissue characteristics, coil design, pulse sequence design, and so forth. The signal-to-noise ratio (SNR) that is equivalent to the ratio of the signal power over the noise power is usually used for measuring noise. Furthermore, g-factor was also used for reconstruction evaluation and its details can be seen in the content of the dissertation.

Artifacts often deteriorate MR Images. Similar to noise, there are many sources of artifacts on MR images, which can be classified into image reconstruction-related artifacts, system-related artifacts, physiology-related artifacts, and so forth. In this dissertation, artifacts are mainly originated from image reconstruction-related artifacts due to limitations intrinsic to the reconstruction algorithm. For example, aliasing artifacts appear by reducing the number of sampled signals on k -space, as shown in Figure 3 (b).

1.2 Parallel MRI

1.2.1 Parallel Imaging

Imaging speed is one of the most important considerations in clinical MRI. In practical acquisition, frequency encoding generally takes only milliseconds, but phase encoding has to be done line by line with appropriate TR elapsing between each line acquisition, which is time consuming and it sometimes takes several minutes depending on the RF sequence. Therefore, as mentioned above, only a portion of k -space data is needed to be

acquired to accelerate imaging speed. Parallel MRI is a technique to increase the speed of MRI acquisition, which can accelerate image acquisition dramatically by extracting spatial information from an array of phased coils simultaneously. Each coil has distinct spatial sensitivities. Fast imaging is achieved by skipping a number of phase-encoding lines and the reduced data from multiple coils are then used to reconstruct the original image. Theories of parallel MRI were developed in late 1980's but the practical implementation was carried out until the end of 1990's. The technique has been widely used in clinical settings since 21st century and its improvement is still undergoing in recent years.

The acquisition time depends on the procedure of filling up the k -space. One way to accelerate this phase encoding step is to reduce the number of phase encoding lines by the acceleration factor R that increases the distance of equidistantly sampled k -space lines. In image space, this kind of undersampling the k -space generates a reduced field of view (FOV) in phase encoding direction associated with fold over artifacts in the coil images presented in Figure 3.

In parallel MRI, undersampling acquisition requires multiple coils to cover the entire FOV. A basic idea of parallel MRI is the following: acquire reduced FOV covering a portion of the entire image by each coil and combine the images from all the coils together to produce full-FOV image. Each element of a phased coil array is related to a separate independent receiver system. So, multiple images can be reconstructed for a given imaging. Roemer et al explored how to combine these images for optimal SNR [4],

whose study shows that, in order to optimally reconstruct array coil images, we need to explicitly know the coil sensitivities at all pixel locations.

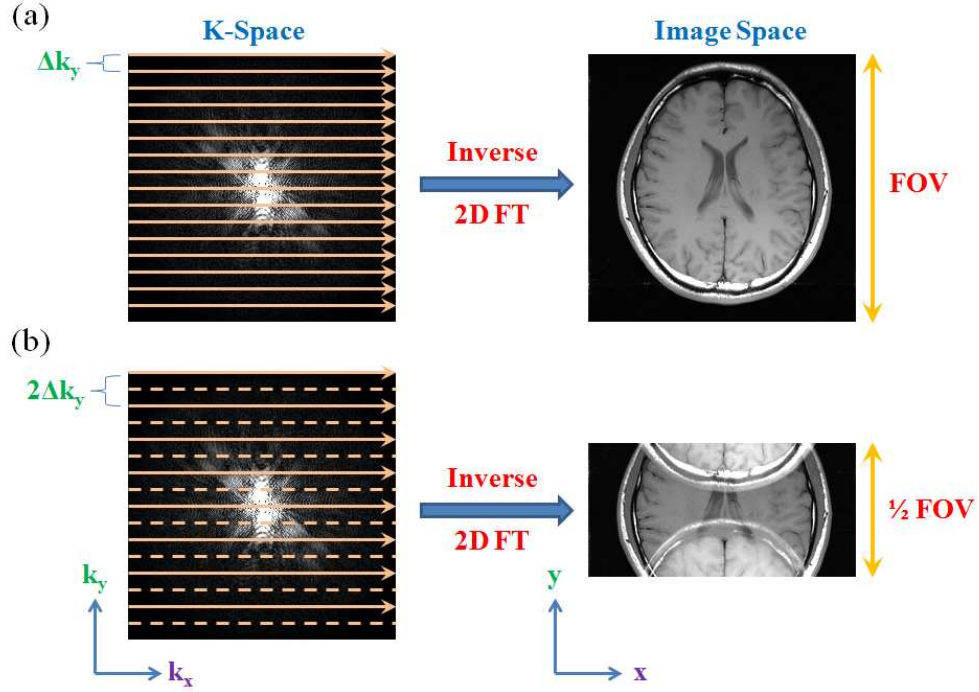


Figure 3. (a) Fully sampled k -space provided by conventional acquisition, resulting in a full FOV image after inverse Fourier transform. (b) Undersampled acquisition ($R=2$), resulting in a reduced FOV (FOV/2) with aliasing artifacts.

Image acquisition is accelerated by the acceleration factor (R -factor). Theoretically, imaging time can be reduced by R -factor equal to the number of array coil elements [5]. However, decreased SNR and increased image non-uniformities limit acceleration factor well below the theoretical maximum R in actual reconstruction methods. Both image-based reconstruction methods and k -space-based reconstruction methods in parallel MRI have this problem.

1.2.2 Image-Based Reconstruction Methods

Reconstruction methods of parallel MRI can be categorized into two groups: image-based parallel imaging reconstruction and k -space-based parallel imaging reconstruction. For image-based parallel imaging reconstruction, the unaliasing process is done after the Fourier transform. To perform the image-based reconstruction, it is necessary to estimate the coil sensitivities. These estimates can be obtained from a separate calibration scan from calibration data, which is contained in the acquisition. The commonly used image-based reconstruction methods contain:

- SENSE [6] (SENSitivity Encoding);
- PILS [7] (Partially Parallel Imaging with Localized Sensitivity).

The typical one of image-based reconstruction methods is SENSE, for which the unaliasing process is done in the image domain. Figure 4 shows the SENSE reconstruction that is done in the image domain after inverse Fourier transforming the k -space data. When the undersampled k -space data is inverse Fourier transformed to the image domain, the field of view (FOV) is reduced by the $1/R$. This means that the same information is contained in a smaller area which leads to fold over type aliasing artifacts.

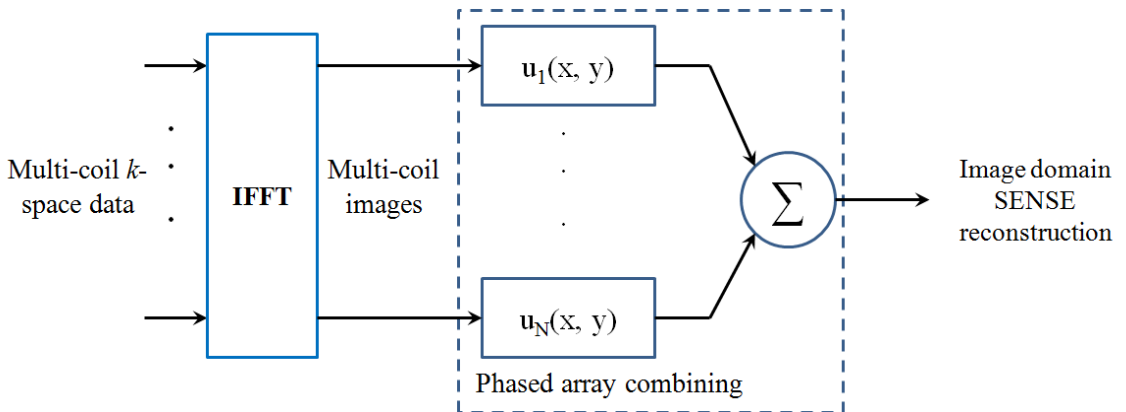


Figure 4. Work flow of SENSE reconstruction

In the following equation, the signal in one pixel at a certain location (x, y) received in the N^{th} component coil image g_N for the case where $R=2$. The g vector represents the complex coil image values including fold over artifacts at the chosen pixel. The matrix S denotes the sensitivities for each coil at the $R=2$. To generate the f vector, the inverse of the S matrix is calculated. This effectively combines the reference data of coil sensitivities and the sub-sampled target data.

$$\begin{bmatrix} g_1(x, y) \\ \vdots \\ g_N(x, y) \end{bmatrix} = \begin{bmatrix} s_1(x, y) & s_1(x, y - \frac{FOV}{2}) \\ \vdots & \vdots \\ s_N(x, y) & s_N(x, y - \frac{FOV}{2}) \end{bmatrix} \begin{bmatrix} f(x, y) \\ \vdots \\ f(x, y - \frac{FOV}{2}) \end{bmatrix} + \begin{bmatrix} n_1(x, y) \\ \vdots \\ n_N(x, y) \end{bmatrix} \quad (1)$$

However, with this sequential processing, any inaccuracy in sensitivity estimation can be propagated to the reconstructed image. Furthermore, sensitivity maps are required to be explicitly estimated for reconstruction.

1.2.3 K-Space-Based Reconstruction Methods

In image-based parallel imaging reconstruction, the unaliasing process is done after the Fourier transform. To perform the image-based reconstruction, it has to estimate the coil sensitivities. These estimates can be obtained from a separate calibration scan from calibration data, which is contained in the acquisition. In contrast, k -space-based reconstruction directly reconstructs missing k -space signals without requiring the prior knowledge of coil sensitivities and then generates images by inverse Fourier transform. The commonly used k -space-based reconstruction methods contain:

- SMASH [8] (SiMultaneous Acquisition of Spatial Harmonics);
- GRAPPA [9] (GeneRalized Autocalibrating Partially Parallel Acquisitions);
- AUTO-SMASH [10] (AUTO-calibration SiMultaneous Acquisition of Spatial Harmonics);
- VD-AUTO-SMASH [11] (VD-AUTO-calibration SiMultaneous Acquisition of Spatial Harmonics).

In conventional GRAPPA, the central k -space of each coil is sampled at the Nyquist rate to obtain ACS data, while the outer k -space is undersampled by some outer reduction factors (ORF). The missing k -space data is estimated by a linear combination of the acquired undersampled data in the neighborhood from all coils, which can be represented mathematically as

$$S_j(k_y + r\Delta k_y, k_x) = \sum_{l=1}^L \sum_{t=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}(l, t, h) \times S_l(k_y + tR\Delta k_y, k_x + h\Delta k_x), \quad j=1, \dots, L, \quad r \neq tR, \quad (2)$$

where $S_j(k_y + r\Delta k_y, k_x)$ denotes the unacquired k -space signal at the target coil, $S_l(k_y + tR\Delta k_y, k_x + h\Delta k_x)$ denotes the acquired undersampled signal, and $w_{j,r}(l, t, h)$ denotes the linear combination coefficients. Here R represents the ORF, l counts all coils, t and h transverse the acquired neighboring k -space data in k_y and k_x directions respectively, and the variables k_x and k_y represent the coordinates along the frequency- and phase-encoding directions, respectively. Figure 5 presents a demonstration of GRAPPA reconstruction as mentioned above, in which the white dots represent the missing k -space signals, the black

dots denote the acquired undersampled data, and the green dots represent auto-calibration signals.

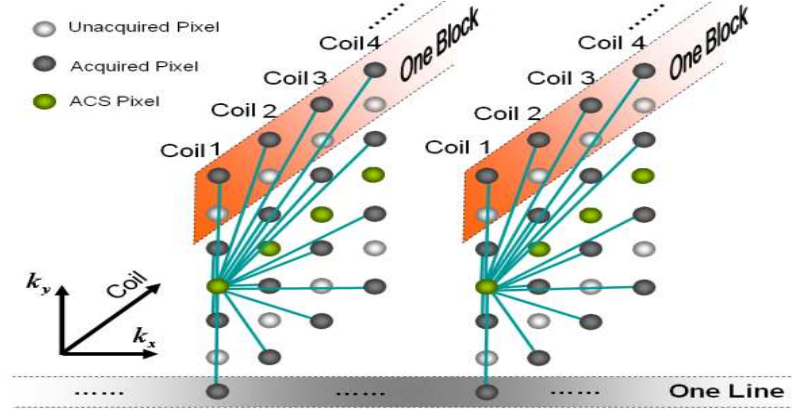


Figure 5. GRAPPA reconstruction demonstration

When the acceleration factor is high, GRAPPA reconstruction can suffer from aliasing artifacts and noise amplifications. Several methods have been developed in recent years to improve GRAPPA, such as localized coil calibration and variable density sampling [12], multicolumn multiline interpolation [13], regularization [14,15], iteratively reweighted least-squares [16], high-pass filtering [17], cross validation [18,19], iterative optimization [20], GRAPPA operator [21,22], virtual coil using conjugate symmetry [23], multi-slice weighting [24], infinite pulse response (IIR) filtering [25], cross sampling [26], and filter bank methods [27,28]. The problems of GRAPPA will be detailed in the next chapter, where the contribution of the dissertation will also be given.

1.3 Compressed Sensing MRI

Compressed sensing [67, 68] is a technique for finding sparse solutions to underdetermined linear systems. In electrical engineering field, particularly in signal processing, compressed sensing is the process of acquiring and reconstructing a signal

that is supposed to be sparse or compressible. Medical images can be compressed with little or no perceptible loss of information. For example, some compression standards including JPEG, JPEG-2000, and MPEG are widely used compression strategies for transform-based compression.

Imaging speed is one of the most challenges in MRI due to an inherently slow data acquisition process. The application of CS to MRI has the potential for significant scan time reductions, with benefits for patients and health care economics. MRI obeys two key requirements for successful application of CS [45, 46]: (1) medical imaging is naturally compressible by sparse coding in an appropriate transform domain (e.g., by wavelet transform); (2) MRI scanners naturally acquire samples of the encoded image in spatial frequency, rather than direct pixel samples. For MRI, CS is also a nonlinear reconstruction process with iterations.

Because parallel MRI and CS reduce sampling based on complementary information, several methods [30, 32, 47, 48] have been developed to combine pMRI and CS for further reduction. Among these methods, the ones that sequentially carry out CS reconstruction for the aliased image of each channel and parallel imaging for the final unfolded image have demonstrated several advantages [32, 47]. Firstly, because the CS and pMRI procedures are clearly decoupled, any CS or pMRI methods are directly applicable without modifications. Secondly, the sampling pattern can be designed to satisfy different requirements for CS and pMRI. Finally, the acceleration factor of the combined approach is a product of the factors achieved by CS and pMRI individually.

1.4 Motivation and Organization

Both SENSE and GRAPPA are both commercially used reconstruction methods. For both techniques, GRAPPA as a k -space-based reconstruction technique has the advantage that it doesn't need to estimate sensitivity maps explicitly and thus avoids error propagation in the whole reconstruction procedure. However, there still exist several issues with GRAPPA.

For example, when acceleration factor is high, reconstruction quality will be deteriorated by noise. Current net acceleration factor of commercially used GRAPPA in SIEMENS MRI scanner is around 2.6, which is far below the number of coils. In addition, more and more channels have been used in parallel MRI for improving SNR. The reconstruction time becomes very significant with a large number of channels. These problems will be analyzed and presented in details in the problem statement of chapter 2. To address these issues, we study nonlinear approaches for GRAPPA reconstruction in this dissertation. Furthermore, the study is also motivated by the success of the previous research in our group which has demonstrated the benefits of nonlinear approaches for SENSE [29-32].

The dissertation is organized as follows. Chapter 1 presents the introduction of the MRI, the parallel MRI techniques and relevant issues for GRAPPA reconstruction. Problem statement and contribution of the dissertation will be given in the chapter 2. After present these basic materials, we will present three novel techniques for improving GRAPPA in chapter 3, 4, and 5, respectively. The work is summarized and future work will be presented in chapter 6. Finally, reference will be given.

Furthermore, within chapter 3, 4, and 5, we will analyze the specific problem firstly, and then improved methods will be proposed via some nonlinear approaches. Finally, experimental results, discussion, or remaining problems to be solved will be presented.

2. Problem Statement and Contribution of the Dissertation

2.1 Problem Statement

2.1.1 Low SNR with High Imaging Acceleration

As pointed out in the chapter 1, SNR is deteriorated when acceleration factor is increasing. GRAPPA reconstruction has the same problem. As a concrete example, noise and aliasing artifacts increase with higher reduction factor as shown in Figure 6. The conventional GRAPPA method [9, 13] reconstructs the missing k -space data by a linear combination of the acquired data, where the coefficients for combination are estimated using some auto-calibration signal (ACS) lines usually acquired in the central k -space. Huang et. al. [49] analyzed two kinds of errors in GRAPPA reconstruction: truncation error and inversion error. Nana et. al. [18, 19] extended the analysis and used more general terms: model error and noise-related error. The first kind of error mainly originates from a limited number of ACS lines and data truncation. When a limited size of k -space signals is observed or inappropriately chosen instead of the whole k -space, model errors occur in GRAPPA reconstruction. This type of error usually varies with the amount of ACS data, reduction factor, and the size of the coefficients to be estimated for reconstruction. For example, a reduction in ACS acquisition usually results in degraded image quality. Therefore a large amount of ACS data is needed to reduce this model error but at the cost of prolonged acquisition time. The second kind of errors originates from noise in the measured data and noise-induced error in estimating the coefficients for linear combination. Regularization [14, 15] has been used in solving the inverse problem

for the coefficients, but significant noise reduction is usually at the cost of increased aliasing artifacts. Iterative reweighted least-squares [16] method reduces the noise-induced error to a greater extent by ignoring noise-induced “outliers” in estimating the coefficients. However, the method is computational expensive.

In general, the coefficients depend on the coil sensitivities and are not known *a priori*. The ACS data are used to estimate these coefficients. Among all the ACS data fully acquired at the central k -space, each location is assumed to be the “missing” point to be used on the left-hand side of Eq. (2). The neighboring locations with a certain undersampling pattern along the phase encoding direction are assumed to be the undersampled points that are used on the right-hand side of Eq. (2). This is repeated for all ACS locations (except boundaries of the ACS region) based on the shift-invariant property to fit GRAPPA coefficients to all ACS data. This calibration process can be simplified as a matrix equation

$$\mathbf{b} = \mathbf{A} \mathbf{x} \quad (3)$$

where \mathbf{A} represents the matrix comprised of the undersampled points of the ACS, \mathbf{b} denotes the vector for the “missing” points of the ACS, and \mathbf{x} represents the coefficients to be fitted. The matrix \mathbf{A} is of size $M \times K$ with M being the total number of ACS data (excluding the boundaries) and K being the number of points in the neighborhood from all coils that are used in reconstruction. The least-squares method is commonly used to calculate the coefficients:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 \quad (4)$$

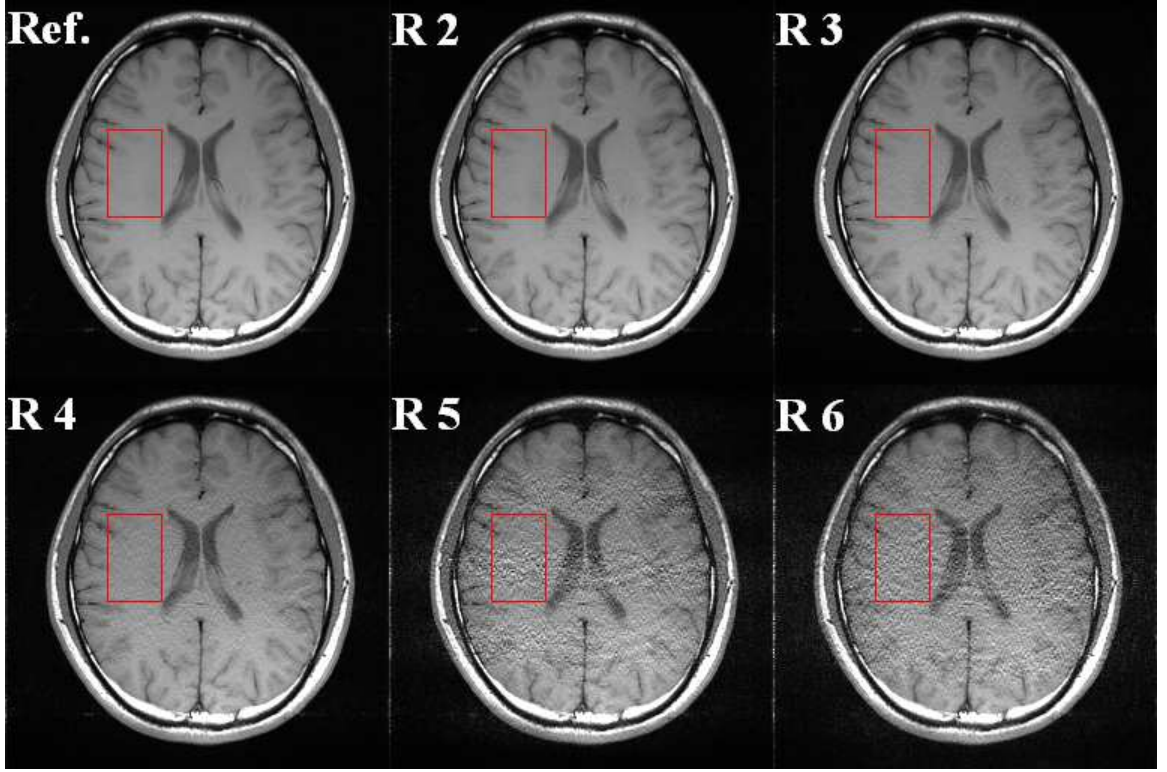


Figure 6. Lower SNR as increasing outer reduction factor. The reconstruction parameters of 8-channel brain data are 40 ACS lines, coefficient size of convolution is 4x5.

When the matrix \mathbf{A} is ill-conditioned, the noise can be greatly amplified in the estimated coefficients. To address the ill-conditioning issue, regularization methods [14, 15] have been used to solve for coefficients using a penalized least-squares method,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{Ax}\|^2 + \lambda R(\mathbf{x}) \quad (5)$$

where $R(\mathbf{x})$ is a regularization function (e.g., $R(\mathbf{x}) = \|\mathbf{x}\|_2$ in Tikhonov) and λ is regularization parameter. Regularization can effectively suppress noise to a certain level. However, aliasing artifacts usually appear in reconstruction at the same time while large noise is suppressed.

Another source of noise-induced error in GRAPPA is “outliers”. Outliers are k -space data with large measurement errors due to noise and low sensitivity. They lead to large deviations in the estimated coefficients from the true ones when the least-squares fitting is used. Iterative reweighted least-squares method [16] has been proposed to minimize the effect of outliers in least-squares fitting. The method iteratively assigns and adjusts weights for the acquired undersampled data. “Outliers” are given less weights or removed in the final estimation so that the fitting accuracy and reconstruction quality is improved. However, the high computational complexity of the method limits its usefulness in practice.

2.1.2 Imaging Acceleration Is Not High Enough

Besides SNR problem, higher acceleration is also another problem of parallel MRI reconstruction methods. As stated previously, following the increasing acceleration, noise deteriorates reconstruction quality in parallel MRI. More efficient reconstruction that needs both high acceleration and high SNR should be achieved simultaneously. GRAPPA reconstruction has been commercially used on SIEMENS scanner, on which the acceleration factor is usually set as 3 and the net acceleration factor provided by general parameter settings is around 2.6 for acceptable reconstruction quality, which is not high enough for even faster imaging.

In the dissertation, although reconstruction quality provided by nonlinear GRAPPA has been improved to some extent, its net accelerations factor is still not high enough. Compressed sensing has been combined with parallel MRI for higher imaging acceleration [30, 32, 47, 48]. For the existing methods that serially combine CS and

parallel MRI, both reconstructed signals by CS and pMRI are not as accurate as acquired MR signals, so that errors of CS reconstruction as the first step will be propagated and even amplified to deteriorate pMRI reconstruction quality in the second step. Therefore, robust method is required to achieve high SNR and high acceleration simultaneously.

2.1.3 Computation Cost of Reconstruction Is Large with Many Coils

Due to success of parallel MRI methods, more and more coils have been used in commercial MRI systems for good reconstruction quality and high acceleration such as 128 coil arrays on research MRI scanners [50]. However, computational cost provided by a large number of element coils increases significantly. Some channel reduction methods [43, 44, 51-54] have been proposed to reduce reconstruction time and keep the almost the same reconstruction quality by using a subset of channels (virtual channels). For example, coil array compression using PCA for MRI with large coil arrays was proposed for SENSE reconstruction [51, 52]. The number of channels can be largely reduced instead of a little SNR loss, so that reconstruction time is accelerated with sacrificing only a very little degeneracy of quality. Furthermore, channel reduction technique by PCA was also successfully applied on GRAPPA as a k -space-based reconstruction approach [43, 44]. Due to reduced channels of having very low sensitivities make little contribution to the SNR of the final reconstructed image [43], two-stage channel compression method has been used for reducing computational cost of GRAPPA reconstruction [44]. However, reconstruction quality is still a little degenerative, although computational and memory costs have been largely reduced. A trade-off exists between loss of information that degenerates reconstruction quality and reduced computational cost. Furthermore, a small

set of neighbor channels were selected based on correlation among channels for reducing computation cost [53]. The reconstruction quality is almost the same in comparison with reconstruction using full channels. A hardware RF signal combiner inline after preamplification was placed to construct an eigencoil array [54]. Optimal SNR can be achieved with receiver channel reduction.

For software-based methods [43, 44, 51-53], due to the loss information caused by channel reduction, reconstruction quality using a reduced subset of virtual coils is equivalent to or a little worse than the reconstruction quality by using full channels. In other words, current software-based channel compression techniques for parallel MRI only focus on accelerating reconstruction time and don't improve reconstruction quality. Therefore, parallel MRI reconstruction with many channels (such as 32-channels, 64-channels) for simultaneously reducing computational cost and achieving high reconstruction quality is another challenge topic that would be solved in the dissertation.

2.2 Dissertation Contribution

Considering the three problems mentioned above, the dissertation is addressing these problems on enhancing SNR at high reduction factors, increasing imaging acceleration, and reducing channels of parallel MRI, respectively.

For SNR problem, we focus on the nature of noise-induced error and develop a novel nonlinear method to reduce such kind of error. Errors-in-variables problem that exists in the conventional GRAPPA reconstruction was analyzed, and nonlinear noise was generalized and suppressed by using kernel method. The proposed nonlinear GRAPPA (NLGRAPPA) not only has the advantage of nonlinear methods in representing

generalized models that include linear ones as a special case, but also maintains the simplicity of linear methods in computation.

In order to achieve even higher imaging acceleration, we proposed a novel serial combination method of CS and NLGRAPPA (CS-NLGRAPPA) reconstruction based on the first work – nonlinear GRAPPA. The generalized kernel regression model of NLGRAPPA can remove error effects of inaccurate signal reconstruction by CS. The experimental results using phantom and in vivo data demonstrate the proposed CS-NLGRAPPA method can significantly improve the reconstruction quality over the existing method and push net reduction factor around 4. The general guideline of selecting sampling patterns is presented based on the discussion of the experimental results.

For computational costs of parallel MRI with many channels, we aimed at combining reduced channels and improving reconstruction quality simultaneously via software-based method. The proposed channel reduction method – kernel PCA (KPCA) based reduction not only reduces computational cost for the conventional GRAPPA reconstruction, but also enhances reconstruction quality.

3. Nonlinear GRAPPA - A Kernel Approach to Parallel MRI Reconstruction

GRAPPA linearly combines the undersampled k -space signals to estimate the missing k -space signals where the coefficients are obtained by fitting to some auto-calibration signals (ACS) sampled with Nyquist rate based on the shift-invariant property. At high acceleration factors, GRAPPA reconstruction can suffer from a high level of noise even with a large number of auto-calibration signals. In this work, we propose a nonlinear method to improve GRAPPA. The method is based on the so-called kernel method which is widely used in machine learning. Specifically, the undersampled k -space signals are mapped through a nonlinear transform to a high-dimensional feature space, and then linearly combined to reconstruct the missing k -space data. The linear combination coefficients are also obtained through fitting to the ACS data but in the new feature space. The procedure is equivalent to adding many virtual channels in reconstruction. A polynomial kernel with explicit mapping functions is investigated in this work. Experimental results using phantom and in vivo data demonstrate that the proposed nonlinear GRAPPA method can significantly improve the reconstruction quality over GRAPPA and its state-of-the-art derivatives.

Kernel methods [33-37] have been successfully applied to a number of real-world problems and are now considered state-of-the-art in various domains such as classification, pattern analysis, machine learning, regression, and so forth. Recently, it

has been successfully applied on signal processing [34] for robust regression and least-squares solution. Due to versatile tools of solving problems kernel methods offer, we study it for GRAPPA reconstruction which is essentially a linear filtering or regression problem.

In this research, we focus on the nature of noise-induced error and develop a novel nonlinear method to reduce noise-induced error. We identify the nonlinear relationship between the bias in the estimated GRAPPA coefficients and the noise in the measured ACS data due to the errors-in-variables problem in the calibration step. This relationship suggests that the finite impulse response (FIR) model currently used in GRAPPA reconstruction is not able to remove the nonlinear noise-induced bias even if regularization is used. We thereby propose a nonlinear approach to GRAPPA using the kernel method, named nonlinear GRAPPA. (Note this kernel is a terminology in machine learning and is different from the GRAPPA kernel for linear combination.) The method maps the undersampled data onto a high dimensional feature space through a nonlinear transform and the data in the new space are then linearly combined to estimate the missing k -space data. Although the relationship between the acquired and missing k -space data is nonlinear, the relationship can be easily and linearly found in the high dimensional feature space using the ACS data. It is worth noting that the nonlinearity of this approach is completely different from that in the GRAPPA operator formulation in Refs. [21, 22] where the former is on the k -space data while the latter is on the GRAPPA coefficients through successive application of linear operators. The proposed method not only has the advantage of nonlinear methods in representing generalized models that

include linear ones as a special case, but also maintains the simplicity of linear methods in computation.

3.1 Errors-in-variables Model of GRAPPA

The conventional GRAPPA formulation in Eq. (2) or (3) models the calibration and reconstruction as a standard linear regression and prediction problem, where the undersampled part of the ACS corresponds to the regressors and the rest is the regressands. With this formulation, if the undersampled points of the ACS (regressors) are measured exactly or observed without noise, and noise is present only in the “missing” ones of the ACS (regressands), then the least-squares solution is optimal and the error in the reconstruction is proportional to the input noise. However this is not the case in GRAPPA because all ACS data are obtained from measurement and thus contain the same level of noise.

To understand the effect of noise in both parts of the ACS data (regressors and regressands), we describe the regression and prediction process of GRAPPA using latent variables [55]. Specifically, if \mathbf{A} and \mathbf{b} are observed variables that come from the ACS data with measurement noise, we assume that there exist some unobserved latent variables value $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{b}}$ representing the true, noise-free counterparts, whose true functional relationship is modeled as a linear function f . We thereby have

$$\begin{cases} \mathbf{A} = \tilde{\mathbf{A}} + \delta_{\mathbf{A}} \\ \mathbf{b} = \tilde{\mathbf{b}} + \delta_{\mathbf{b}} \\ f : \tilde{\mathbf{b}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} \end{cases} \quad (6)$$

where $\delta_{\mathbf{A}}$ and $\delta_{\mathbf{b}}$ represent measurement noise that are present in the ACS data and assumed to be independent of the true value $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{b}}$, and $\tilde{\mathbf{x}}$ denotes the latent true coefficients for the linear relationship between $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{b}}$ without the hidden noise.

In the standard regression process, the coefficients \mathbf{x} is estimated by fitting to the observed data in \mathbf{A} and \mathbf{b} :

$$\mathbf{b} = \mathbf{A}\mathbf{x} \rightarrow \tilde{\mathbf{b}} + \delta_{\mathbf{b}} = (\tilde{\mathbf{A}} + \delta_{\mathbf{A}})\mathbf{x} \quad (7)$$

Therefore, there is a bias $\delta_{\mathbf{x}} = \mathbf{x} - \tilde{\mathbf{x}}$ in the coefficients estimated from the least-squares fitting, where

$$\mathbf{x} = [(\tilde{\mathbf{A}} + \delta_{\mathbf{A}})^T (\tilde{\mathbf{A}} + \delta_{\mathbf{A}})]^{-1} (\tilde{\mathbf{A}} + \delta_{\mathbf{A}})^T (\tilde{\mathbf{b}} + \delta_{\mathbf{b}}) \quad (8)$$

For example, consider the simplest case where \mathbf{x} is a scalar and \mathbf{b} and \mathbf{A} are both column vectors whose elements b_t and a_t represent measurements at index t . The estimated coefficient is given by

$$\mathbf{x} = \sum_{t=1}^T a_t b_t / \sum_{t=1}^T a_t^2 \quad (9)$$

which deviates from the true coefficient $\tilde{x} = \tilde{b} / \tilde{a}$. When the number of measurements T increases without bound, the estimated coefficient converges to

$$x = \tilde{x} / (1 + \sigma_b^2 / \sigma_A^2) \quad (10)$$

where the noise in \mathbf{A} and \mathbf{b} is assumed to have zero mean and variance of σ_A^2 and σ_b^2 respectively. It suggests that even if there are an infinite number of measurements, there is still a bias in the least-squares estimator. Since the bias depends on the noise in both \mathbf{A} and \mathbf{b} , its effects on the estimated coefficients \mathbf{x} are also noise-like. In the multivariable

case, the bias of GRAPPA coefficients is not easily characterized analytically, but is known to be upper bounded by (Theorem 2.3.8 in Ref [56])

$$\frac{\|\delta_x\|_2}{\|\mathbf{x}\|_2} \leq \kappa(\tilde{\mathbf{A}}) \left(\frac{\|\delta_A\|}{\|\tilde{\mathbf{A}}\|} + \frac{\|\delta_b\|}{\|\mathbf{b}\|} + \frac{\|\delta_A\| \|\delta_b\|}{\|\tilde{\mathbf{A}}\| \|\mathbf{b}\|} \right) \quad (11)$$

where $\kappa(\tilde{\mathbf{A}})$ is the condition number of matrix $\tilde{\mathbf{A}}$. The bound in Eq. (11) suggests that the bias in GRAPPA coefficients can be large at high reduction factors due to ill-conditioned \mathbf{A} [14]. In addition, the bias is not a linear function of the noise in the ACS. This is known as the errors-in-variable problem in regression. Figure 7 uses an example to demonstrate the nonlinearity of the bias for GRAPPA coefficients as a function of noise in the ACS data. Specifically, a set of brain data with simulated coil sensitivities (obtained from <http://www.nmr.mgh.harvard.edu/~fhlin/>) was used as the noise-free signal. We calculated the bias for GRAPPA coefficients (with the coefficients obtained from the noise-free signal as reference) when different levels of noise were added on all 24 lines of the ACS data. We plotted the normalized bias for GRAPPA coefficients as a function of the normalized noise level added to the ACS data. It is seen that the bias is not a linear function of noise level. However, when the noise is sufficiently low, the curve is well approximated by a straight line and the bias-noise relationship is approximately linear. Total least squares [57, 58] is a linear method used to alleviate the problem by solving Eq. (7) using the total least squares instead of least squares. It addressed the error-in-variable problem to some extent when the noise is low. In the reconstruction step of GRAPPA, when the biased coefficients \mathbf{x} are applied upon the noisy undersampled data \mathbf{A} to estimate the missing data in outer k -space, errors presented in the reconstruction are given by

$$\mathbf{y} - \tilde{\mathbf{A}}\tilde{\mathbf{x}} = (\tilde{\mathbf{A}} + \delta_{\mathbf{A}})(\tilde{\mathbf{x}} + \delta_{\mathbf{x}}) - \tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{A}}\delta_{\mathbf{x}} + \tilde{\mathbf{x}}\delta_{\mathbf{A}} + \delta_{\mathbf{A}}\delta_{\mathbf{x}} \quad (12)$$

It shows the effect of biased coefficients on the estimated missing k -space data is also nonlinear and noise-like. A comprehensive statistical analysis of noise in GRAPPA reconstruction can be found in [59].

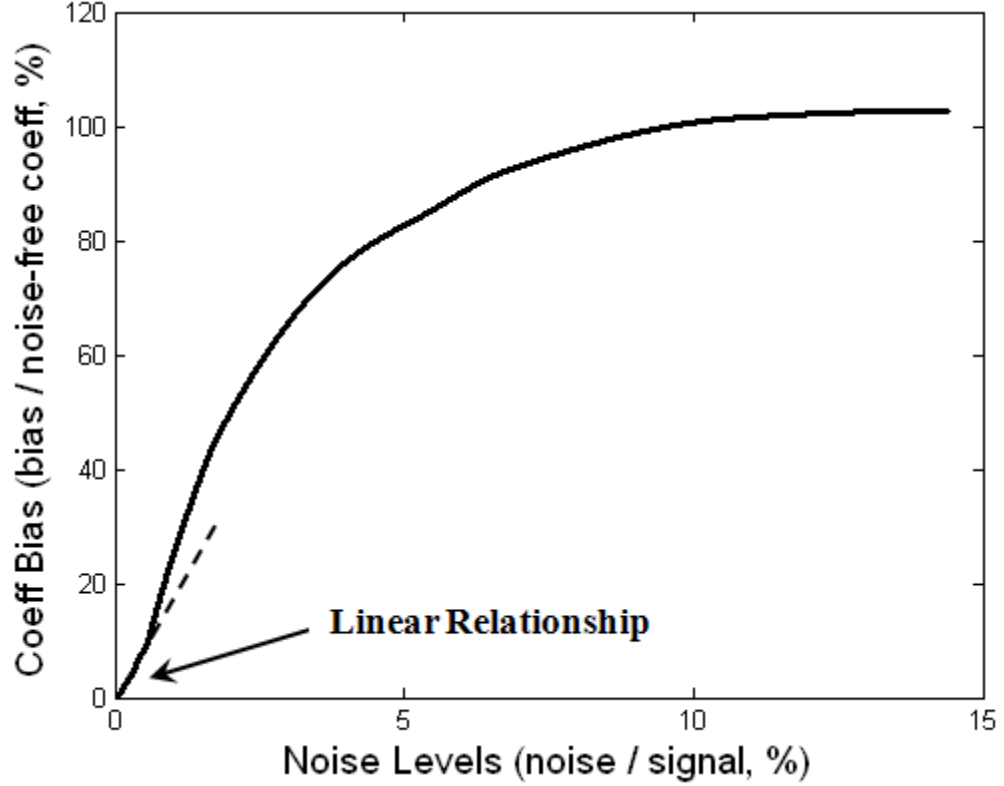


Figure 7. Nonlinearity of the bias for GRAPPA coefficients as a function of noise in ACS data.

3.2 Proposed Nonlinear GRAPPA

All existing GRAPPA derivatives are based on the linear model in Eq. (2) without considering the nonlinear bias due to noise in the ACS data. To address the nonlinear, noise-like errors in GRAPPA reconstruction, a kernel method is proposed to describe the

nonlinear relationship between the acquired undersampled data and the missing data in present of noise-induced errors. Please note that the kernel used in this paper is different from the kernel usually used in GRAPPA literature to represent the k -space neighborhood for linear combination.

3.2.1 General Formulation Using Kernel Method

Kernel method [33-36] is an approach that is widely used in machine learning. It allows nonlinear algorithms through simple modifications from linear ones. The idea of kernel method is to transform the data nonlinearly to a higher dimensional space such that linear operations in the new space can represent a class of nonlinear operations in the original space. Specifically, given a linear algorithm, we map the data in the input space A to the feature space H via a nonlinear mapping $\Phi(\cdot): A \rightarrow H$, and then run the algorithm on the vector representation $\Phi(\mathbf{a})$ of the data. However, the map may be of very high or even infinite dimensions and may also be hard to find. In this case, the kernel becomes useful to perform the algorithms without explicitly computing $\Phi(\cdot)$. More precisely, a kernel is related to the mapping Φ in that

$$k(\mathbf{a}_1, \mathbf{a}_2) = \langle \Phi(\mathbf{a}_1), \Phi(\mathbf{a}_2) \rangle, \quad \forall \mathbf{a}_1, \mathbf{a}_2 \in A \quad (13)$$

where \langle, \rangle represents the inner product. Many different types of kernels are known [36] and the most general used ones include polynomial kernel [60] and Gaussian kernel [61].

To introduce nonlinearity into GRAPPA, we apply a nonlinear mapping to the undersampled k -space data $\mathbf{a}_i = \{ S_l(k_y + tR\Delta k_y, k_x + h\Delta k_x) \}$ in the neighborhood of each missing point where l counts all coils and t and h transverse the acquired neighboring k -

space data in k_y and k_x directions respectively. Under such a mapping, Eq. (3) is transformed to the following new linear system of \mathbf{x} :

$$\mathbf{b} = \Phi(\mathbf{A})\mathbf{x} \quad (14)$$

where $\Phi(\mathbf{A}) = [\Phi(\mathbf{a}_1), \Phi(\mathbf{a}_2), \dots, \Phi(\mathbf{a}_M)]^T$ with \mathbf{a}_i being the i th row vector of the matrix \mathbf{A} defined in Eq. (2). The new matrix $\Phi(\mathbf{A})$ is of $M \times N_K$, where N_K is the dimension in the new feature space which is usually much higher than K . Equation (14) means the missing data in \mathbf{b} is a linear combination of the new data in feature space which are generated from the original undersampled k -space data \mathbf{A} . Although Eq. (14) is still a linear equation of the coefficients \mathbf{x} , it mathematically describes the nonlinear relationship between the undersampled and missing data because of the nonlinear mapping function $\Phi(\cdot)$. With the ACS data, the regression process to find the coefficients \mathbf{x} in Eq. (14) for the proposed nonlinear GRAPPA can still be solved by a linear, least-squares algorithm in feature space

$$\hat{\mathbf{x}} = (\Phi^H(\mathbf{A})\Phi(\mathbf{A}))^{-1} \Phi^H(\mathbf{A})\mathbf{b} \quad (15)$$

Once the coefficients are estimated in Eq. (15), they are plugged back in Eq. (14) for the prediction process to reconstruct the missing data in outer k -space, like the conventional GRAPPA does. Figure 8 summarizes the above procedure and illustrates the nonlinear and linear parts of the proposed method (NLGRAPPA) in comparison to the conventional GRAPPA. It can be seen that the proposed method introduces an additional nonlinear mapping step into GRAPPA to pre-process the acquired undersampled data while the computational algorithm to find the coefficients is still the linear least-squares method.

Other linear computational algorithms such as reweighted least-squares and total least-squares methods can also be used here.

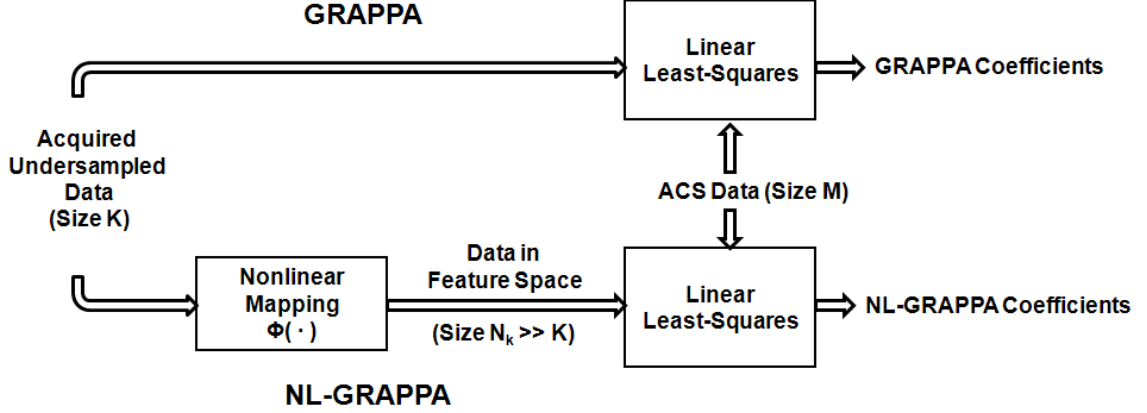


Figure 8. Illustration of the calibration procedure for GRAPPA and nonlinear (NL) GRAPPA.

3.2.2 Choice of nonlinear mapping $\Phi(\cdot)$

To choose the optimal kernel or feature space is not trivial. For example, Gaussian kernel has been proved to be universal, which means that linear combinations of the kernel can approximate any continuous function. However, overfitting of the calibration data may arise as a result of this powerful representation. Given the success of GRAPPA, we want the nonlinear mapping to be a smooth function that includes the linear one as a special case when the dimension of the feature space is as low as the original space. Since polynomials satisfy the desired properties, we choose an inhomogeneous polynomial kernel of the following form

$$\kappa(\mathbf{a}_i, \mathbf{a}_j) = (\gamma \mathbf{a}_i^T \mathbf{a}_j + r)^d \quad (16)$$

where γ and r are scalars and d is the degree of the polynomial. Another advantage of polynomial kernel lies in the fact that its corresponding nonlinear mapping $\Phi(\mathbf{a})$ such that $k(\mathbf{a}_1, \mathbf{a}_2) = \langle \Phi(\mathbf{a}_1), \Phi(\mathbf{a}_2) \rangle$ has explicit representations. For example, if $\gamma = r = 1$ and $d = 2$, $\Phi(\mathbf{a})$ is given by [62]

$$\Phi(\mathbf{a}) = [1, \sqrt{2}a_1, \dots, \sqrt{2}a_K, a_1^2, \dots, a_K^2, \sqrt{2}a_1a_2, \dots, \sqrt{2}a_1a_j, \dots, \sqrt{2}a_{K-1}a_K]^T \quad (17)$$

where a_1, a_2, \dots, a_K are components of the vector \mathbf{a} and there are $(K+2)(K+1)/2$ terms in total. It is seen that the vector includes the linear terms in the original space as well as the constant and second-order terms.

When all possible terms in $\Phi(\mathbf{a})$ are included, direct use of the kernel function may be preferred over the use of nonlinear mapping in Eq. (14) for the sake of computational complexity. However our experiment (See Figure 14 in Results) shows that the reconstruction using kernel functions suffers from blurring and aliasing artifacts. This is because the model is excessively complex and represents a too broad class of functions, and thus the model has been overfit during calibration but poorly represents the missing data. This overfitting problem can be addressed by reducing the dimension of the feature space [63]. The reduction of feature space is achieved here by keeping the constant term and all first-order terms $\sqrt{2}a_1, \dots, \sqrt{2}a_K$, but truncating the second-order terms in vector $\Phi(\mathbf{a})$. Specifically, we sort the second-order terms according to the following order. We first have the square terms within each coil, and then the product terms between the nearest neighbors, the next-nearest neighbors, and so on so forth in k -space. The above order is then repeated for terms that are across different coils. With the sorted terms, we can truncate the vector $\Phi(\mathbf{a})$ according to the desired dimension of the feature space.

The performance of the proposed method depends on the number of second-order terms. If the number is too low, prediction is inaccurate because the feature space is not complex enough to accurately describe the true relationship between the calibration and undersampled data in presence of noise, and thus the reconstruction resembles GRAPPA and still suffers from noise-like errors. On the other hand, if the dimension is too high, the model is overfit by the calibration data but poorly represents the missing data, thus leading to aliasing artifacts and loss of resolution in reconstruction. This is known as the bias-variance tradeoff and is demonstrated using an example in Results section.

3.2.3 Explicit implementation of nonlinear GRAPPA

We find heuristically (elaborated in Results) that it is sufficient to keep the number of the second-order terms to be about three times that of the first-order terms. That is, the feature space is reduced to

$$\tilde{\Phi}(\mathbf{a}) = [1, \sqrt{2}a_1, \sqrt{2}a_2, \dots, \sqrt{2}a_K, \underbrace{a_1^2, a_2^2, \dots, a_K^2}_K, \underbrace{a_1a_2, \dots, a_ia_j, \dots, a_{K-1}a_K}_{\sim K}, \underbrace{a_1a_3, \dots, a_pa_q, \dots, a_{K-2}a_K}_{\sim K}] \quad (18)$$

where (a_i, a_j) are nearest neighbors and (a_p, a_q) are next-nearest neighbors in k -space along k_x within each coil. We also find that a slight increase in the number of second-order terms does not change the reconstruction quality, but increases the computation.

After plugging the above truncated mapping vector $\tilde{\Phi}(\mathbf{a})$ in Eq. (18) into the matrix representation in Eq. (14) and changing to the notations in conventional GRAPPA, the proposed nonlinear GRAPPA method is thereby formulated as

$$\begin{aligned}
S_j(k_y + r\Delta k_y, k_x) = & w_{j,r}^{(0)} \times 1 + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}^{(1)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}^{(2,0)}(l, b, h) \times S_l^2(k_y + bR\Delta k_y, k_x + h\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2-1} w_{j,r}^{(2,1)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \times S_l(k_y + bR\Delta k_y, k_x + (h+1)\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2-2} w_{j,r}^{(2,2)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \times S_l(k_y + bR\Delta k_y, k_x + (h+2)\Delta k_x),
\end{aligned} \tag{19}$$

where the same notations are used as in Eq. (2).

The above nonlinear formulation represents a more general model for GRAPPA, which includes the conventional GRAPPA as a special case. It is seen that the first-order term of nonlinear GRAPPA in Eq. (19) is equivalent to the conventional GRAPPA, which mainly captures the linear relationship between the missing and acquired undersampled data in the absence of noise and approximations. The second-order terms of Eq. (19) can be used to characterize other nonlinear effects in practice such that noise and approximation errors are suppressed in reconstruction. The proposed formulation is a nonlinear model in the sense that nonlinear combination of acquired data contributes to estimation of missing k -space data. However the computational algorithm is still linear because the new system equation in Eq. (19) is still a linear function of the unknown coefficients and can still be solved by the linear least-squares method.

To better interpret the nonlinear GRAPPA method, we can consider the nonlinear terms as additional virtual channels as done in Ref. [23]. For example, the first-order terms in Eq. (19) represent a linear combination of L physical channels, while each second-order term represents a set of additional L virtual channels. Therefore there are $4L$ channels in total when Eq. (19) is used. More second-order terms provide more virtual

channels. It is worth noting that different from the true physical channel, there is no equivalent concept of coil sensitivities for the virtual channels. This is because the additional virtual channels are nonlinear function of the original physical channels. For example, the “square channel” takes the square of the k -space data point-by-point. In image domain, this is equivalent to the sensitivity-modulated image convolves with itself. Therefore the resulting image cannot be represented as the product of the original image and another independent “sensitivity” function. Another point to be noted is that the virtual channels are not necessarily all independent. Only adding channels that are linearly independent can improve the reconstruction performance. Choosing independent channels needs further study in our future work.

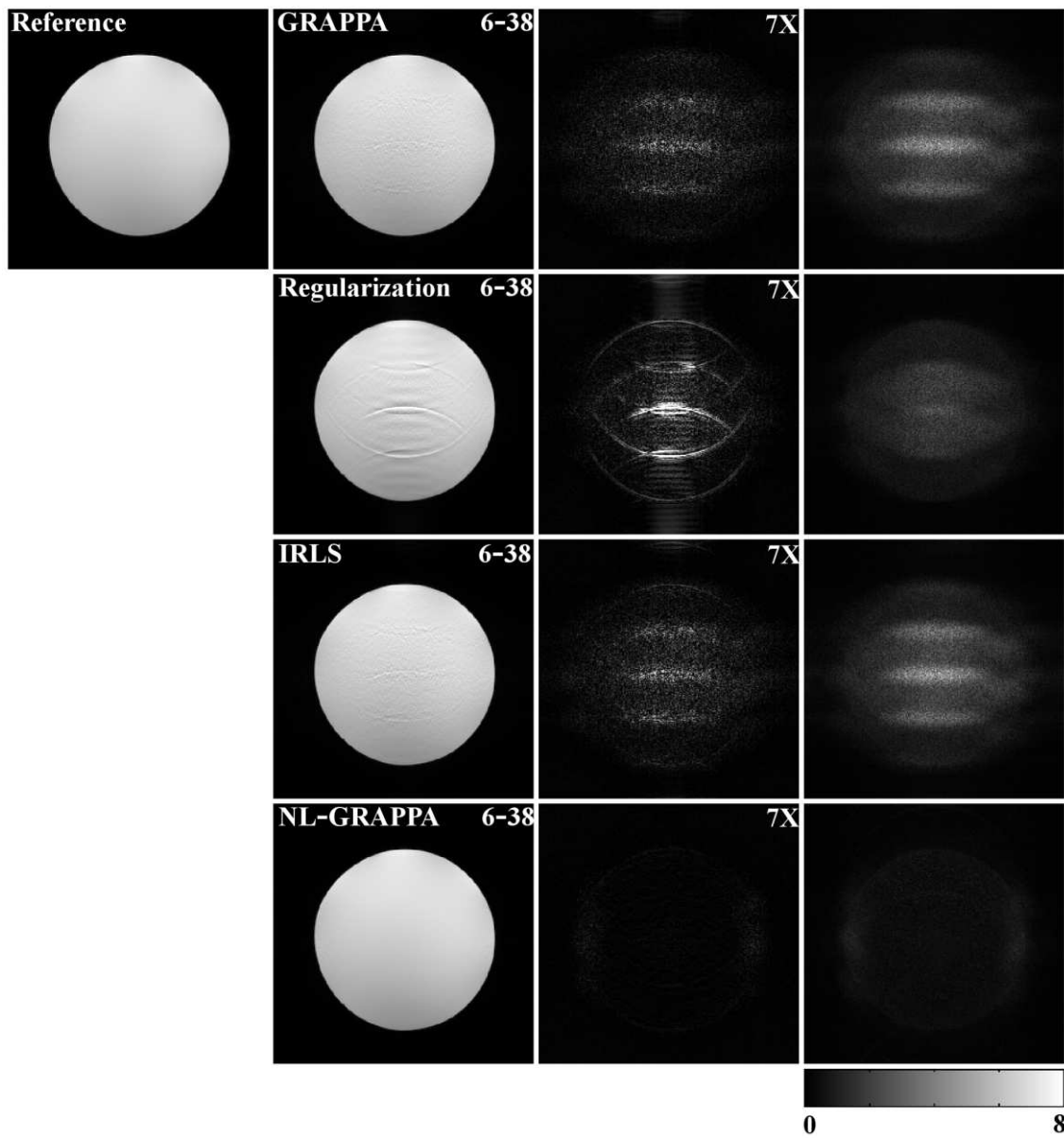


Figure 9. Phantom images reconstructed from an eight-channel dataset with an ORF of 6 and 38 ACS lines (denoted as 6-38 on the right corner of each image). With the SOS reconstruction as the reference, the proposed nonlinear GRAPPA (NL-GRAPPA) method is compared with conventional GRAPPA, regularized GRAPPA, and IRLS methods. The corresponding difference images with the reference ($7\times$ amplification) and g-factor maps are also shown on the right two columns respectively.

3.3 Experimental Results

3.3.1 Experiment Settings

The performance of the proposed method was validated using four scanned datasets. The first three scanned datasets were all acquired on a GE 3T scanner (GE Healthcare, Waukesha, WI) with an 8-channel head coil, and the last one was acquired on a Siemens 3T scanner (Siemens Trio, Erlangen, Germany). In the first dataset, a uniform water phantom was scanned using a gradient echo (GRE) sequence (TE/TR = 10/100 ms, 31.25 kHz bandwidth, matrix size = 256×256 , FOV = 250 mm²). The second dataset was an axial brain image acquired using a 2D spin echo (SE) sequence (TE/TR = 11/700 ms, matrix size = 256×256 , FOV = 220 mm²). The third one was a sagittal brain dataset acquired using a 2D spin echo (SE) sequence (TR = 500 ms, TE = min full, matrix size = 256×256 , FOV = 240 mm²). In the fourth dataset, cardiac images were acquired using a 2D trueFISP sequence (TE/TR=1.87/29.9 ms, bandwidth 930 Hz/pixel, 50 degree flip angle, 6mm slice thickness, 34 cm FOV in readout direction, 256x216 acquisition matrix) with a 4-channel cardiac coil. Informed consents were obtained for all *in vivo* experiments in accordance with the institutional review board policy.

The proposed method was compared with conventional GRAPPA, as well as two existing methods that improve the SNR, Tikhonov regularization [14] and iterative reweighted least-squares (IRLS) [16]. The root sum of squares (SOS) reconstruction from the fully sampled data of all channels was shown as the reference image for comparison. The size of the coefficients (blocks by columns) was chosen optimally for each individual method by comparing the mean-squared errors resulting from different sizes. The g-factor

map was calculated using Monte Carlo simulations as described in [64] and used to show noise amplification. It is worth noting that for nonlinear algorithms, the SNR loss depends on the input noise level, and the g-factors shown in Results are valid only in a small range around the noise level used in this study. Difference images were used to show all sources of error, including blurring, aliasing and noise. All methods were implemented in MATLAB (Mathworks, Natick, MA). To facilitate visual comparison, difference images from the reference and zoomed-in patches were also shown for some reconstructions. A software implementation of the proposed nonlinear GRAPPA method is available at https://pantherfile.uwm.edu/leiyang/www/index_files/software.htm.

3.3.2 Results of Phantom and In Vivo Datasets

Figure 9 shows the reconstructions of the phantom using SOS, GRAPPA, Tikhonov regularization, IRLS, and the proposed nonlinear GRAPPA for an ORF of 6 and the ACS of 38 (net acceleration of 3.41). The size of the coefficients was chosen optimally for each individual method, though the image quality is not sensitive to the change of size within a large range of the optimal choice. The size of the coefficients for nonlinear GRAPPA was 2 blocks and 15 columns and that for the other methods was 4 blocks and 9 columns. It is seen that the conventional GRAPPA suffers from serious noise. Tikhonov regularization and IRLS can both improve the SNR to some extent but at the cost of aliasing artifacts. The proposed nonlinear GRAPPA method suppresses most of the noise without additional artifacts or loss of resolution. In addition, difference images with the reference and g-factor maps shown in Fig. 9 also suggest that the noise-like

errors have quite different distributions spatially and they are more uniformly distributed in nonlinear GRAPPA than in other methods.

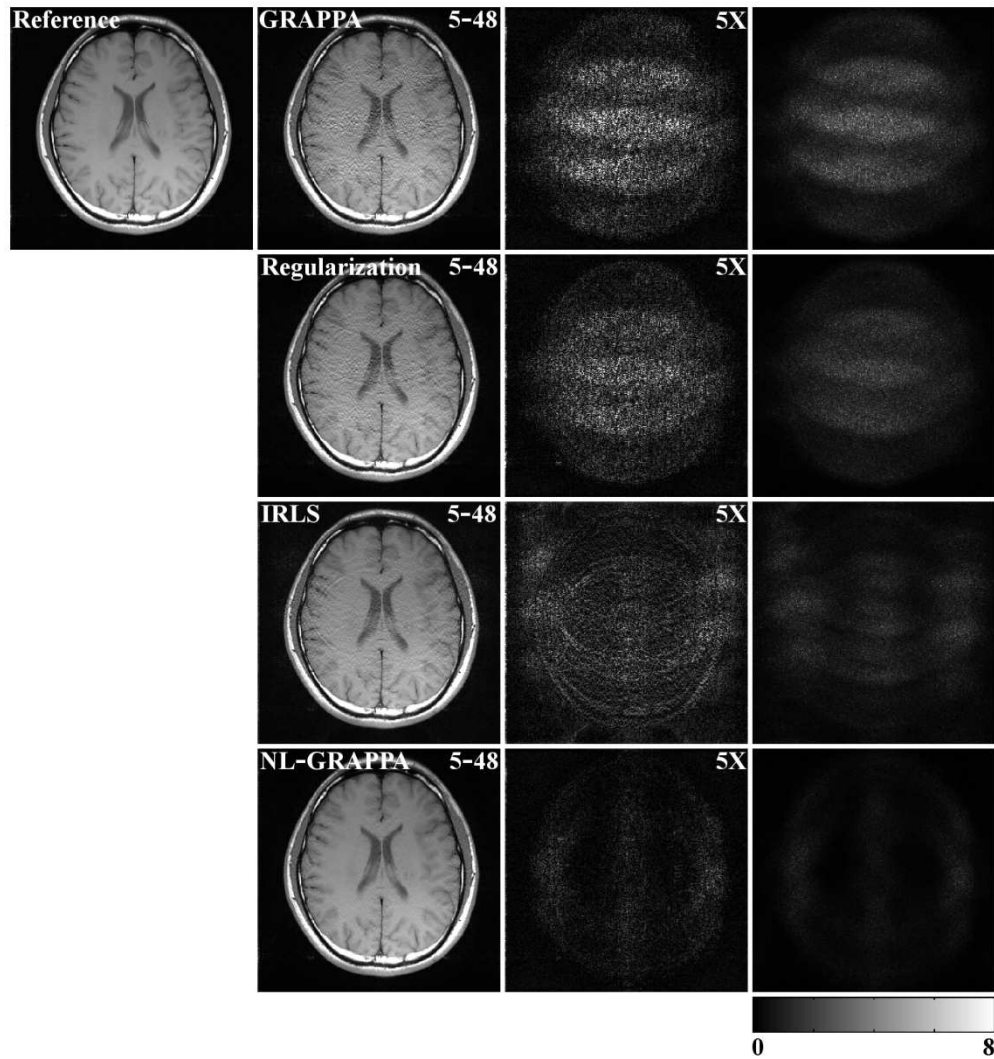


Figure 10. Axial brain images reconstructed from a set of eight-channel data with an ORF of 5 and 48 ACS lines using GRAPPA, regularized GRAPPA, IRLS, and the proposed nonlinear method. The corresponding difference images with the reference (5 \times amplification) are shown on the middle column and g-factor maps on the right column.

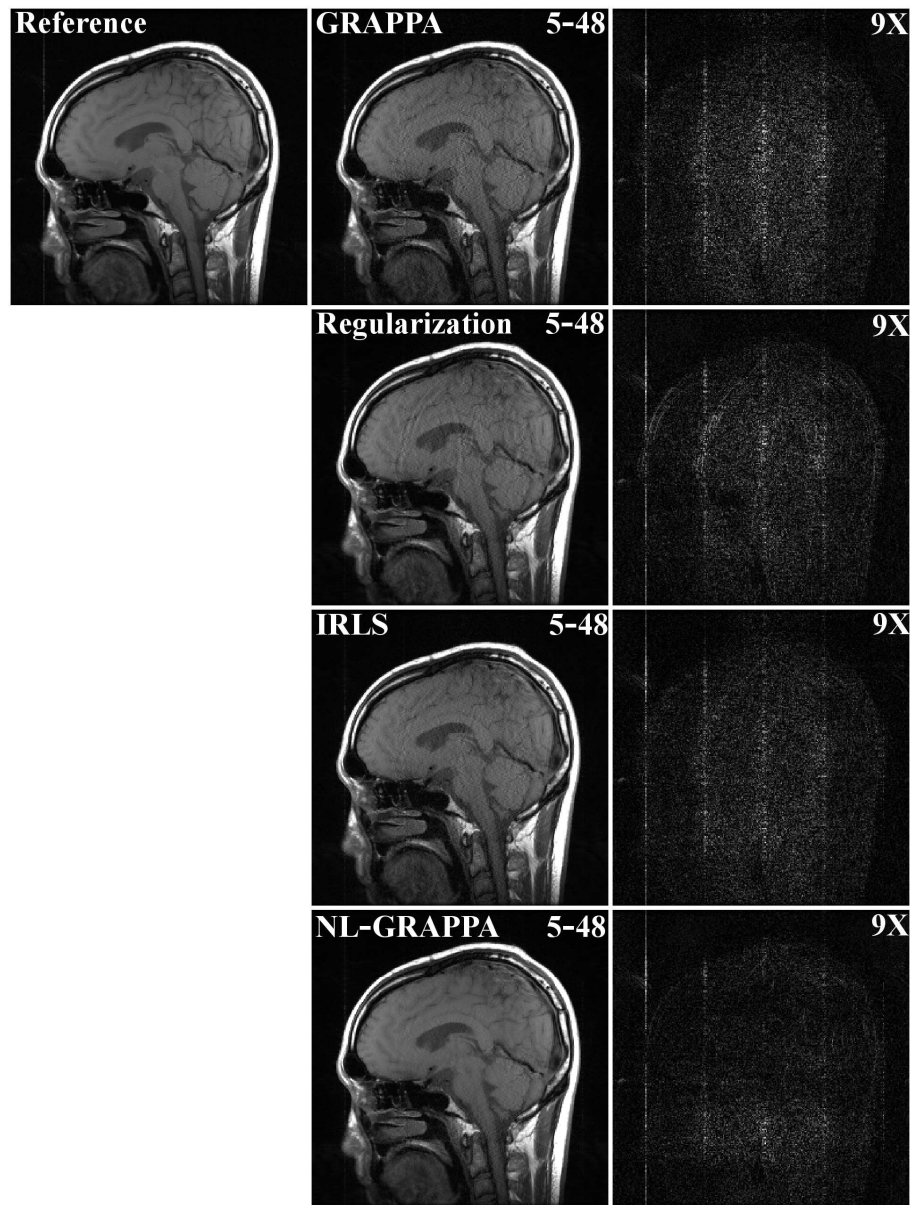


Figure 11. Sagittal brain images reconstructed from a set of eight-channel data with an ORF 5 and 48 ACS lines and their corresponding difference images on the right. The proposed nonlinear GRAPPA suppresses most noise without aliasing artifacts.

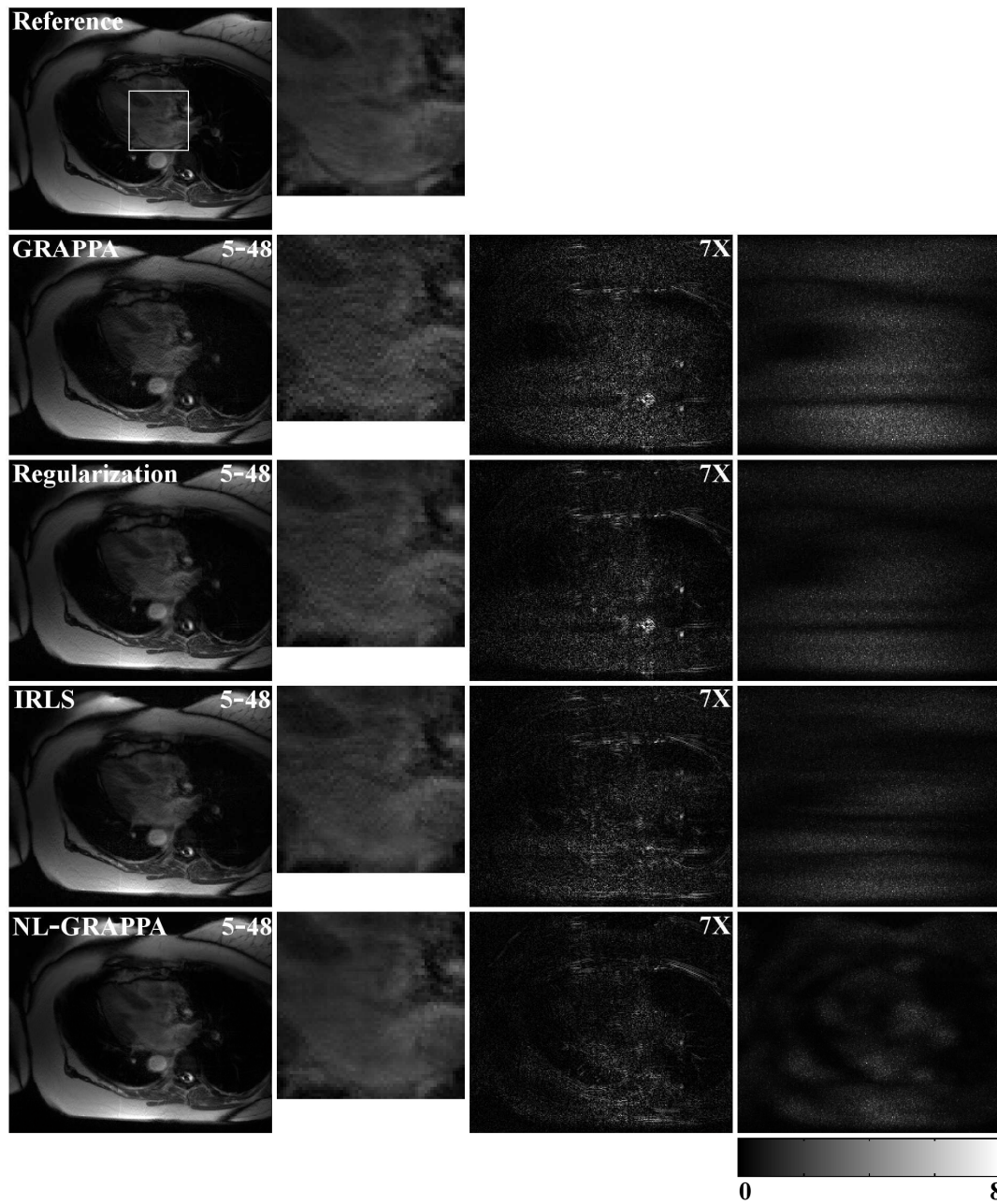


Figure 12. Results from the four-channel cardiac dataset with an ORF 5 and 48 ACS lines. The reconstructed images, zoomed ROI, difference images, and g-factor maps are shown from left to right respectively. They show that the proposed method can remove more noise than other methods while still preserving the resolution.

Figures 10 and 11 show the reconstruction results for the two *in vivo* brain datasets, axial and sagittal respectively. An ORF of 5 and the ACS of 48 were used with a net acceleration of 2.81. Nonlinear GRAPPA used a size of 2 blocks and 15 columns, while the other methods used that of 4 blocks and 9 columns. The difference images with the reference are also shown (amplified 5 and 9 times for display) in both Figs. 10 and 11 and g-factor maps are shown for the axial dataset in Fig. 10. It is seen that the reconstruction using the proposed method achieves a quality superior to all other methods. The proposed method effectively removes the spatially-varying noise in the conventional GRAPPA reconstruction without introducing aliasing artifacts as Tikhonov regularization and IRLS methods do. Furthermore, the proposed method also preserves the resolution of the axial image without blurring. There is only a slight loss of details in the sagittal image due to the tradeoff between noise suppression and resolution preservation (discussed later in Fig. 14).

Figure 12 show the results for the *in vivo* cardiac dataset in long axis. The ORF is 5 and number of ACS lines is 48 (net acceleration of 2.60). The size of the nonlinear GRAPPA coefficients was 4 blocks and 15 columns. The other methods used a size of 4 blocks and 3 columns. The ventricle areas are zoomed to show more details. Both the difference images and the g-factor maps are shown for all methods. The same conclusion can be made that the nonlinear GRAPPA method can significantly suppress the noise in GRAPPA and still preserve the resolution and avoid artifacts.

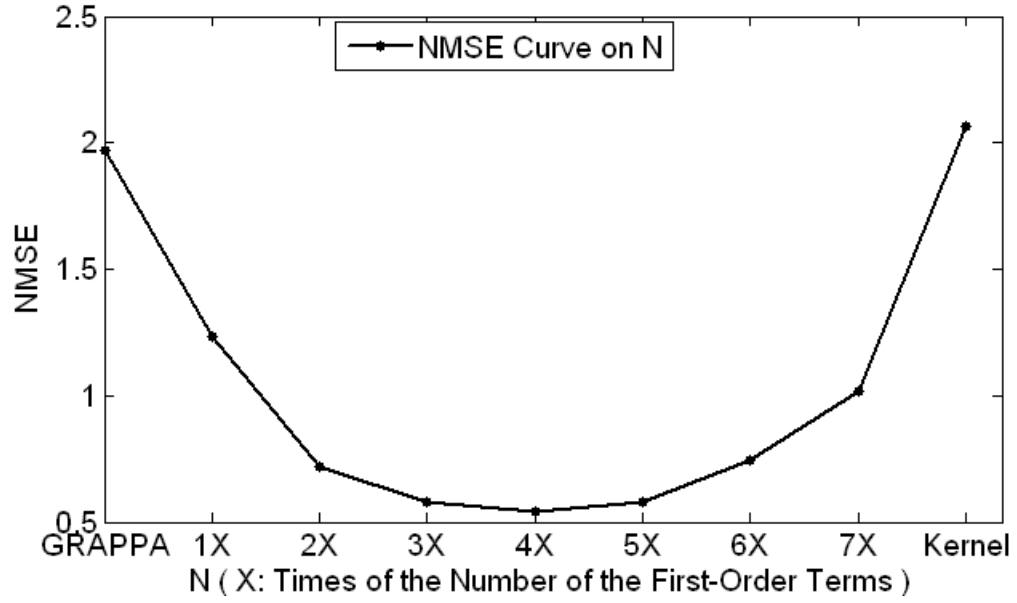


Figure 13. NMSE curves of the proposed method as a function of the number of the second-order terms using the long-axis cardiac dataset. The “U” shape of the curve suggests that some intermediate number should be chosen.

We also used the cardiac dataset to study how the number of second-order terms affects the nonlinear GRAPPA reconstruction quality. Specifically, we truncate all the sorted second-order terms to keep the number to be N times (e.g., 3 times in Eq. (19)) that of the first-order terms. The normalized mean squared errors (NMSE) was calculated and plotted as the function of the number of the first-order terms in Figure 13. In consideration of computational complexity, only the central 64 columns of the 48 ACS lines were used here for calibration. The two endpoints of the curve are the extreme cases of the proposed method. The left one corresponds to conventional GRAPPA without the second-order terms, and the right one is the case where all second-order terms are included (implemented efficiently using kernel representation directly). Figure 14 shows the corresponding reconstructions at some points of the curve. Both the curve and the

images suggest that too small or too large N deteriorates reconstruction quality. When the number N increases, noise is gradually suppressed, but the resolution gradually degrades and aliasing artifacts gradually appear due to the overfitting problem. The optimal range for the value of N to balance the tradeoff between noise, resolution, and aliasing artifacts is seen to be 3-4 times of the number of the first-order terms, according to both the NMSE curve and the images. Because the value of N directly affects the computational complexity, $N = 3$ was chosen and shown to work well for all datasets tested in this study.

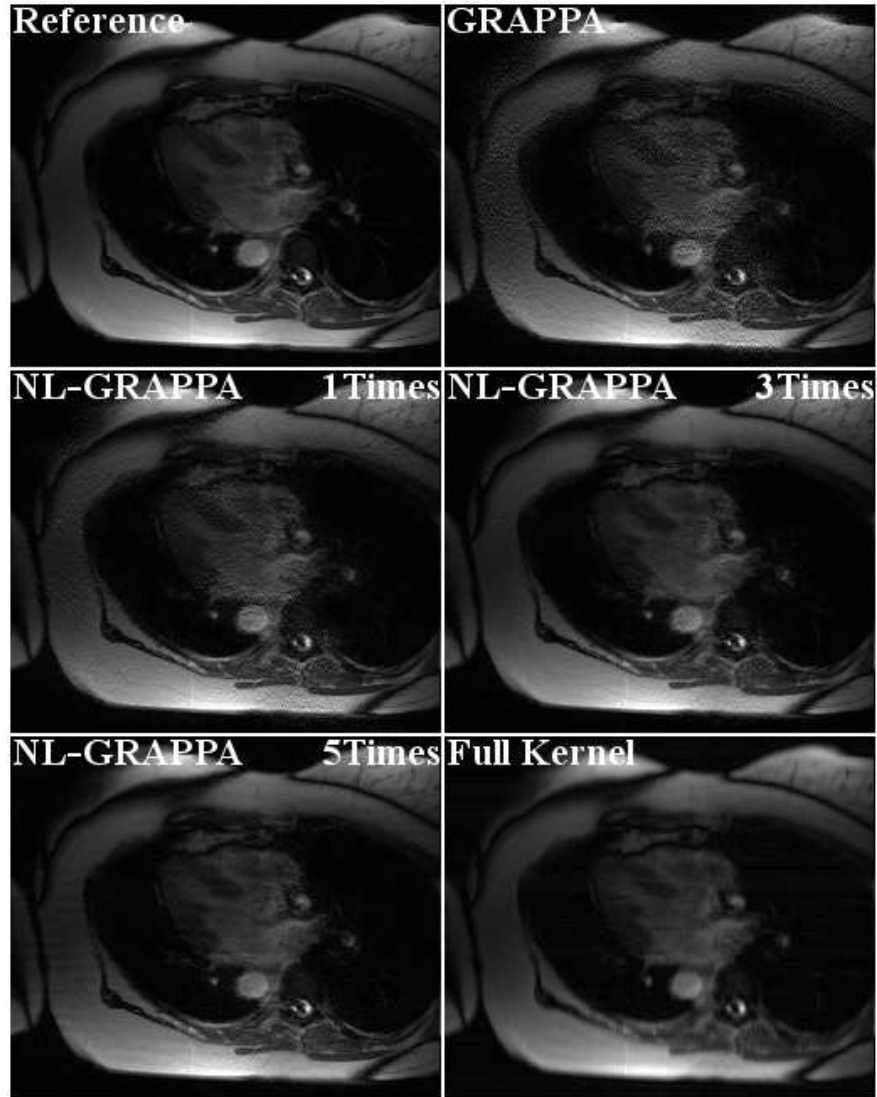


Figure 14. The nonlinear GRAPPA reconstructions with an increasing number of the second-order terms show that the noise is gradually removed but artifacts gradually increase. In the extreme case when all second-order terms are included, both blurring and aliasing artifacts are serious.

3.4 Discussion

We have shown in results that the proposed nonlinear GRAPPA method can outperform GRAPPA at high ORFs but also with a large number of ACS lines. It is interesting to see how the method behaves at lower ORFs or with fewer ACS lines. In Figure 15, we compare GRAPPA and nonlinear GRAPPA with decreasing ORFs when the number of ACS lines is fixed to be 40. At a low ORF of 2, both methods perform similarly well. The proposed method has a slightly lower level of noise. As ORF increases, GRAPPA reconstruction begins to deteriorate due to the increased level of noise. In contrast, the nonlinear GRAPPA method can maintain a similar SNR. Therefore the benefit of nonlinear GRAPPA becomes more obvious at high ORFs. On the other hand, as ORF increases, the required number of ACS lines usually needs to increase to avoid aliasing artifacts.

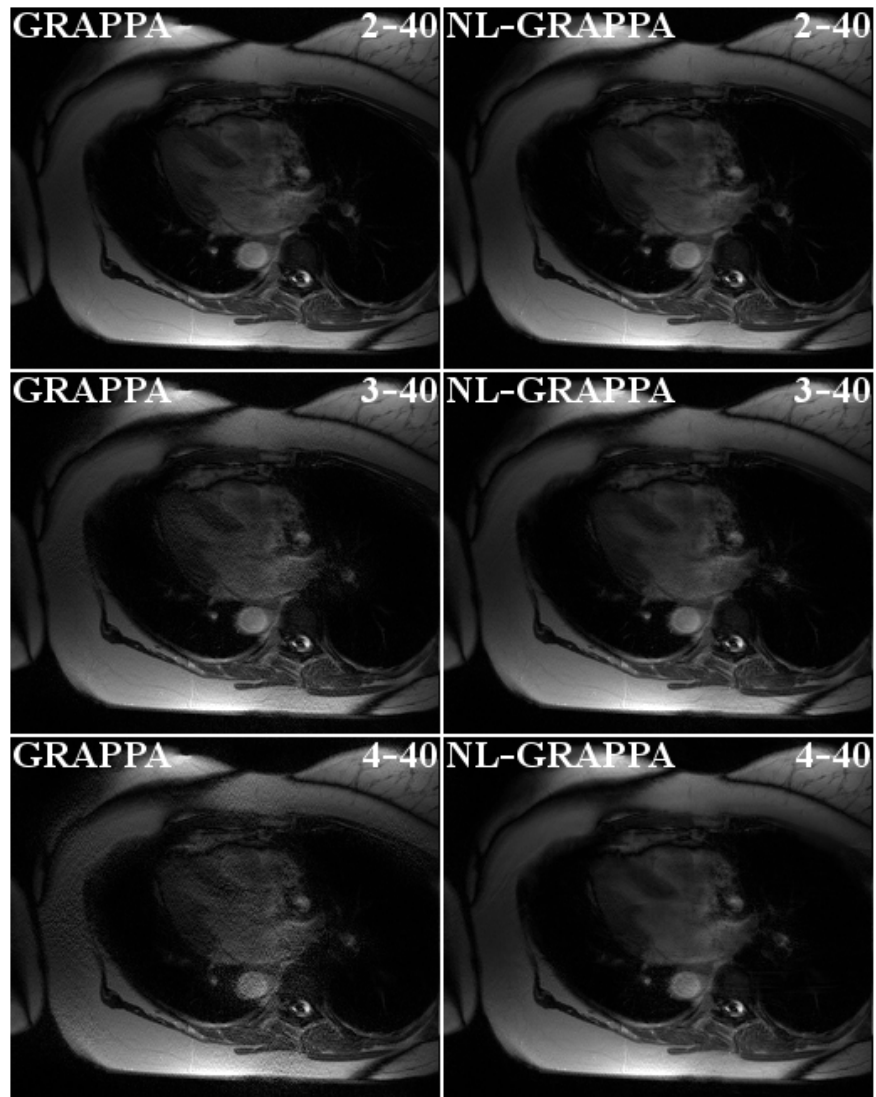


Figure 15. Comparison between GRAPPA and nonlinear GRAPPA when ORF increases with fixed ACS lines. Contrary to GRAPPA, noise in the proposed method does not increase with the ORF.

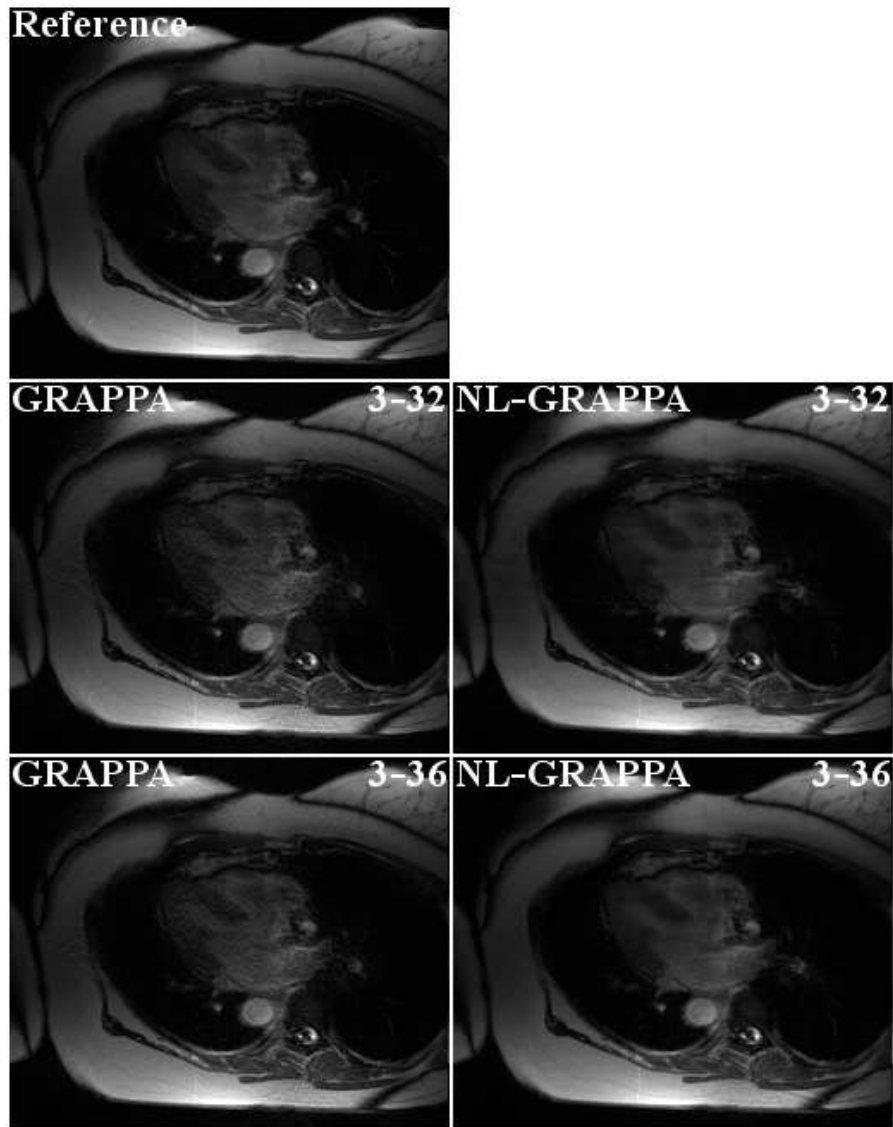


Figure 16. Comparison of GRAPPA and nonlinear GRAPPA reconstructions when ACS increases with fixed ORF. It shows nonlinear GRAPPA needs more ACS lines than GRAPPA to avoid aliasing artifacts, but GRAPPA has more noise than nonlinear GRAPPA.

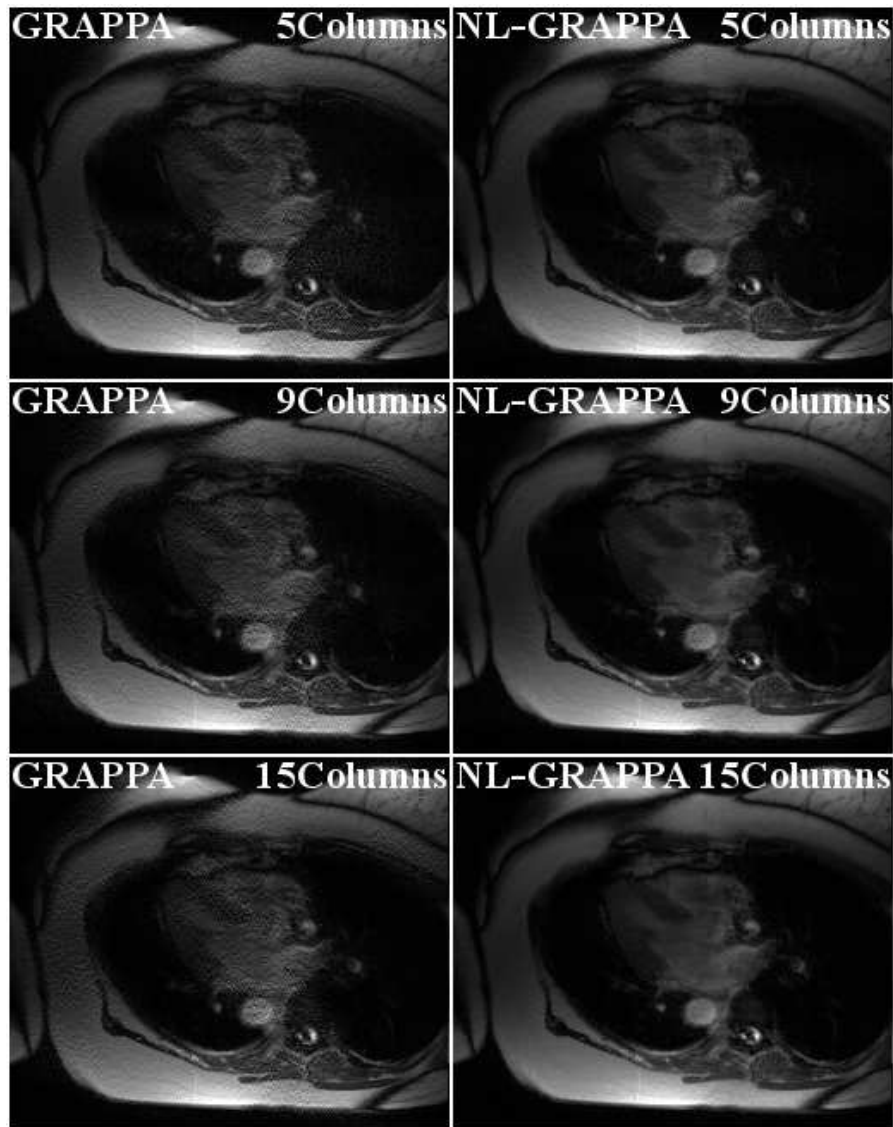


Figure 17. Comparison between GRAPPA and nonlinear GRAPPA when different numbers of columns are chosen for the coefficients. Contrary to GRAPPA, the use of more columns in nonlinear GRAPPA can suppress more noise in reconstruction.

Theoretically the proposed method needs more ACS lines than GRAPPA to set up sufficient number of equations to avoid the aliasing artifacts. This is because there are more unknown coefficients to be solved for in the high dimensional feature space. Figure 16 shows how the reconstruction quality improves when increasing the number of ACS

lines. The improvement of GRAPPA is primarily in terms of noise suppression while the improvement of nonlinear GRAPPA is in aliasing-artifacts reduction. Although more number of ACS lines is needed in the proposed method to avoid artifacts, the ORF can be pushed much higher than GRAPPA and thereby the net acceleration factor can remain low. For example, the combination of ORF of 5 and 48 ACS lines (denoted as 5-48) in Fig. 12 has a higher net acceleration factor to that of 4-40 (net acceleration of 2.57) in Fig. 15 and 3-32 (net acceleration of 2.30) in Fig. 16. The nonlinear GRAPPA reconstruction with 5-48 is always superior to GRAPPA with 5-48, 4-40, or 3-32 combinations.

In GRAPPA, it is known that the size of the coefficients also affects the reconstruction quality. More columns usually improves the data consistency and reduces aliasing artifacts, but at the cost of SNR and computation efficiency [65]. A rather small number of columns (e.g., 3-5 columns) are typically used to balance the trade-off. In the proposed nonlinear GRAPPA, the size of coefficients also plays an important role. Figure 17 shows the GRAPPA and nonlinear GRAPPA reconstructions with 5, 9, and 15 columns of coefficients. Contrary to the observation in GRAPPA, more columns in nonlinear GRAPPA can improve the SNR due to the higher degree of freedom in calibration. In consideration of the computation cost, 15 columns were chosen to be used in our experiments.

The computation time of the proposed method is about 2-5 times that of conventional GRAPPA and Tikhonov-regularized GRAPPA, while IRLS is the most time consuming among all. Furthermore, regularization can also be easily incorporated into the proposed reconstruction method.

4. Accelerating Nonlinear GRAPPA by Compressed Sensing

As experimental results presented in nonlinear GRAPPA reconstruction, although reconstruction quality can be improved by increasing SNR, net acceleration factor is still around 2.8 for 256 phase encoding k -space lines. Compressed sensing (CS) and parallel imaging (PI) have been widely studied for accelerating MRI reconstruction. Furthermore, the serial combination methods of CS and PI have been proposed for even higher speed of reconstruction. However, both reconstructed signals by CS and PI are not as accurate as acquired MR signals, so that errors of CS reconstruction as the first step will be propagated and even amplified to deteriorate PI reconstruction quality in the second step. Based on our previous work – nonlinear GRAPPA (NLGRAPPA), we proposed a novel serial combination of CS and NLGRAPPA (CS-NLGRAPPA) reconstruction. The generalized kernel regression model of NLGRAPPA can remove error effects of inaccurate signal reconstruction by CS. Experimental results using phantom and in vivo data demonstrate that the proposed CS-NLGRAPPA method can significantly improve the reconstruction quality over the existing method and push net reduction factor around 4.

4.1 Combination Methods of CS and pMRI

MR imaging speed is usually limited by the large number of samples needed along the phase encoding direction. In conventional MRI using Fourier encoding, the required number of samples is determined by the field of view and the resolution of the image

based on the Shannon sampling theory. To accelerate conventional MRI, both parallel MRI (pMRI) and compressed sensing (CS) have been used to reduce the number of acquired data. In pMRI, due to the availability of multi-channel coils, the MR images can be reconstructed from multi-channel k -space data sampled below the Nyquist sampling rate using reconstruction methods such as SENSE [6] and GRAPPA [9]. Theoretically, the maximum reduction factor can be up to the number of channels under ideal conditions. However, this maximum usually cannot be achieved due to practical limitations such as noise and imperfect coil geometry. CS-MRI [45, 46] is based on CS theory [67, 68], a new framework for data sampling and signal recovery. CS-MRI takes advantage of the fact that MRI meets two conditions of CS. One is the MR images are sparse or compressible after certain transformations. The other is the Fourier encoding is incoherent with these sparse transformations. Therefore, the MR images can be reconstructed using a nonlinear convex program from data sampled at a rate close to their intrinsic information rate which is well below the Nyquist rate.

Because pMRI and CS reduce sampling based on complementary information, several methods [30, 32, 47, 48] have been developed to combine pMRI and CS for further reduction. Among these methods, the ones that sequentially carry out CS reconstruction for the aliased image of each channel and parallel imaging for the final unfolded image have demonstrated several advantages [32, 47]. First, because the CS and pMRI procedures are clearly decoupled, any CS or pMRI methods are directly applicable without modifications. Second, the sampling pattern can be designed to satisfy different

requirements for CS and pMRI. Finally, the acceleration factor of the combined approach is a product of the factors achieved by CS and pMRI individually.

4.2 Serial Combination of CS and NLGRAPPA

However, an issue with the sequential method is error propagation. The error in CS reconstruction may propagate and be amplified in the pMRI step according to the g-factor. For example in CS-SENSE [32], SENSE with inaccurate coil sensitivities accentuates the CS reconstruction error. In CS-GRAPPA [47], the g-factor of GRAPPA increases dramatically with the acceleration factors and thus magnifies the CS error. We have recently developed a nonlinear GRAPPA (NLGRAPPA) method [66, 69] using kernel method. The method has shown to reduce the g-factor significantly at high acceleration factors.

In this research, we propose a novel sequential approach that performs CS followed by nonlinear GRAPPA. The proposed method, named CS-NLGRAPPA, first employs CS to reconstruct a set of multi-channel aliased images (i.e., k -space data on a uniformly undersampled grid in each channel), and then applies NLGRAPPA to reconstruct the missing k -space data for the final root sum-of-squares (SOS) reconstruction. Due to the nonlinear characterization ability in the generalized regression model, the kernel method can remove error effects generated by CS reconstruction to certain extent. Experimental results demonstrate that the proposed method outperforms CS-GRAPPA in suppressing the spatial-varying noise and aliasing artifacts with high accelerations.

4.2.1 Sampling Pattern

The sampling pattern for the proposed method has two parts. The first part is the same as the sampling used in CS-SENSE [32]. Specifically, a variable-density random sampling scheme [45] is employed to further undersample the data along the phase encoding direction that is already undersampled in pMRI, as shown in Figure 18 from (a) to (b), where the white lines represent the uniformly undersampled k -space locations, while black regions denote the unacquired locations. The other part acquires the fully sampled auto-calibration signal (ACS) at the central k -space. Because the variable-density sampling usually acquires the central k -space in full as well, the two parts have many overlapping lines. The union of the randomly undersampled k -space lines (Figure 18 (b)) and the auto-calibration signal (ACS) lines (Figure 18 (c)) produces the ultimate sampling pattern as shown in Figure 18 (e) for the proposed method. The net reduction factor is calculated by the total number of acquired lines divided by the number of lines when full sampling is employed. As we can see, the net reduction factor for the proposed method in Figure 18 (e) is higher than that for conventional GRAPPA which uses the sampling in Figure 18 (d).

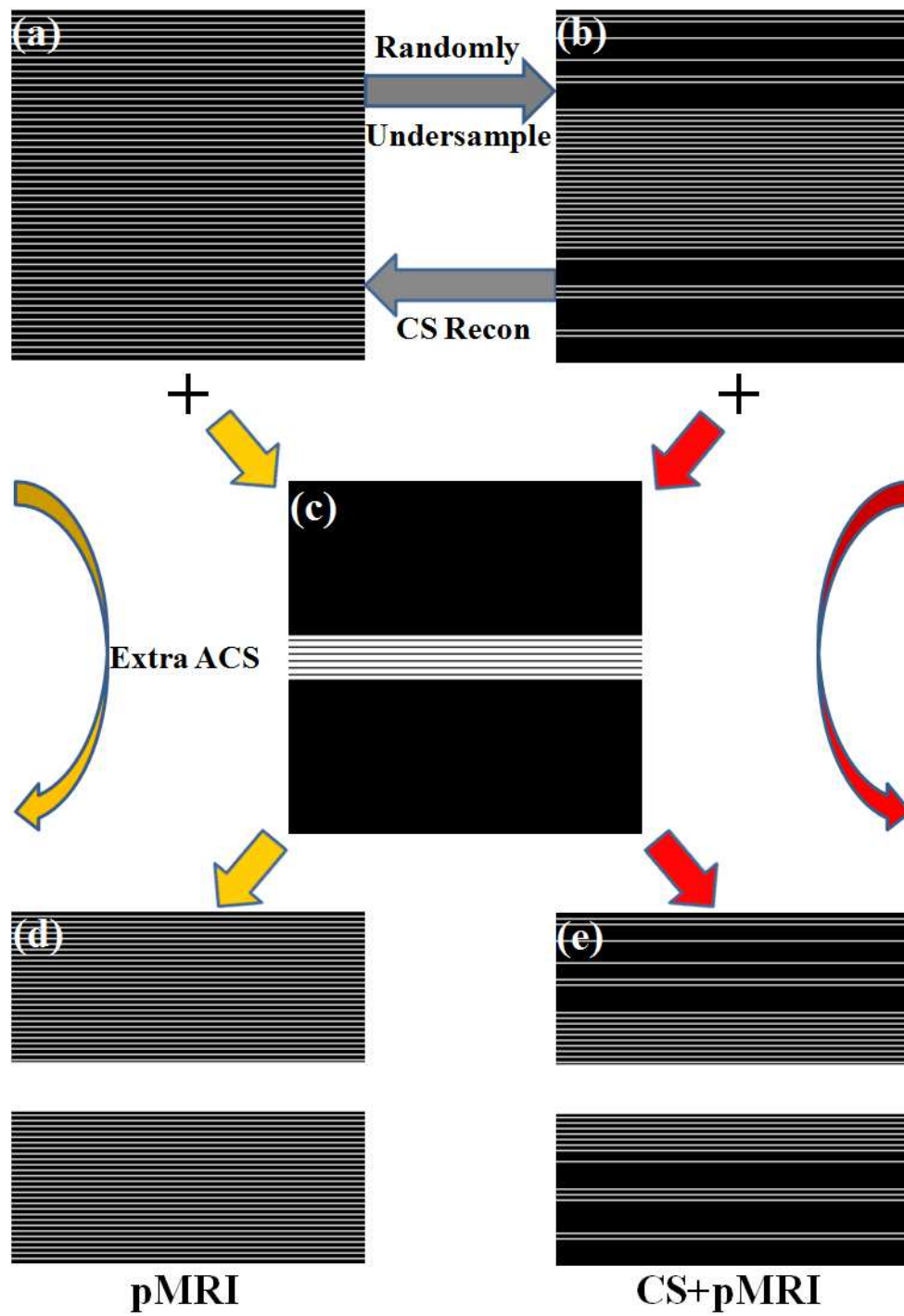


Figure 18. Sampling pattern and CS reconstruction demonstration of the proposed method.

4.2.2 Sequential Reconstruction

The proposed method carries out CS and NLGRAPPA reconstructions sequentially. In the CS step, the uniformly undersampled data are reconstructed from the randomly undersampled data as shown in Figure 18 from (b) to (a). To this end, CS is used to reconstruct the aliased image in each channel and then Fourier transformed to obtain the uniformly undersampled k -space data. If \mathbf{f}_l^A is a vector for the aliased reduced FOV image from the l -th channel, the data consistence comes from

$$\mathbf{F}^u \mathbf{f}_l^A = \mathbf{d}_l^u \quad (20)$$

where \mathbf{F}^u represents Fourier encoding operation and \mathbf{d}_l^u is the undersampled k -space data from l -th channel. Therefore, the aliased image \mathbf{f}_l^A at each channel can be reconstructed using CS

$$\arg \min_{\mathbf{f}_l^A} \left\{ \left\| \mathbf{F}^u \mathbf{f}_l^A - \mathbf{d}_l^u \right\|_2^2 + \lambda_1 \left\| \boldsymbol{\Psi} \mathbf{f}_l^A \right\|_1 + \lambda_2 \text{TV}(\mathbf{f}_l^A) \right\} \quad (21)$$

where $\boldsymbol{\Psi}$ adopts wavelet transform and TV represents total variation.

In the NLGRAPPA steps, the ACS data in Figure 18 (c) are used for calibration and the uniformly undersampled data in Figure 18 (a) obtained from the CS step are used for reconstruction of the final image. NLGRAPPA is used for calibration and reconstruction. The method has shown to suppress the nonlinear error in GRAPPA due to the errors-in-variables problem. In NLGRAPPA, a generalized, nonlinear model is exploited to characterize the relationship between the acquired data and missing data. Kernel method is used to represent the nonlinear relationship such that the linear least squares algorithm

is still applicable for the underlying nonlinear problem. Specifically, NLGRAPPA transforms original input space (original k -space here) to a high dimensional feature space using a nonlinear mapping. A polynomial kernel is used here so that the acquired k -space data are nonlinearly mapped using a truncated second-order polynomial function. The unacquired k -space signal S_j is obtained by a nonlinear combination of the acquired k -space signals S_l :

$$\begin{aligned}
S_j(k_y + r\Delta k_y, k_x) = & w_{j,r}^{(0)} \times 1 + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}^{(1)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2} w_{j,r}^{(2,0)}(l, b, h) \times S_l^2(k_y + bR\Delta k_y, k_x + h\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2-1} w_{j,r}^{(2,1)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \times \\
& S_l(k_y + bR\Delta k_y, k_x + (h+1)\Delta k_x) \\
& + \sum_{l=1}^L \sum_{b=B_1}^{B_2} \sum_{h=H_1}^{H_2-2} w_{j,r}^{(2,2)}(l, b, h) \times S_l(k_y + bR\Delta k_y, k_x + h\Delta k_x) \times \\
& S_l(k_y + bR\Delta k_y, k_x + (h+2)\Delta k_x)
\end{aligned} \tag{19}$$

where w denotes the coefficient set, R represents the outer reduction factor, j is the target coil, l counts all coils, b and h transverse the acquired neighboring k -space data in k_y and k_x directions respectively, and the variables k_x and k_y represent the coordinates along the frequency- and phase-encoding directions, respectively. The nonlinear formulation of Eq. (19) represents a more general model for GRAPPA, which includes the conventional GRAPPA as a special case. It is seen that the first-order terms of nonlinear GRAPPA is equivalent to the conventional GRAPPA, which mainly captures the linear relationship between the missing and acquired undersampled data in the absence of noise and approximations. The second-order terms of Eq. (19) can be used to characterize other

nonlinear effects in practice such that noise and approximation errors are suppressed in reconstruction. The proposed formulation is a nonlinear model in the sense that nonlinear combination of acquired data contributes to estimation of missing k -space data.

Under such a mapping, we still only need to solve the following linear system equation in the feature space:

$$\mathbf{b} = \Phi(\mathbf{A})\mathbf{x} \quad (14)$$

where \mathbf{x} represents the reconstruction coefficients, $\Phi(\mathbf{A})$ is a matrix representing the nonlinear mapping. It is seen that the calibration process to find \mathbf{x} can still be easily and linearly solved in the high dimensional feature space by a linear algorithm:

$$\hat{\mathbf{x}} = \left((\Phi(\mathbf{A}))^H (\Phi(\mathbf{A})) \right)^{-1} (\Phi(\mathbf{A}))^H \mathbf{b} \quad (15)$$

although the relationship between the acquired and missing k -space data is nonlinear. Other linear computational algorithms such as reweighted least-squares and total least-squares methods can also be used here. Once the coefficients are estimated in Eq. (15), they are plugged back in Eq. (14) to reconstruct the missing data in outer k -space, like the conventional GRAPPA does.

4.3 Experimental Results

4.3.1 Experiment Settings

The proposed method was tested on both phantom and *in vivo* data sets. An 8-channel phantom dataset was acquired using a Gradient Echo sequence (TE/TR = 10/100 ms, 31.25 kHz bandwidth, matrix size = 256×256, FOV = 250 mm²). An 8-channel brain

dataset was obtained using a 2D spin echo (SE) sequence ($TE/TR = 11/700$ ms, matrix size = 256×256 , $FOV = 220 \text{ mm}^2$). A set of cardiac cine images was acquired using a 2D trueFISP sequence ($TE/TR=1.67/50.1$ ms, bandwidth 903 Hz/pixel, 45 degree flip angle, 8mm slice thickness, 34 cm FOV in readout direction, 256×146 acquisition matrix) with a 12-channel cardiac coil but combined to 4 channels. All data sets were acquired in full and then manually undersampled to simulate the accelerated acquisition. The sum-of-squares (SoS) reconstruction from the fully sampled data of all channels was used as the reference image for comparison. The reconstructions from reduced data using the CS-GRAPPA [15], the proposed CS-NLGRAPPA method, and SPIRiT method [16] were compared. In both CS-GRAPPA and CS-NLGRAPPA methods, the number of blocks (size of b) is 2 and the number of columns (size of h) is 15 for all datasets. For the eight-channel phantom data, a net reduction factor of 4 was achieved by a combination of 1.5 for CS and outer reduction factor (ORF) = 6 with 42 ACS for GRAPPA/NLGRAPPA. For the eight-channel brain data, the net reduction factor is 3.93 with CS 1.7, ORF = 4, and ACS = 36. In the four-channel cardiac experiment, the net reduction factor is 4.12 with CS 2.0, ORF = 4, and 40 ACS lines.

4.3.2 Results of Phantom and In Vivo Datasets

Figure 19 shows the reconstructions of the phantom using SOS, CS-GRAPPA, and the proposed CS-NLGRAPPA for a net reduction factor 4.00. The size of the coefficients for GRAPPA and nonlinear GRAPPA was 2 blocks and 15 columns. The size of the coefficients could be chosen optimally for each individual method, but the image quality is not sensitive to the change of size within a large range of the optimal choice, so we

fixed its size for both GRAPPA and nonlinear GRAPPA. Furthermore, a small region is zoomed to show more details. It is seen that the CS-GRAPPA suffers from large noise and aliasing artifacts at a net reduction factor of 4.00. The proposed CS-NLGRAPPA method can suppress both noise and artifacts at such a high acceleration without additional artifacts or loss of resolution.

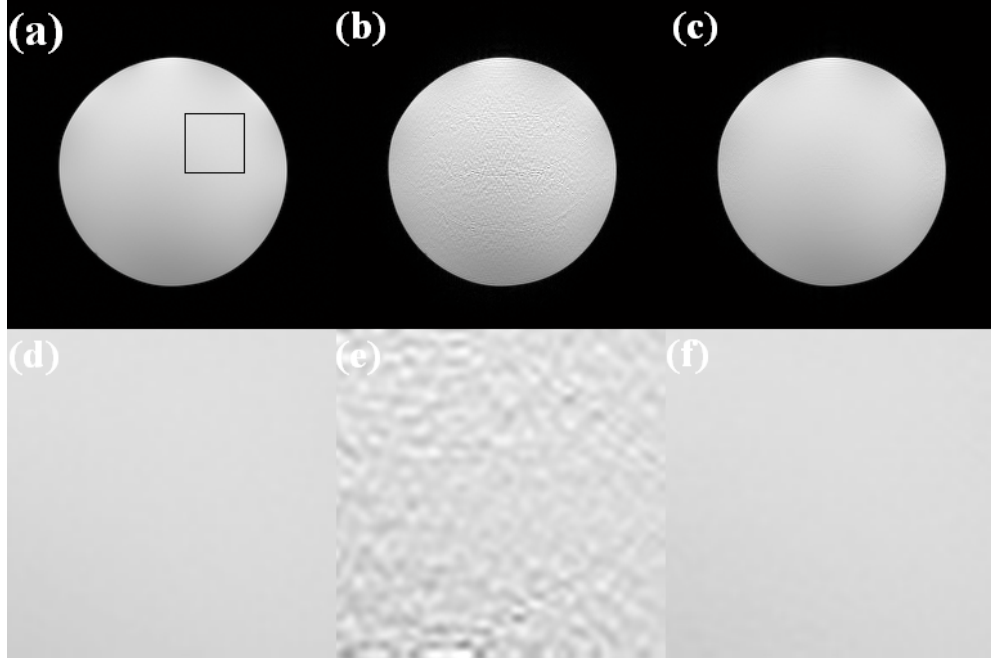


Figure 19. With (a) the SoS reconstruction of the 8-channel phantom image as the reference, we compare (b) CS-GRAPPA and (c) the proposed method at a net reduction factor of 4. Subfigures (d) – (f) show zoomed-in regions of (a) – (c) respectively.

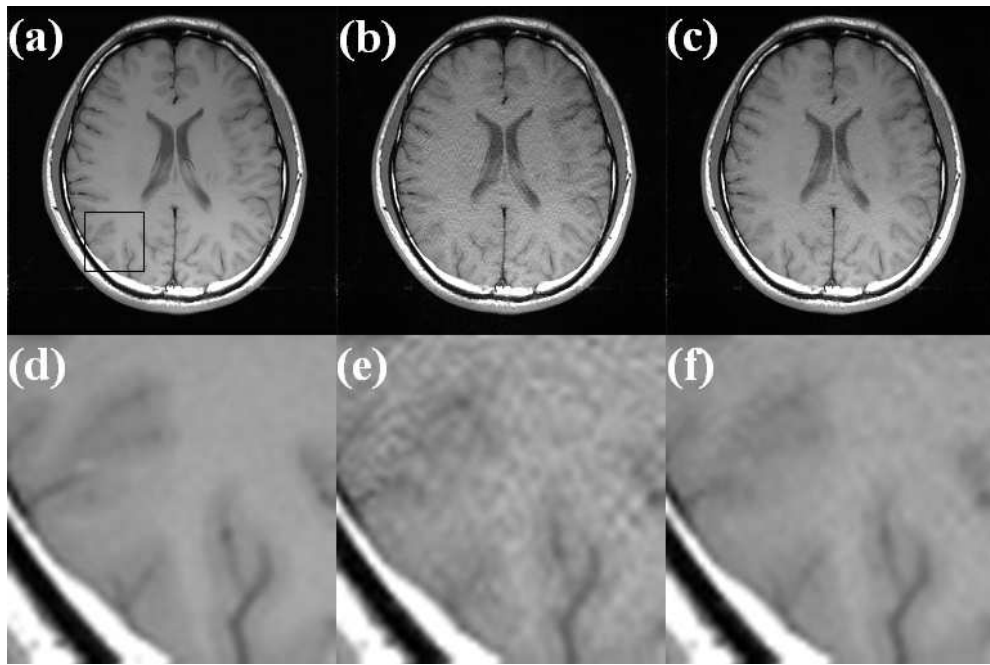


Figure 20. With (a) the SoS reconstruction of the 8-channel brain image as the reference, we compare (b) CS-GRAPPA and (c) the proposed method at a net reduction factor of 3.93.

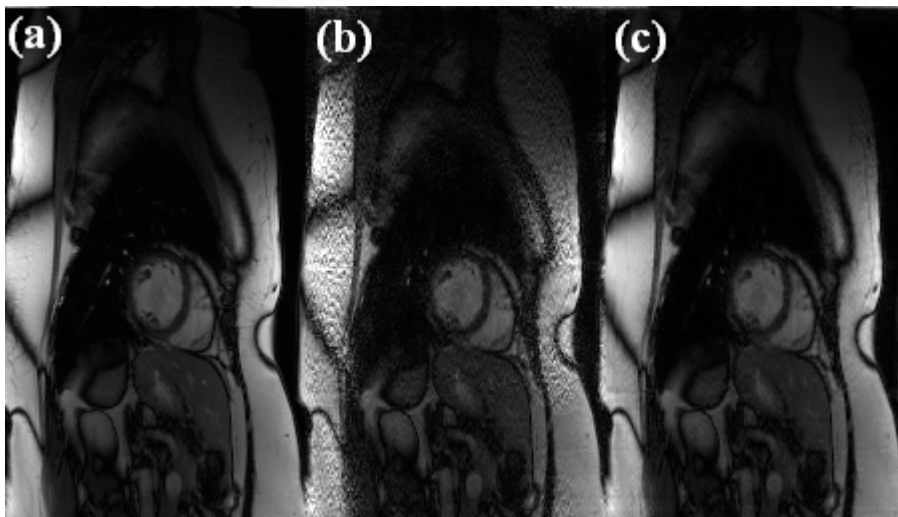


Figure 21. Comparison of (a) SoS, (b) CS-GRAPPA and (c) the proposed method in reconstructing the four-channel cardiac image at a net reduction of 4.12.

Figure 20 shows the reconstruction results from a set of *in vivo* axial brain data. An ORF of 4 and the ACS of 36 were used with reduction for CS 1.7, so that a net acceleration of 3.93 can be achieved. Both GRAPPA and nonlinear GRAPPA still used a size of 2 blocks and 15 columns. It is seen that the reconstruction using the proposed method achieves a quality superior to CS-GRAPPA method. As shown in the zoomed regions, the proposed method not only effectively removes the spatial-varying noise and aliasing artifacts in the CS-GRAPPA reconstruction, but also preserves details well without blurring boundaries. The CS-NLGRAPPA can push the net reduction factor to 3.93.

The reconstruction results of the cardiac dataset in long axis are presented in Figure 21. The ORF is 4 with reduction for CS 2.0 and number of ACS lines is 40 (net acceleration of 4.12). The size of the GRAPPA and the nonlinear GRAPPA coefficients were 4 blocks and 15 columns. The difference images with the reference are also shown (amplified 5 times for display) in Fig. 4. In contrast to the CS-GRAPPA reconstruction, the reconstruction by the proposed method is much cleaner with a slight loss of resolution at a high net reduction factor of 4.12. The same conclusion can be made that the CS-NLGRAPPA method can significantly suppress the noise in CS-GRAPPA and still preserve the resolution and avoid artifacts at very high net acceleration factors. Furthermore, all methods were implemented in MATLAB programming environment, and codes were run on a HP workstation with four 3.2-GHz processors and 32-GB RAM. The computational time of both methods is compared in Table 1. The proposed method takes about 1.7 - 2 times of CS-GRAPPA for reconstruction.

	Phantom	Brain	Cardiac
CS-GRAPPA	208 s	140 s	67 s
CS-NLGRAPPA	429 s	267 s	116 s

Table 1 Comparison of computational time between CS-GRAPPA and CS-NLGRAPPA

We also compare the proposed method with SPIRiT [48] - an integrated approach of CS+pMRI reconstruction, which needs to modify the form of the traditional GRAPPA that undersamples k -space along 1D phase encoding direction. Because SPIRiT method has different sampling pattern that adopts 2D random sampling on k -space, central 32×32 calibration area is fully sampled here, while outside k -space is randomly sampled. But the net acceleration factor of the SPIRiT method is strictly set as 3 and 4 for the phantom and brain datasets for comparison with the proposed method. As we can see in Figure 22, due to random sampling on 2D direction, SPIRiT method displays incoherent aliasing artifacts in the reconstructed image. On the other hand, the proposed CS-NLGRAPPA method can outperform the SPIRiT reconstruction without incoherent aliasing artifacts and with higher SNR.

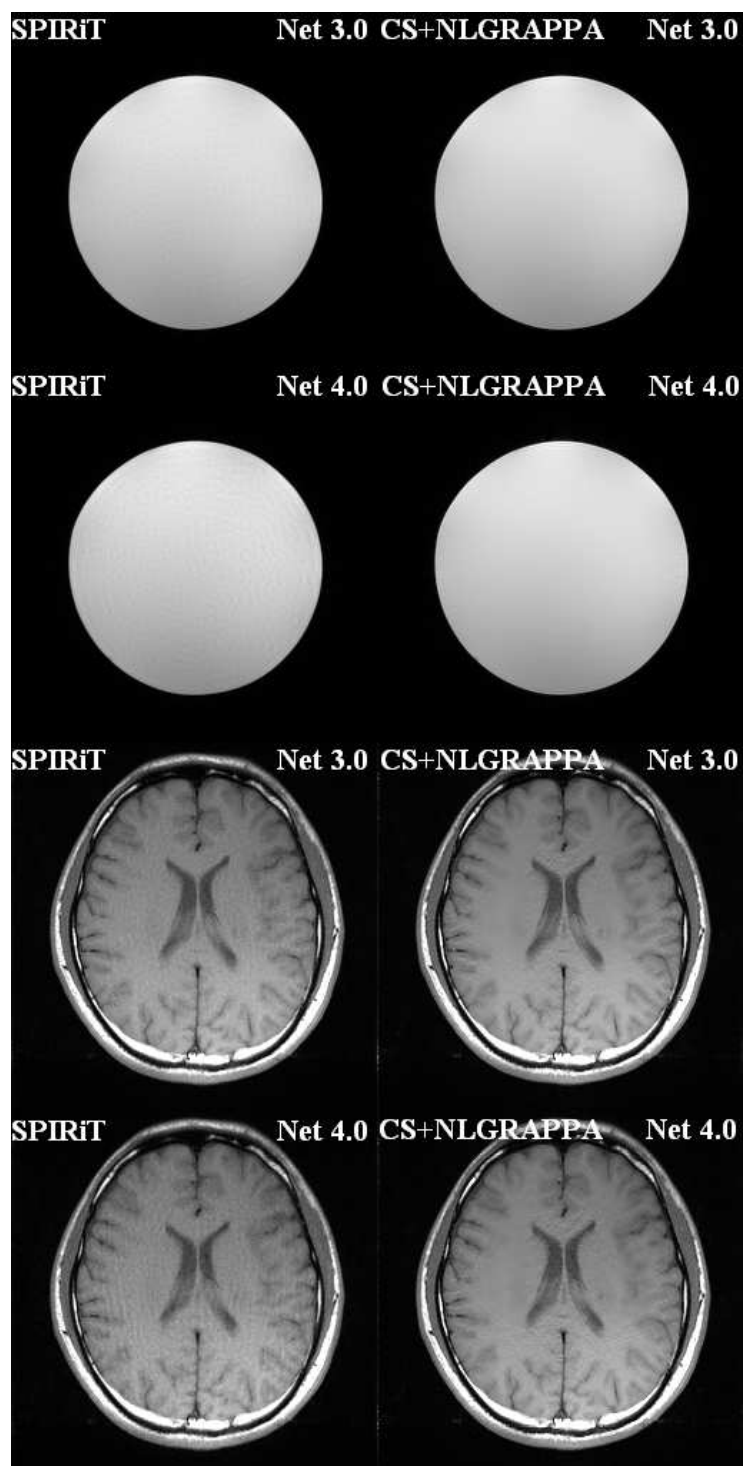


Figure 22. The performance comparison between SPIRiT and the proposed CS-NLGRAPPA method. Reconstructions of the phantom and the brain data sets are presented at net reduction factor 3.0 and 4.0, respectively.

4.4 Discussion

We investigate how different numbers are to be distributed among the CS, ORF and ACS to achieve the optimal quality at a fixed net reduction factor. It is known that high accelerations in CS can lead to aliasing artifacts and loss of resolution. On the other hand, high ORFs usually deteriorate SNR and require more ACS to avoid aliasing. The study was carried out on the eight-channel brain dataset. The results are shown in Figure 23. For a fixed net reduction factor of 3.93 and ORF of 4 for the NLGRAPPA step, we tested on three different combinations of ACS and reduction for CS (R_{CS}). When 26 ACS lines and $R_{CS} = 1.4$ were used, aliasing artifacts appear due to insufficient number of ACS lines. The appearance of artifacts is in form of folding that is typical of parallel imaging. When 44 ACS lines and $R_{CS} = 2.0$ were used, aliasing artifacts appear due to propagation of CS reconstruction error. The appearance is rather incoherent due to the random sampling. The intermediate combination of 36 ACS and $R_{CS} = 1.7$ generates the optimal reconstruction for this dataset.

On the other hand, we fixed ACS to observe reconstruction results by various ORFs and reductions for CS as shown in Figure 24. For the fixed 44 ACS lines, ORF of 4 and 5 were used on the horizontal direction; reductions for CS were sets as 1.5, 2.0, and 2.5 on vertical direction, respectively. Following the increasing ORF, aliasing artifacts in form of folding gradually appear as we can see from left to right side on the horizontal

direction, which is also a typical parallel imaging observation. From the top to bottom in the same column, some additional incoherent aliasing artifacts deteriorate the reconstructed image due to randomly sampling for CS reduction, which has the same conclusion in Figure 23. Therefore, both kinds of aliasing artifacts caused by parallel imaging and compressed sensing exist in the proposed method. If reconstruction quality is acceptable, both ORF and reduction for CS can be pushed even higher for fast imaging with balance of both aliasing artifacts. In general, the reduction in CS cannot be high due to error propagation. A factor of lower than 2 is usually used.

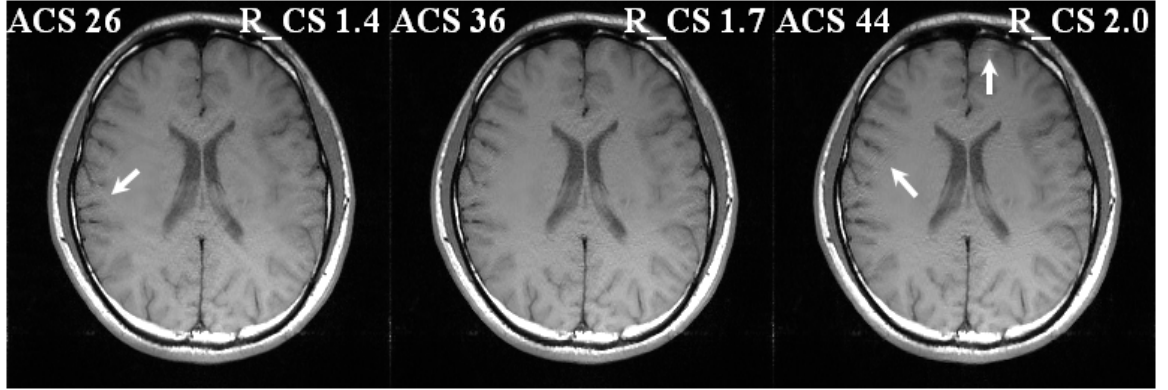


Figure 23. CS-NLGRAPPA reconstructions with three different combinations of ACS and R_{CS} with fixe ORF of 4 and a net reduction factor of 3.93.

In this research, we propose a novel method that combines CS with nonlinear GRAPPA sequentially for fast imaging. NLGRAPPA is applied on the k -space data reconstructed by CS. The experimental results demonstrate that the proposed CS-NLGRAPPA [70] method is superior to CS-GRAPPA in suppressing both noise and artifacts when a high net reduction factor is used. However, more experiments are needed to be done to study noise source in k -space, based on which, general guidelines of

choosing sampling pattern may be concluded. Iterative strategy will also be studied for even better CS-NLGRAPPA reconstruction.

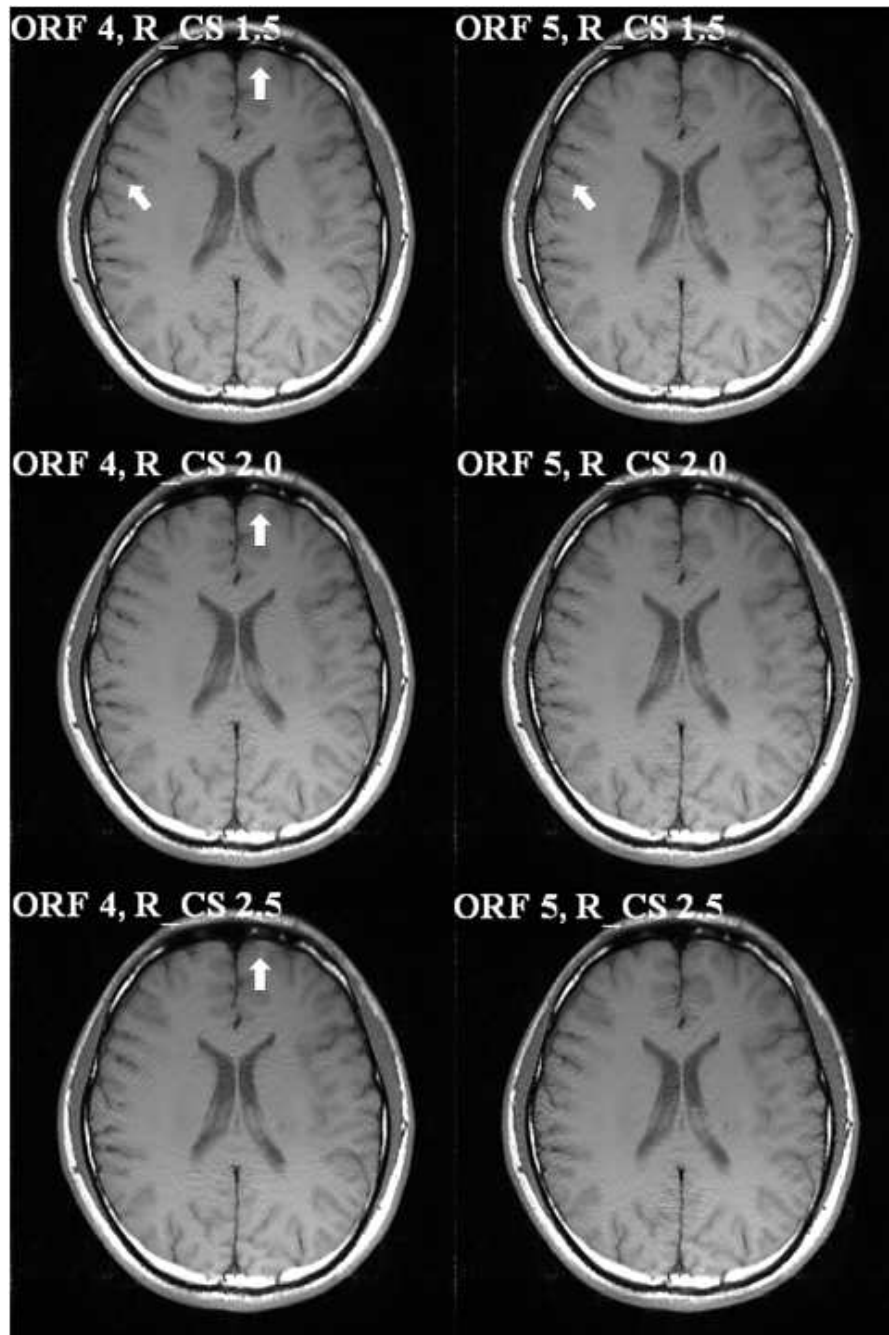


Figure 24. CS-NLGRAPPA reconstructions with six different combinations of RCS with various ORF used the fixed ACS.

5. Nonlinear Channel Reduction for Efficient GRAPPA Reconstruction

5.1 Review of Channel Compression

Due to success of pMRI methods, more and more coils have been used in MRI systems [50, 72-74] for higher reconstruction quality and acceleration such as 128 coil arrays on research MRI scanners [50].

However, computational cost provided by a large number of element coils can be significant. Some channel reduction methods [43-44, 51-54, 75-84] have been proposed to reduce reconstruction time. A hardware RF signal combiner inline after preamplification was placed to construct an eigencoil array [54, 75] by using the noise covariance of the receiver array. Optimal SNR can be achieved with receiver channel reduction, and the almost the same reconstruction quality can be kept by using a subset of channels (virtual channels). However, it needs hardware design and is not suitable for commonly used coils that have existed in the scanner.

In contrast, software-based channel reduction methods provide a flexible way to reduce the computation cost for parallel MRI reconstruction. For example, coil array compression using principal component analysis (PCA) for MRI with large coil arrays was proposed for SENSE reconstruction [51] that is a sensitivity-based parallel imaging reconstruction method. The combination of coils can be expressed as a linear transformation matrix which reduces the number of physical coil channels to a subset of virtual channels. The number of channels can be largely reduced instead of only a little SNR loss, so that reconstruction time is accelerated with sacrificing only a very little

degeneracy of quality. Doneva et al [52] proposed an automatic coil selection approach using singular value decomposition (SVD) in SENSE-based parallel imaging reconstruction. The approach makes use of the coil sensitivity information and takes reduction factor and phase encoding direction into account. Although the proposed method is computationally efficient without remarkable image quality degradation, it still can't improve reconstructed image's quality. Furthermore, channel reduction technique was also successfully applied on k -space-based reconstruction approach such as GRAPPA. Due to reduced channels of having very low sensitivities make little contribution to the SNR of the final reconstructed image [43], two-stage channel compression and incorporation of all acquired data into the final reconstruction have been used for reducing computational time of GRAPPA reconstruction [44]. However, reconstruction quality is still a little degenerative, although computational and memory costs have been largely reduced. A trade-off exists between loss of information that degenerates reconstruction quality and reduced channels that reduce computational cost. Furthermore, the increase in computation time and memory provided by channel-by-channel reconstruction can also be addressed by Direct Virtual Coil (DCV) technique [76] that synthesizes unaccelerated data for a virtual coil from undersampled data acquired by multiple source coils. Through the image quality comparison between channel-by-channel parallel imaging reconstruction and the DVC reconstruction approach [77], the DVC reconstructions were determined to be nearly identical to the channel-by-channel reconstructions. Similarly, a single synthetic target coil [78] was proposed to combine images prior to the training instead of the channel-by-channel reconstruction, so that

significant computation gain can be achieved since image combination is performed on much smaller matrix size. Convolution based training and aliasing only need to be performed once a time instead of for each channel.

In addition, a dynamic coil selection algorithm [79] is presented by a subset of receive coils that are selected for reconstruction using a coil ranking based on the distance to the current slice or catheter position. The center-of-sensitivity coordinates and relative signal intensities have to be determined for each coil in a prescan. Animal experiments involving catheter manipulation in the aorta and the right heart chamber demonstrate that the anatomy was successfully visualized at frame rates of about 5Hz using active catheter tracking in interventional applications. Similarly, since the channel sensitivity is localized in large arrays, channel reduction is achieved through channel cross correlation for channel-by-channel reconstruction [53]. The almost the same reconstruction quality compared with reconstruction using all channels can be achieved. Furthermore, computational burden was reduced by omitting information from physical coils or “virtual coils” that don’t contribute significantly to image SNR in ROI [80]. The sampled data from coil arrays with low signal energy content are discarded.

Due to the success of software channel reduction methods have been applied on reconstruction by 2D data acquisition, these methods are also extended to 3D datasets [81, 82]. A PCA-based software channel compression method was proposed for 3D datasets, and the compressed data was then reconstructed by GRAPPA with sliding blocks. A novel channel compression technique for 3D Cartesian sampling dataset was proposed recently [82] that exploits the spatial varying coil sensitivities along fully sampled

directions for better compression and computation reduction. From the experimental results, 3D datasets from a 32 channel pediatric coil can be effectively compressed into 6 virtual coils without noticeable compression loss.

Besides the reduction of the computation load, channel reduction can also have the denoising effect [83]. For the regions of low signal intensity in the straight root of sum of squares reconstruction of MR data, increased number of channels implies more noise. The property of PCA is to concentrate the useful information into fewer channels. The channels with low Eigen values (for example less than 1% of the maximum) can be considered as noise and hence the corresponding channels can be left unused and discarded. However, such denoising effect was only briefly discussed in Ref. [83]. No significant improvements have been reported. Furthermore, channel compression techniques have been also extended to other imaging methods such as BLADE [84]. In this dissertation, I apply kernel PCA to suppress noise and aliasing artifacts as well as compress channels for accelerating reconstruction time.

5.2 The Proposed Method: Kernel PCA-based GRAPPA

5.2.1 PCA-based GRAPPA

PCA [38] is a basis transformation to diagonalize an estimate of the covariance matrix of the data. The new coordinates in the Eigenvector basis are called principal components. PCA finds an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so

on. When applied to channel reduction, the ACS data is used to obtain the transformation which is then applied to all acquired data to obtain a new set of data in the new coordinate system. Mathematically, the linear transformation \mathbf{W} can be obtained by the eigen-decomposition of the covariance matrix of the ACS data:

$$\mathbf{A}^H \mathbf{A} = \mathbf{W}^H \mathbf{\Sigma} \mathbf{W} \quad (22)$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_L]$ is composed of vector \mathbf{a}_l formed from the ACS data of the l^{th} channel (L channels in total) after removing the mean, and \mathbf{W} and $\mathbf{\Sigma}$ are matrices with eigenvectors and eigenvalues. The new coordinates in the eigenvector basis are known as principal components. It is assumed that the directions of largest variance represent interesting information and those of smallest variance represent uninteresting noise. For reduction, only the first few eigenvectors which correspond to the largest eigenvalues are kept to form the linear transformation \mathbf{T} . This transformation matrix is then applied to the acquired k -space data to obtain the orthogonal projections onto the eigenvectors, which gives a new set of reduced virtual channels. Conventional GRAPPA reconstruction is then applied on the new dataset to reconstruct missing data in the transform domain. In Ref. [43], it is noted that the number of source channels (N_{sch}) and the number of target channels (N_{tch}) after PCA reduction can be different and one might be larger than the other to obtain the best results with the same computation time. The final root sum-of-square (SOS) combines all the images from the virtual channels to generate the final image.

Kernel principal component analysis (kernel PCA) is an extension of principal component analysis (PCA) by using techniques of kernel methods. Using a kernel, the

originally linear operations of PCA are done in a reproducing kernel Hilbert space with a nonlinear mapping. It can be considered as a natural extension of linear PCA and is very well suited to extract interesting nonlinear structures in the data. It has proved useful for various applications, such as denoising [39-41] and as a pre-processing step in regression problems [42]. Since PCA only linearly transforms data to new coordinates in the given space so that data relationship can't be described more accurately than that in input space. KPCA can be used by explicit or implicit mapping [39]. The first one directly nonlinearly maps original k -space data into new coordinates by a nonlinear mapping Φ to construct the covariance matrix, and then uses conventional linear PCA to decompose the covariance matrix to obtain the compressed signals, which has the almost the same procedures with linear PCA. The second one uses kernel trick for predicting missing k -space data, which needs longer computational time and more memory, since kernel trick matrix is generally very large. Since PCA has been successfully applied on channel reduction for GRAPPA reconstruction [43, 44], we studies kernel PCA for better channel reduction with the almost same reconstruction time and improved reconstruction quality.

As shown in Figure 25, compressed signals produced by PCA are input into GRAPPA reconstructor and then final image is obtained by sum-of-square (SOS) to combine each image that is produced by IFT each virtual channel. Our investigation is replacing PCA with kernel PCA for better GRAPPA reconstruction with almost the same computational costs.

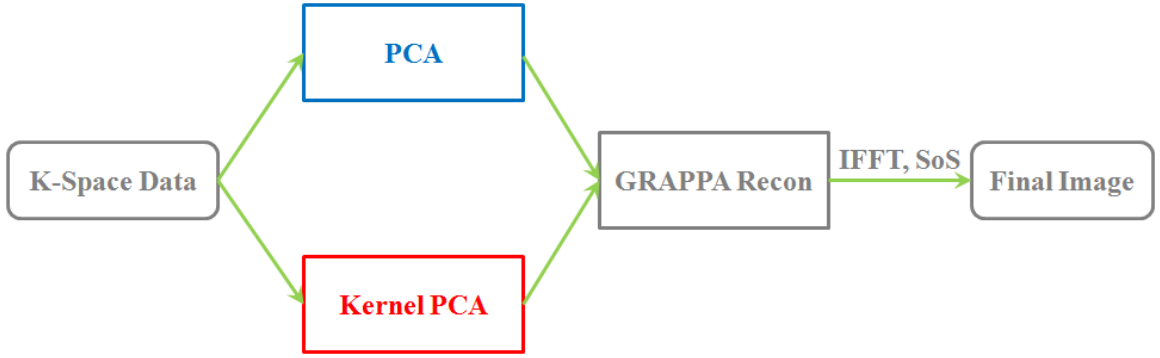


Figure 25. Demonstration of PCA and KPCA roles in channel reduction for GRAPPA reconstruction.

5.2.2 Proposed Kernel PCA-based GRAPPA

KPCA can be used by explicit or implicit mapping [40]. In this paper, we adopt explicit mapping in KPCA. To apply KPCA to channel reduction, we map the ACS data matrix \mathbf{A} into a new feature space $\Phi(\mathbf{A})$ through a nonlinear transformation, and then the data in the new space is ensured to have zero-mean by subtracting the average value among all ACS data points. PCA can be worked in the feature space constructed by products of vector elements, which takes into account higher-order statistics. We then do linear PCA for the new data in the feature space. Specifically, for the covariance matrix $\Phi(\mathbf{A})^T \Phi(\mathbf{A})$, eigenvectors that correspond to the largest few eigenvalues are maintained in matrix \mathbf{T} , which will be used to project the data in feature space to a new space with reduced number of virtual channels. Parallel imaging reconstruction methods such as GRAPPA can then be applied to this new dataset, where the computation time can be reduced due to fewer channels.

Since polynomials are widely used to approximate smooth unknown functions, we choose an inhomogeneous polynomial kernel for Φ mapping. It has the following form

$$\kappa(\mathbf{a}, \mathbf{b}) = (\gamma \mathbf{a}^T \mathbf{b} + r)^d \quad (23)$$

where λ and r are scalars and d is the degree of the polynomial. Another advantage of polynomial kernel lies in the fact that its corresponding nonlinear mapping $\Phi(\mathbf{a})$ such that $k(\mathbf{a}, \mathbf{b}) = \langle \Phi(\mathbf{a}), \Phi(\mathbf{b}) \rangle$ has explicit representations. For example, if $d = 2$, $\Phi(\mathbf{A})$ maps the original L -channel data \mathbf{A} to:

$$\Phi(\mathbf{A}) = [r^2, \sqrt{2\lambda r} \mathbf{a}_1, \dots, \sqrt{2\lambda r} \mathbf{a}_L, \lambda \mathbf{a}_1^{(2)}, \dots, \lambda \mathbf{a}_L^{(2)}, \sqrt{2\lambda} \mathbf{a}_1 \square \mathbf{a}_2, \dots, \sqrt{2\lambda} \mathbf{a}_i \square \mathbf{a}_j, \dots, \sqrt{2\lambda} \mathbf{a}_{L-1} \square \mathbf{a}_L]^T \quad (24)$$

where $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_L$ are vectors representing different channels, superscript $^{(2)}$ means piecewise square, and \square denotes piecewise multiplication. It is seen that the vector includes the linear terms in the original space as well as the constant and second-order terms, and $\Phi(\mathbf{A})$ has $(L+2)(L+1)/2$ terms in total.

Since there are too many terms, we truncate the second order terms to save computation. Specifically, we sort the second-order terms according to the following order. We first have the square terms within each coil, and then the product terms between the nearest neighbors, the next-nearest neighbors, and so on so forth in k -space. The above order is then repeated for terms that are across different coils. With the sorted terms, we can truncate the vector $\Phi(\mathbf{a})$ according to the desired dimension of the feature space. If we truncate all second-order terms, the proposed method is equivalent to linear PCA-based channel reduction.

The target channels are defined as the data that correspond to the left hand side of Eq. (2) and source channels as those for the right hand side of (2). Since the target channel and source channel in GRAPPA are rather independent [44], different feature spaces can

be used. We choose the original space for the target channel to avoid the complication of transforming data back from the feature space to the original space. The source channels are only used for estimation and thereby do not need to be transformed back to the original space. In this study, we choose the number of second-order terms to be three times of that of the first-order terms for constructing the source channels.

The parameters r and λ control the nonlinearity incorporated into the new coordinates. If the nonlinearity is too high, original channel information will be lost and the reconstructed image will be seriously distorted. On the other hand, if the nonlinearity is too small, it is almost equivalent to original PCA-based channel reduction method, in which case the noise is not suppressed. We let $\sqrt{2\lambda r}=1$ and adjust λ for optimal performance. Specifically, we first find the maximum absolute value M_{2nd} of the second-order terms to be used for constructing the feature space. We then let λ_{source} take any value in the range between $(1/ M_{2nd}, 10/ M_{2nd})$. It is found empirically that the performance is insensitive to the values within this range.

5.3 Analysis of Computational Expense

The computational expense can be quantitatively evaluated by using matrix multiplication measure [85]. For the 2D GRAPPA [13], the dominant calibration cost is the matrix inversion step that requires $N_p N_x (N_c d_x d_y)^2$ complex-valued multiplications, where N_p is the number of phase-encoding lines that are possible fit locations along phase-encoding direction, N_x is the number of points along the frequency-encoding direction, N_c is the total number of all channels for original k -space data, and d_x and d_y are the convolution size of 2D GRAPPA along the frequency-encoding and phase-

encoding directions respectively. Furthermore, the dominant computational expense for the synthesis step of GRAPPA needs $N_u N_x N_c^2 d_x d_y (R-1)$ complex-valued multiplications, where N_u is the number of phase-encoding lines to be synthesized at a particular phase-encoding offset location, and R is outer reduction factor (ORF). Therefore, the whole computational expense for 2D GRAPPA reconstruction [85] is $N_p N_x (N_c d_x d_y)^2 + N_u N_x N_c^2 d_x d_y (R-1)$. Due to PCA-based channel reduction for GRAPPA reconstruction, the computational expense can be reduced to $N_p N_x (N_{sch} d_x d_y)^2 + N_u N_x N_{sch} N_{tch} d_x d_y (R-1)$ [44], where N_{sch} and N_{tch} are the number of source channels and target channels reduced by PCA.

Dominant computational expense of PCA or KPCA hasn't been analyzed so far. Similar to GRAPPA reconstruction, channel reduction is also an essential part of the whole reconstruction procedure. The whole computational cost depends on both of them. If we can prove that the dominant computational expense of PCA and KPCA only take a very small portion of the computational cost of GRAPPA reconstruction, we can explain why KPCA-GRAPPA could achieve the almost the same reconstruction time to PCA-GRAPPA, since GRAPPA takes the dominant reconstruction time in the whole procedure.

The dominant computational expense of KPCA focuses on the construction of covariance matrix and SVD. For the first one, the construction of covariance matrix in KPCA requires $N_{acs} N_x (4N_c)^2$ complex-valued multiplications. Furthermore, the computational cost of SVD for a complex-valued matrix with $m \times n$ size is $16m^2n + 24mn^2 + 29n^3$ [86]. Therefore, set $m=n=4N_c$, and total computational expense for KPCA is $N_{acs} N_x (4N_c)^2 + 4416N_c^3$. Similarly, the computational expense for PCA is

$N_{acs}N_x(N_c)^2+69N_c^3$. After obtain these computational expenses, ratios of computational expenses can be obtained as follows.

$$\frac{CE_{GRAPPA}}{CE_{KPCA}} = \frac{N_p N_x (d_x d_y N_{sch})^2 + N_u N_x N_{sch} N_{tch} d_x d_y (R-1)}{N_{acs} N_x (4N_c)^2 + 4166N_c^3} \quad (25)$$

, and

$$\frac{CE_{KPCA}}{CE_{PCA}} = \frac{N_{acs} N_x (4N_c)^2 + 4166N_c^3}{N_{acs} N_x N_c^2 + 69N_c^3} \quad (26)$$

If we adopt commonly used reconstruction parameters [65] that $N_{acs}=40$, $N_x=256$, $d_x=9$, $d_y=4$, $R=4$, $N_u=168$, $N_{sch}=N_{tch}=16$, $N_c=32$, $SlideBlock=2$, $N_p \approx SlideBlock \times N_{acs} \times (R-1) = 240$, then the ratios are

$$\frac{CE_{GRAPPA}}{CE_{KPCA}} \approx 70.9 \text{ and } \frac{CE_{KPCA}}{CE_{PCA}} \approx 23.9 \quad (27)$$

, respectively. Although the computational expense of KPCA is about 24 times of PCA, computational expense of GRAPPA using reduced channels is much larger than KPCA. Therefore, both PCA and KPCA only take a very tiny portion of the whole computational expense and they have the almost same computational cost in the whole reconstruction process.

5.4 Experimental Results

5.4.1 Experiment Settings

The performance of the proposed method was validated using three scanned datasets. The first scanned dataset was acquired on a GE 3T scanner (GE Healthcare, Waukesha, WI) with a 32-channel coil, and the last two datasets were acquired on a Siemens 3T scanner with an 32-channel coil (Siemens Trio, Erlangen, Germany). In the first dataset, a uniform water phantom was scanned using a gradient echo (GRE) sequence ($TE/TR = 10/100$ ms, 15.63 kHz bandwidth, matrix size = 256×256 , $FOV = 250 \text{ mm}^2$, Slice Thickness=3.0mm). The other second datasets were an axial brain image and coronary brain image acquired using a 2D spin echo (SE) sequence ($TE/TR = 2.29/100$ ms, matrix size = 256×256 , $FOV = 240 \text{ mm}^2$, Slice Thickness=3.0mm). Informed consents were obtained for all in vivo experiments in accordance with the institutional review board policy.

The proposed method was compared with the conventional GRAPPA [13] and PCA-based GRAPPA [43, 44]. The root sum of squares (SOS) reconstruction from the fully sampled data of all channels was shown as the reference image for comparison. The size of the coefficients (blocks by columns) was chosen as 2×15 for each individual method by comparing the mean-squared errors resulting from different sizes. Difference images from the reference were used to show all sources of error, including blurring, aliasing and noise. All methods were implemented in MATLAB (Mathworks, Natick, MA). To facilitate visual comparison, zoomed-in patches were also shown for some reconstructions.

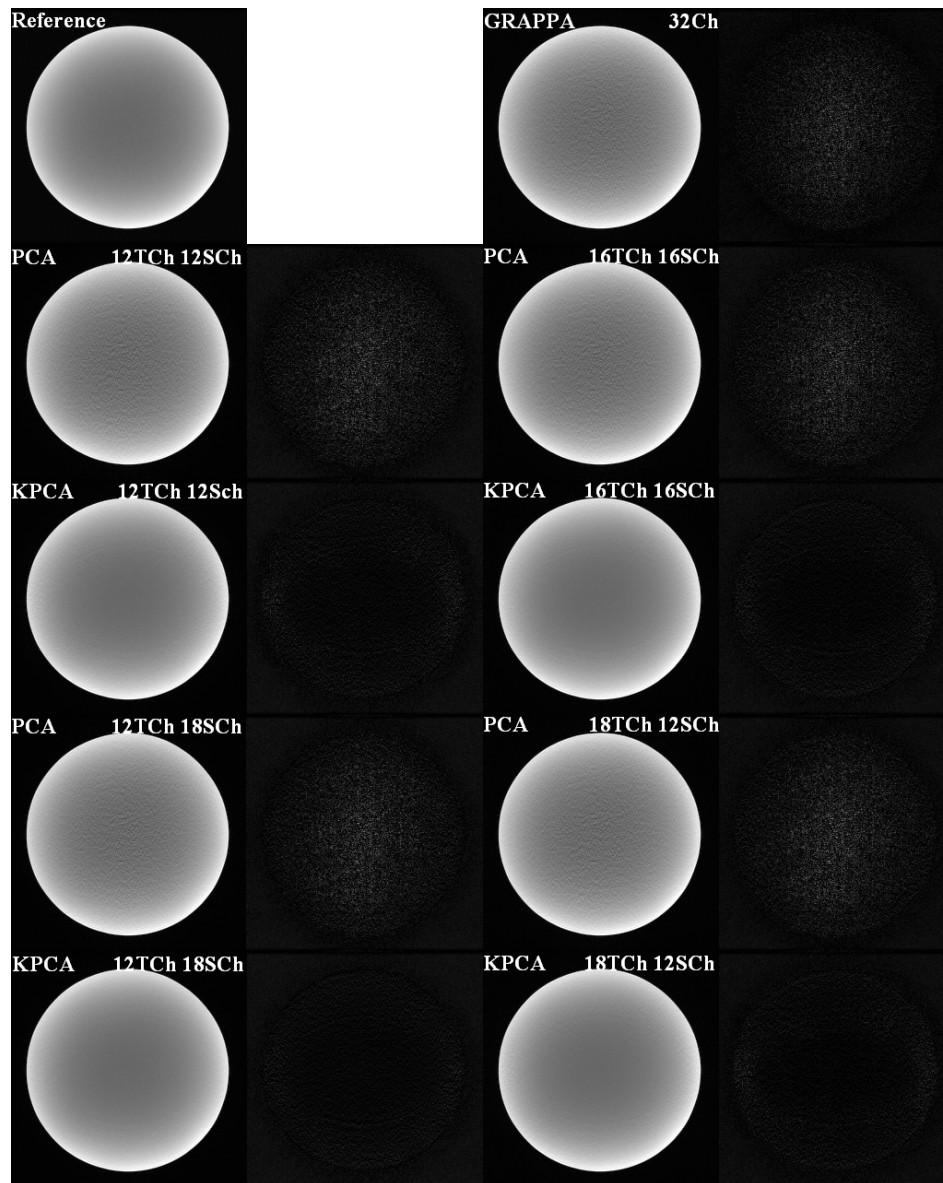
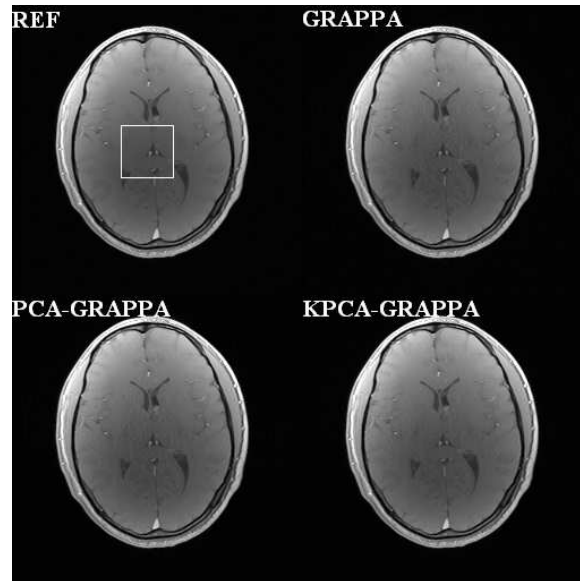


Figure 26. For the 32-channel phantom dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively. Their difference maps are also given for performance comparison.

5.4.2 Results of Phantom and In Vivo Datasets

Figure 26 shows the reconstructions of the phantom using the conventional GRAPPA, GRAPPA using PCA-based channel reduction, and the proposed GRAPPA using KPCA-

based channel reduction for an ORF of 5 and the ACS of 42 (net acceleration of 3.01). The size of the coefficients was 2 blocks and 15 columns. Both target and source have the same number of virtual channels, which are set as 12 and 16 for performance comparison. Furthermore, they have two combinations of $N_{\text{tch}}=12$, $N_{\text{sch}}=18$ and $N_{\text{tch}}=18$, $N_{\text{sch}}=12$ for testing performance. Similar to the conventional GRAPPA, it is seen that GRAPPA with PCA-based channel reduction suffers from serious noise. The proposed KPCA-based channel reduction can suppresses most of the noise without additional artifacts or loss of resolution. In addition, the proposed method has almost the same reconstruction time in comparison with PCA-based channel reduction, which is much faster than the conventional GRAPPA reconstruction. More details of the CPU time can see the Table 2 (a). It can be seen that the proposed method not only can reduce the reconstruction time to as low as around 6% of the conventional GRAPPA, but also it can achieve better reconstruction quality in comparison with the conventional GRAPPA and PCA-based GRAPPA.



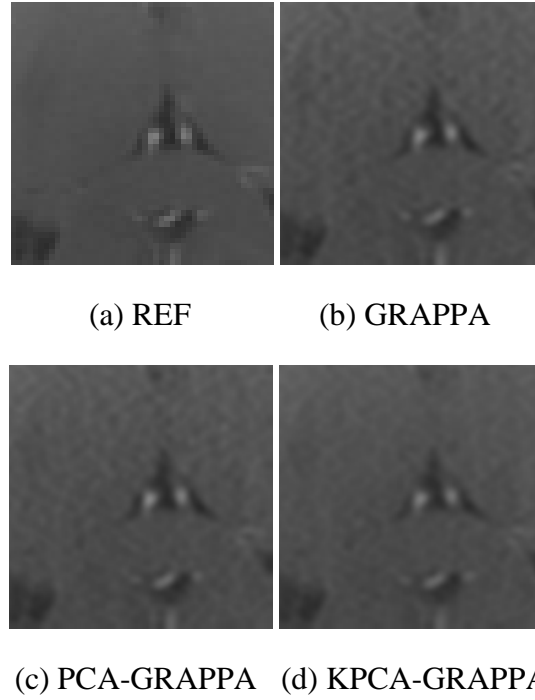


Figure 27. For the 32-channel brain (axial) dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively. Their difference maps are also given for performance comparison.

Figure 27 shows the reconstruction results for a vivo dataset of axial brain. An ORF of 4 and the ACS of 48 were used with a net acceleration of 2.56. The numbers of target and sources channels are both 16, and λ_{source} is $1.23\text{e-}9$. It is seen that the reconstruction using the proposed method is able to suppress the spatially-varying noise in both conventional GRAPPA and PCA-reduced GRAPPA reconstructions. Furthermore, the proposed method also preserves the details of the image without blurring. The proposed method takes almost the same time (863 seconds) as the PCA-based channel reduction method, and it takes only about 11% reconstruction time of the conventional GRAPPA (7771 seconds) to reconstruct the image but with better quality.

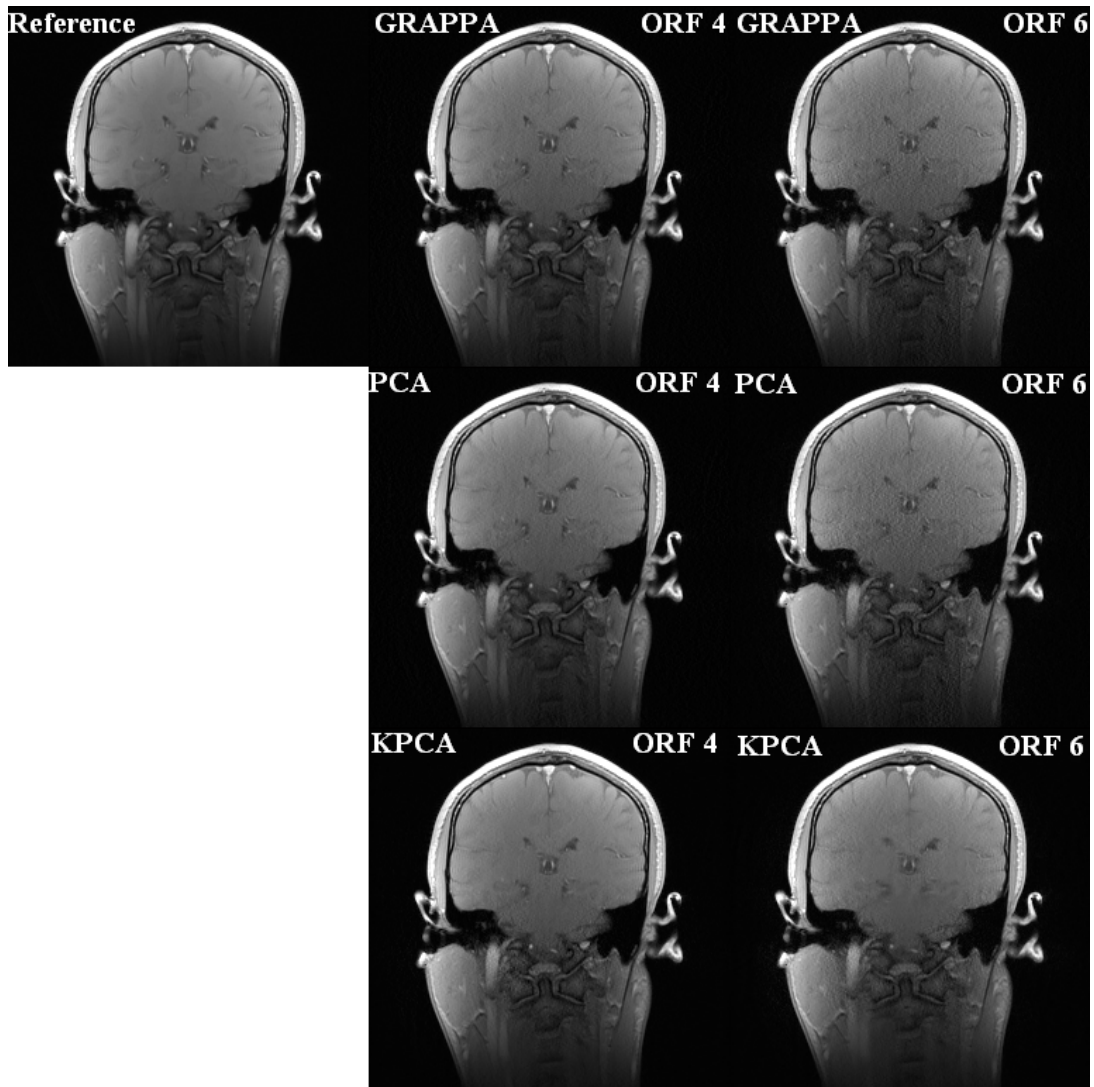


Figure 28. For the 32-channel brain (coronary) dataset, reference and reconstructions by the conventional GRAPPA, PCA-GRAPPA, KPCA-GRAPPA are presented, respectively.

Figure 28 shows the reconstruction results for a vivo dataset of coronary brain. Reconstruction parameters are 48 ACS lines, ORF 4 and 6, Coefficient size 2×15 , original channels 32, number of reduced channels 20, and net reduction factor 2.56 and 3.01. As we can see, the proposed method still outperforms both conventional GRAPPA using full channels and PCA-GRAPPA method. The CPU time is presented in Table 2 (b) for comparison.

Furthermore, NMSE and reconstruction time have been presented in Figure 29. The left sub-figure in Figure 29 shows the normalized mean-squared error (NMSE) for the conventional GRAPPA, PCA-GRAPPA, and KPCA-GRAPPA at ORF 4, 5, and 6, respectively. The right sub-figure in Figure 29 presents the corresponding reconstruction times in the left figure. It can be seen that reconstructions by the conventional GRAPPA have the longest reconstruction times and the worse reconstruction quality. Following the increasing ORF, the proposed method can still preserve lowest NMSE, but reconstructions of GRAPPA and PCA-GRAPPA are gradually deteriorated. Furthermore, the proposed method can reduce reconstruction time to a large extent. Although PCA-GRAPPA can also accelerate reconstruction time as well as the proposed method, its reconstruction quality is the worst among the all methods. The proposed method not only reduces the reconstruction time to around 15% of the conventional GRAPPA, but also keeps lowest NMSE in comparison with the conventional GRAPPA and PCA-GRAPPA.

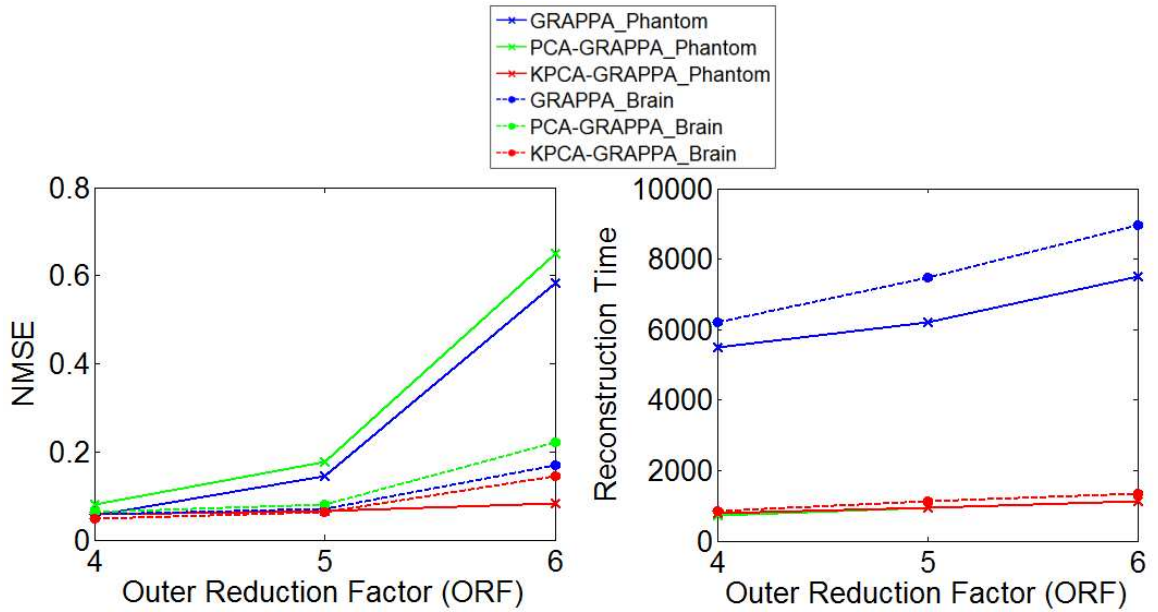


Figure 29. Quantitative comparison among GRAPPA, PCA-GRAPPA, and KPCA-GRAPPA at ourter reduction factors 4, 5, and 6.

(a)

ORF=5	N=32	$N_{tch}=N_{sch}=12$	$N_{tch}=N_{sch}=16$	$N_{tch}=12, N_{sch}=18$	$N_{tch}=18, N_{sch}=12$
GRAPPA	6197 s				
PCA		400 s	930 s	601 s	785 s
KPCA		397 s	939 s	600 s	782 s

(b)

	ORF = 4	ORF = 6
GRAPPA	5735 s	9587 s
PCA-GRAPPA	1448 s	2269 s
KPCA-GRAPPA	1452 s	2273 s

Table 2. (a) CPU times for 32-channel phantom dataset using different combinations of source and target channels, (b) reconstruction times for 32-channel brain (coronary) dataset.

Furthermore, computational costs of PCA and KPCA are extracted independently for evaluating their proportions in the whole reconstruction procedure. For 32-channel

datasets and 48 ACS lines, if the number of the reduced set of target and source channels is 8, 12, 16, 20, 24 respectively, the computational costs via CPU time of PCA and KPCA are presented as shown in Table 3. Although KPCA consumes more computation costs than PCA, it only takes a very small portion of the CPU times in the whole reconstruction procedure that are presented in Table 2. For this reason, whole reconstruction times of PCA-GRAPPA and KPCA-GRAPPA are almost the same.

	PCA	KPCA
$N_{tch}=N_{sch}=8$	0.30 s	3.93 s
$N_{tch}=N_{sch}=12$	0.33 s	4.23 s
$N_{tch}=N_{sch}=16$	0.37 s	4.34 s
$N_{tch}=N_{sch}=20$	0.39 s	4.52 s
$N_{tch}=N_{sch}=24$	0.43 s	4.75 s

Table 3. The computational costs via CPU time of PCA and KPCA at reduced 8, 12, 16, 20, and 24 channels.

5.5 Discussion

The parameter λ_{source} have different selections to control the reconstruction quality. The sensitivity map of one channel is generated from the division of the individual low-resolution image and the SOS of all channels. For 32 channels, we select 8 channels: 1, 5, 9, 13, 17, 21, 25, and 29 to demonstrate the changes of sensitivity maps, in which each

sensitivity maps is displayed using the same scale (bright means large magnitude value and dark denotes low magnitude value). As shown in Figure 30, sensitivity maps of original channels (without channel reduction) display normal magnitudes and each channel has sensitivity distribution with almost the same magnitudes. On the other hand, many channels corresponding to small eigenvalues after PCA reduction present very little magnitudes of sensitivity maps that contribute little to the SNR of the final reconstructed image. Generally, the first channel that corresponds to the largest eigenvalue has the largest magnitudes on the sensitivity map. Following the decreasing eigenvalues, magnitudes of the sensitivity map become smaller and make little contribution to SNR of the final reconstructed image. For KPCA, if λ_{source} is too small, sensitivity distribution for each channel after channel reduction is almost the same to that of PCA channel reduction. Following the increasing value of λ_{source} , sensitivity maps are gradually changed in comparison with PCA channel reduction, because new coordinates have been changed nonlinearly.

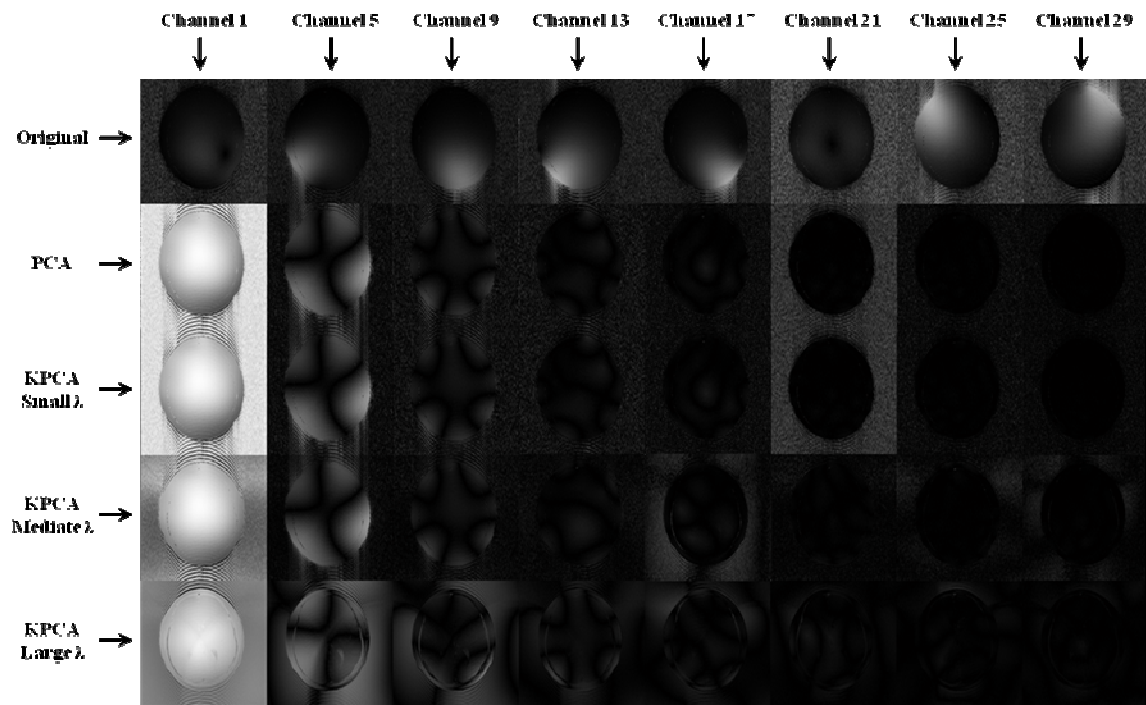


Figure 30. Comparison of sensitivity maps among original channels, reduced channels by PCA, reduced channels by KPCA with different parameters.

6. Conclusion

In summary, the problems of GRAPPA reconstruction are low SNR at high acceleration factor, low imaging acceleration, and significant computational costs with many coils. Since nonlinear approaches have been proposed and successfully applied on SENSE reconstruction in our group, in order to improve GRAPPA reconstruction efficiently, three nonlinear approaches have been proposed for the contribution in the dissertation.

In order to enhance SNR of GRAPPA reconstruction at high acceleration factors, we proposed a novel kernel-based nonlinear reconstruction algorithm for GRAPPA. The proposed method [66, 69] provides a more general model to characterize the noise behavior in GRAPPA reconstruction and thereby improves the SNR significantly. Experimental results demonstrate that the nonlinear GRAPPA is superior to conventional GRAPPA and other two methods at the same net acceleration factor. We anticipate that the proposed nonlinear approach can bring further benefits to current applications of conventional GRAPPA.

Compressed sensing (CS) and parallel MRI (pMRI) have been widely studied for accelerating MRI reconstruction. However, both reconstructed signals by CS and PI are not as accurate as acquired MR signals, so that errors of CS reconstruction as the first step will be propagated and even amplified to deteriorate pMRI reconstruction quality in the second step. We proposed a novel method [70] that combines CS with nonlinear GRAPPA sequentially for fast imaging. NLGRAPPA is applied on the k -space data reconstructed by CS. The generalized kernel regression model of NLGRAPPA can remove error effects of inaccurate signal reconstruction by CS. The experimental results

demonstrate that the proposed CS-NLGRAPPA method is superior to CS-GRAPPA in suppressing both noise and artifacts when a high net reduction factor is used.

Furthermore, in order to reduce the computational costs caused by many coils in reconstruction, we studied channel reduction technique for fast reconstruction. We proposed a novel nonlinear channel reduction for GRAPPA reconstruction. A novel kernel-PCA-based channel reduction method is proposed for parallel MRI. The method maps the data onto a higher dimensional space through a nonlinear transformation and PCA is performed to generate a reduced set of new channels. Experimental results demonstrate that the proposed KPCA-based channel reduction method is able to not only reduce the computational time, as the PCA-based method does, but also suppress noise in GRAPPA reconstruction.

7. Reference

- [1] Z.P. Liang, and P.C. Lauterbur, “Principles of magnetic resonance imaging: a signal processing perspective”, Wiley-IEEE press, 1999.
- [2] D.B. Twieg, “The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods”, *Medical Physics*, vol. 10, pp. 610-621, 1983.
- [3] S. Ljunggren, “A simple graphical representation of Fourier-based imaging methods”, *Journal of Magnetic Resonance*, vol. 54, pp. 338-343, 1983.
- [4] P.B. Roemer, W.A. Edelstein, C.E. Hayes, S.P. Souza, and O.M. Mueller, “The NMR phased array”, *Magnetic Resonance in Medicine*, vol. 16, pp. 192-225, 1990.
- [5] L. Ying, and Z.P. Liang, “Parallel MRI using phased array coils – multichannel sampling theory meets spin physics”, *IEEE Signal Processing Magazine*, vol. 27, pp. 90-98, 2010.
- [6] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, and P. Boesiger, “SENSE: Sensitivity encoding for fast MRI”, *Magnetic Resonance in Medicine*, vol. 42, pp. 952-962, 1999.
- [7] M.A. Griswold, P.M. Jakob, M. Nittka, J.W. Goldfarb, and A. Haase, “Partially parallel imaging with localized sensitivities (PILS)”, *Magnetic Resonance in Medicine*, vol. 44, pp. 602-609, 2000.
- [8] D.K. Sodickson, and W.J. Manning, “Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays”, *Magnetic Resonance in Medicine*, vol. 38, pp. 591-603, 1997.

- [9] 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)", *Magnetic Resonance in Medicine*, vol. 47, pp. 1202-1210, 2002.
- [10] P.M. Jakob, M.A. Griswold, R.R. Edelman, and D.K. Sodickson, "AutoSMASH: a self-calibrating technique for SMASH imaging", *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 7, pp. 42-54, 1998.
- [11] R.M. Heidemann, M.A. Griswold, A. Haase, and P.M. Jakob, "VD-AutoSMASH imaging", *Magnetic Resonance in Medicine*, vol. 45, pp. 1066-1074, 2001.
- [12] J. Park, Q. Zhang, V. Jellus, O. Simonetti, and D. Li, "Artifact and noise suppression in GRAPPA imaging using improved k -space coil calibration and variable density sampling", *Magnetic Resonance in Medicine*, vol. 53, pp. 186-193.
- [13] Z. Wang, J. Wang, and J.A. Detre, "Improved data reconstruction method for GRAPPA", *Magnetic Resonance in Medicine*, vol. 54, pp. 738-742, 2005.
- [14] P. Qu, C. Wang, and G.X. Shen, "Discrepancy-based adaptive regularization for GRAPPA reconstruction", *Journal of Magnetic Resonance Imaging*, vol. 24, pp. 248-255, 2006.
- [15] F.H. Lin, "Prior-Regularized GRAPPA Reconstruction", *Proceedings of the 14th Annual Meeting of ISMRM*, Seattle, 3656, 2006.
- [16] D. Huo, and D.L. Wilson, "Robust GRAPPA reconstruction and its evaluation with the perceptual difference model", *Journal of Magnetic Resonance Imaging*, vol. 27, pp. 1412-1420, 2008.

- [17] F. Huang, Y. Li, S. Vijayakumar, S. Hertel, and G.R. Duensing, "High-pass GRAPPA: an image support reduction technique for improved partially parallel imaging", *Magnetic Resonance in Medicine*, vol. 59, pp. 642-649, 2008.
- [18] R. Nana, T. Zhao, K. Heberlein, S.M. LaConte, and X. Hu, "Cross-validation-based kernel support selection for improved GRAPPA reconstruction", *Magnetic Resonance in Medicine*, vol. 59, pp. 519-825, 2008.
- [19] R. Nana, and Hu X. Data consistency criterion for selecting parameters for k-space-based reconstruction in parallel imaging. *Magnetic Resonance Imaging*, vol. 28, pp. 119-128, 2010.
- [20] T. Zhao, and X. Hu, "Iterative GRAPPA (iGRAPPA) for improved parallel imaging reconstruction", *Magnetic Resonance in Medicine*, vol. 59, pp. 903-907, 2008.
- [21] M.A. Griswold, M. Blaimer, F. Breuer, R.M. Heidemann, M. Mueller, and P.M. Jakob, "Parallel magnetic resonance imaging using the GRAPPA operator formalism", *Magnetic Resonance in Medicine*, vol. 54, pp. 1553-1556, 2005.
- [22] M. Bydder, and Y. Jung, "A nonlinear regularization strategy for GRAPPA calibration", *Magnetic Resonance Imaging*, vol. 27, pp. 137-141, 2009.
- [23] M. Blaimer, M. Gutterlet, P. Kellman, F.A. Breuer, H. Kostler, and M.A. Griswold, "Virtual coil concept for improved parallel MRI employing conjugate symmetric signals", *Magnetic Resonance in Medicine*, vol. 61, pp. 93-102, 2009.
- [24] M. Honal, S. Bauer, U. Ludwig, and J. Leupold, "Increasing efficiency of parallel imaging for 2D multislice acquisitions", *Magnetic Resonance in Medicine*, vol. 61, pp. 1459-1470, 2009.

- [25] Z. Chen, J. Zhang, R. Yang, P. Kellman, L.A. Johnston, and G.F. Egan, "IIR GRAPPA for Parallel MR Image Reconstruction", *Magnetic Resonance in Medicine*, vol. 63, pp. 502-509, 2010.
- [26] H. Wang, D. Liang, K.F. King, G. Nagarsekar, Y. Chang, and L. Ying, "Improving GRAPPA using cross-sampled autocalibration data", *Magnetic Resonance in Medicine*, vol. 67, pp. 1042-1053, 2011.
- [27] L. Ying, and E. Abdelsalam, "Parallel MRI reconstruction: a filter-bank approach", *Proceedings of the IEEE Engineering in Medicine and Biology Conference (EMBC)*, pp. 1374-1377, 2005.
- [28] B. Sharif, and Y. Bresler, "Distortion-optimal self-calibrating parallel MRI by blind interpolation in subsampled filter banks", *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 52-56, 2011.
- [29] L. Ying, and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)", *Magnetic Resonance in Medicine*, vol. 57, pp. 1196-1202, 2007.
- [30] B. Liu, F.M. Sebert, Y. Zou, and L. Ying, "SparseSENSE: randomly-sampled parallel imaging using compressed sensing", *Proceedings of the 16th Annual Meeting of ISMRM, Toronto*, 3154, 2008.
- [31] B. Liu, K.F. King, M. Steckner, J. Xie, J. Sheng, and L. Ying, "Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations", *Magnetic Resonance in Medicine*, vol. 61, pp. 145-152, 2009.
- [32] D. Liang, B. Liu, J. Wang, and L. Ying, "Accelerating SENSE using compressed sensing", *Magnetic Resonance in Medicine*, vol. 62, pp. 1574-1584, 2009.

- [33] P. Tomaso, and S. Steve, “The mathematics of learning: dealing with data”, Notices of the American Mathematical Society, vol. 50, pp. 537-544, 2003.
- [34] W. Liu, P.P. Pokharel, and J.C. Principe, “The kernel least-mean-square algorithm”, IEEE Transactions on Signal Processing, vol. 56, pp. 543-554, 2008.
- [35] G. Camps-Valls, J.L. Rojo-Alvarez, and M. Martinez-Ramon, “Kernel methods in bioengineering, signal and image processing”, Idea Group Publishing, London, 2007.
- [36] B. Scholkopf, and A.J. Smola, “Learning with kernels: support vector machines, regularization, optimization, and beyond (adaptive computation and machine learning)”, The MIT Press, Boston, 2001.
- [37] H. Takeda, S. Farsiu, and P. Milanfar, “Kernel regression for image processing and reconstruction”, IEEE Transactions on Image Processing, vol. 16, pp. 349-366, 2007.
- [38] I.T. Jolliffe, “Principal component analysis”, Springer Series in Statistics, Springer, 2nd Edition, 2010.
- [39] B. Scholkopf, A. Smola, and K.R. Muller, “Kernel principal component analysis”, International Conference on Artificial Neural Networks, LNCS, vol. 1327, pp. 583-588, 1997.
- [40] S. Mika, B. Scholkopf, A. Smola, K.R. Muller, M. Scholz, and G. Ratsch, “Kernel PCA and de-noising in feature space”, Neural Information Processing Systems Conference (NIPS), 1998.
- [41] P. Gruber, K. Stadlthanner, M. Bohm, F.J. Theis, and E.W. Lang, “Denoising using local projective subspace methods”, Neurocomputing, vol. 69, pp. 1485-1501, 2006.

- [42] M. Berar, M. Desvignes, G. Bailly, Y. Payan, and B. Romaniuk, "Missing data estimation using polynomial kernels", International Conference on Advances in Pattern Recognition, LNCS, vol. 3686, pp. 390-399, 2005.
- [43] F. Huang, S. Vijayakumar, Y. Li, S. Hertel, and G. Duensing, "A software channel compression technique for faster reconstruction with many channels", Magnetic Resonance Imaging, vol. 26, pp. 133-141, 2008.
- [44] F. Huang, W. Lin, G.R. Duensing, and A. Reykowski, "A hybrid method for more efficient channel-by-channel reconstruction with many channels", Magnetic Resonance in Medicine, Early View Online, 2011.
- [45] M. Lustig, D. Donoho, and J.M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," Magnetic Resonance in Medicine, vol. 58, pp. 1182-1195, 2007.
- [46] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly, "Compressed sensing MRI", IEEE Signal Processing Magazine, vol. 25, pp. 72-82, 2008.
- [47] K.F. Kevin, D. Xu, A.C. Brau, P. Lai, P.J. Beatty, and L. Marinelli, "A new combination of compressed sensing and data driven parallel imaging", Proceedings of the 18th Annual Meeting of ISMRM, Toronto, 4881, 2010.
- [48] M. Lustig, and J.M. Pauly, "SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space," Magnetic Resonance in Medicine, vol. 64, pp. 457-471, 2010.
- [49] F. Huang, and G.R. Duensing, "A theoretical analysis of errors in GRAPPA", Proceedings of the 14th Annual Meeting of ISMRM, Seattle, 2468, 2006.

- [50] M. Schmitt, A. Potthast, D.E. Sosnovik, J.R. Polimeni, G.C. Wiggins, C. Triantafyllou, and L.L. Wald, “A 128-channel receive-only cardiac coil for highly accelerated cardiac MRI at 3 Tesla”, *Magnetic Resonance in Medicine*, vol. 59, pp. 1431-1439, 2008.
- [51] M. Buehrer, K.P. Pruessmann, P. Boesiger, and S. Kozerke, “Array compression for MRI with large coil arrays”, *Magnetic Resonance in Medicine*, vol. 57, pp. 1131-1139, 2007.
- [52] M. Doneva, and P. Bornert P, “Automatic coil selection for channel reduction in SENES-based parallel imaging”, *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 21, pp. 187-196, 2008.
- [53] S. Feng, Y. Zhu, J. Ji, “Efficient large-array k-domain parallel MRI using channel-by-channel array reduction”, *Magnetic Resonance Imaging*, vol. 29, pp. 209-215, 2011.
- [54] S. King, S. Varosi, G. Duensing, “Optimum SNR data compression in hardware using an eigencoil array”, *Magnetic Resonance in Medicine*, vol. 63, pp. 1346-1356, 2010.
- [55] M.A. Carreira-Perpiñán, “Continuous latent variable models for dimensionality reduction and sequential data reconstruction”, PhD thesis, University of Sheffield, UK, 2001.
- [56] D. Watkins, “Fundamentals of matrix computations”, 2nd edition. Wiley-Interscience, New York, USA, 2002.
- [57] G. Golub, “Some modified matrix eigenvalue problems”, *Society for Industrial and Applied Mathematics (SIAM) Review*, vol. 15, pp. 318-344, 1973.

- [58] G. Golub, and C.V. Loan, “An analysis of the total least squares problem”, Society for Industrial and Applied Mathematics (SIAM) Journal on Numerical Analysis, vol. 17, pp. 883-893, 1980.
- [59] S. Aja-Fernández, A. Tristán-Vega, W.S. Hoge, “Statistical noise analysis in GRAPPA using a parameterized noncentral Chi approximation model”, Magnetic Resonance in Medicine, vol. 65, pp. 1195-1206, 2011.
- [60] T. FrieB, and R. Harrison, “A kernel-based adaline”, Proceedings of the 7th European Symposium on Artificial Neural Networks, Bruges, Belgium, pp. 245-250, 1999.
- [61] K. Kim, M.O. Franz, and B. Scholkopf, “Iterative kernel principal component analysis for image modeling”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp. 1351-1366, 2005.
- [62] Y.W. Chang, C.J. Hsieh, K.W. Chang, M. Ringgaard, and C.J. Lin, “Training and testing low-degree polynomial data mappings via linear SVM”, Journal of Machine Learning Research, vol. 11, pp. 1471-1490, 2010.
- [63] T. Zhang T, “Learning bounds for kernel regression using effective data dimensionality”, Neural Computation, vol. 17, pp. 2077-2098, 2005.
- [64] P.M. Robson, A.K. Grant, A.J. Madhuranthakam, R. Lattanzi, D.K. Sodickson, C.A. McKenzie, “Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions”, Magnetic Resonance in Medicine, vol. 60, pp. 895-907, 2008.

- [65] S. Bauer, M. Markl, M. Honal, and B.A. Jung, “The effect of reconstruction and acquisition parameters for GRAPPA-based parallel imaging on the image quality”, *Magnetic Resonance in Medicine*, vol. 66, pp. 402-409, 2011.
- [66] Y. Chang, D. Liang, and L. Ying, “Nonlinear GRAPPA: a kernel approach to parallel MRI reconstruction”, *Magnetic Resonance in Medicine*, vol. 68, pp. 730-740, September, 2012.
- [67] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, pp. 489-509, 2006.
- [68] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*., vol. 52, pp. 1289 - 1306, 2006.
- [69] Y. Chang, D. Liang, and L. Ying, “A kernel approach to parallel MRI reconstruction”, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 389-392, 2011.
- [70] Y. Chang, K. King, D. Liang, Y. Wang, and L. Ying, “A kernel approach to Compressed Sensing parallel MRI”, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 78-81, 2012.
- [71] A.N. Tikhonov, V.Y. Arsenin, “Solutions of ill-posed problems”, New York: Winston & Sons, 1977.
- [72] J.R. Porter, S.M. Wright, A. Reykowski, “A 16-element phased array head coil”, *Magnetic Resonance in Medicine*, vol. 40, pp. 272-279, 1998.

- [73] Y. Zhu, C. Hardy, D. Sodickson, R. Giaquinto, C. Dumoulin, G. Kenwood, T. Niendorf, H. Lejay, C. McKenzie, M. Ohliger, and N. Rofsky, "Highly parallel volumetric imaging with a 32-element RF coil array", *Magnetic Resonance in Medicine*, vol. 52, pp. 869-877, 2004.
- [74] M. McDougall, and S. Wright, "64-channel array coil for single echo acquisition magnetic resonance imaging", *Magnetic Resonance in Medicine*, vol. 54, pp. 386-392, 2005.
- [75] S. King, G. Duensing, "The MRI Eigencoil: 2N-channel SNR with N-receivers", *Proceedings of the 11th Annual Meeting of ISMRM*, Toronto, 712, 2003.
- [76] P. Beatty, W. Sun, A. Brau, "Direct virtual coil (DVC) reconstruction for data-driven parallel imaging", *Proceedings of the 16th Annual Meeting of ISMRM*, Toronto, 8, 2008.
- [77] P. Beatty, J. Holmes, S. Chang, E. Bayram, J. Brittain, S. Reeder, A. Brau, "Coil-by-coil vs. direct virtual coil (DVC) parallel imaging reconstruction: an image quality comparison for contrast-enhanced liver imaging", *Proceedings of the 18th Annual Meeting of ISMRM*, Stockholm, 2879, 2008.
- [78] W. Chen, P. Hu, C. Meyer, "Rapid partially parallel reconstruction using single synthetic target coil", *Proceedings of the 16th Annual Meeting of ISMRM*, Toronto, 1296, 2008.
- [79] S. Muller, R. Umathum, P. Speier, S. Zuhlsdorff, S. Ley, W. Semmler, M. Bock, "Dynamic coil selection for real-time imaging in interventional MRI", *Magnetic Resonance in Medicine*, vol. 56, pp. 1156-1162, 2006.

- [80] L. Marinelli, C.J. Hardy, “Coil selection optimization using mean-field annealing and its application to 128-channel imaging”, Proceedings of the 15th Annual Meeting of ISMRM, Berlin, 750, 2007.
- [81] F. Huang, S. Vijayakumar, J. Akao, “Software compression for partially parallel imaging with multi-channels”, Proceedings of the IEEE Engineering in Medicine and Biology Society, Shanghai, pp. 1348-1351, 2005.
- [82] T. Zhang, J. Pauly, S. Vasanawala, M. Lustig, “Coil compression for accelerated imaging with Cartesian sampling”, Magnetic Resonance in Medicine, Early View Online, 2012.
- [83] S. Vijayakumar, F. Huang, J.H. Akao, M.K. Limkeman, G.R. Duensing, “Channel compression and denoising”, Proceedings of the 15th Annual Meeting of ISMRM, Berlin, 1908, 2007.
- [84] A. Stemmer, V. Jellus, S. Kannengiesser, B. Kiefer, “Channel compression for BLADE”, Proceedings of the 16th Annual Meeting of ISMRM, Toronto, 1274, 2008.
- [85] A. Brau, P.J. Beatty, S. Skare, R. Bammer, “Comparison of reconstruction accuracy and efficiency among autocalibrating data-driven parallel imaging methods”, Magnetic Resonance in Medicine, vol. 59, pp. 382–395, 2008.
- [86] M. Arakawa, “Computational workloads for commonly used signal processing kernels”, Lincoln Laboratory Project Report SPR-9, MIT, May 28 2003, Reissued Nov. 30 2006.

8. Curriculum Vitae

Education

Ph.D., Electrical Engineering, University of Wisconsin - Milwaukee, USA, Expect 12/2012

Concentrations: Magnetic Resonance Imaging, Biomedical Imaging, Signal Processing

Dissertation: A Study of Nonlinear Approaches to Parallel Magnetic Resonance Imaging

M.S., Pattern Recognition & Intelligent Systems, Shanghai Jiao Tong Univ., China, 2006

Concentrations: Pattern Recognition, Image Processing, Computer Vision, Machine Learning

Thesis: Color Image Segmentation with Application on Content-Based Image Retrieval

B.S., Automatic Control, Northwestern Polytechnical University, China, 2003

Graduation Design: Unix/Linux Shell Programming

Experience

C++ Software Engineer, 06/2012 – 08/2012

The Neat Company, Philadelphia, Pennsylvania, USA

Research Assistant, 01/2009 – 12/2011

University of Wisconsin – Milwaukee, Milwaukee, Wisconsin, USA

Teaching Assistant, 09/2006 – 12/2006 & 09/2008 – 12/2008

Brigham Young University, Provo, Utah, USA

Research Assistant, 01/2007 – 08/2008

Brigham Young University, Provo, Utah, USA

Intern, 04/2006 – 07/2006

Microsoft Research Asia, Beijing, China

Intern, 10/2005 – 03/2006

Hitachi (China) Research & Development Corp., Shanghai, China

Intern, 05/2005 – 08/2005

Centaline China Real Estate Corp., Shanghai, China

Publication

- [1] Yuchou Chang, Dong Liang, and Leslie Ying, “Nonlinear GRAPPA: A Kernel Approach to Parallel MRI Reconstruction”, *Magnetic Resonance in Medicine*, vol. 68, pp. 730-740, September, 2012.
- [2] Haifeng Wang, Dong Liang, Kevin F. King, Gajanan Nagarsekar, Yuchou Chang, and Leslie Ying, “Improving GRAPPA Using Cross-Sampled Autocalibration Data”, *Magnetic Resonance in Medicine*, vol. 67, pp. 1042-1053, April, 2012.
- [3] Dong Liang, Haifeng Wang, Yuchou Chang, and Leslie Ying, “Sensitivity Encoding Reconstruction with Nonlocal Total Variation Regularization”, *Magnetic Resonance in Medicine*, vol. 65, pp. 1384-1392, May, 2011.
- [4] Yuchou Chang, Kevin F. King, Dong Liang, Yong Wang, and Leslie Ying, “A Kernel Approach to Compressed Sensing Parallel MRI”, *IEEE Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, Barcelona, Spain, May 2 – 5, 2012.
- [5] Haifeng Wang, Yihang Zhou, Yuchou Chang, Yong Wang, “Smoothed Random-Like Trajectory for Compressed Sensing MRI”, *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, San Diego, California, Aug. 28 – Sep. 1, 2012.
- [6] Yong Wang, Dong Liang, Yuchou Chang, and Leslie Ying, “A Hybrid Total-Variation Minimization Approach to Compressed Sensing”, *IEEE Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, Barcelona, Spain, May 2 – 5, 2012.
- [7] Yihang Zhou, Yuchou Chang, Dong Liang, and Leslie Ying, “K-T CSPI: A Dynamic MRI Reconstruction Framework for Combining Compressed Sensing and Parallel Imaging”, *IEEE Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, Barcelona, Spain, May 2 – 5, 2012.
- [8] Yuchou Chang, Kevin F. King, Dong Liang, and Leslie Ying, “Combining Compressed Sensing and Nonlinear GRAPPA for Highly Accelerated Parallel MRI”, *20th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM)*, Melbourne, Australia, May 5 – 11, 2012.

- [9] Yuchou Chang, and Leslie Ying, “A Nonlinear ARMA Model for GRAPPA Reconstruction”, 20th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM), Melbourne, Australia, May 5 – 11, 2012.
- [10] Yihang Zhou, Yuchou Chang, and Leslie Ying, “k-t CS-NLG: Dynamic Imaging Reconstruction with Compressed Sensing and Nonlinear GRAPPA”, 20th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM), Melbourne, Australia, May 5 – 11, 2012.
- [11] Haifeng Wang, Yuchou Chang, Dong Liang, Kevin F. King, and Leslie Ying, “Cross Sampled Nonlinear GRAPPA for Parallel MRI”, 20th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM), Melbourne, Australia, May 5 – 11, 2012.
- [12] Yuchou Chang, Dong Liang, and Leslie Ying, “Kernel GRAPPA: A General Nonlinear Framework for GRAPPA Regularization”, 19th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM), Quebec, Canada, 4394, May 7-13 2011.
- [13] Yuchou Chang, Dong Liang, and Leslie Ying, “A Kernel Approach to Parallel MRI Reconstruction”, IEEE Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Chicago, Illinois, pp. 389-392, Mar. 30 – Apr. 2 2011.
- [14] Yuchou Chang, Dong Liang, and Leslie Ying, “Nonlinear GRAPPA Using Volterra Filter”, 2nd Madison-Freiburg Workshop on Accelerated Medical Imaging, Madison, Wisconsin, Jun. 17-18 2010.
- [15] Yuchou Chang, Dong Liang, and Leslie Ying, “A Nonlinear GRAPPA Method for Improving SNR”, 18th Annual Meeting in International Society for Magnetic Resonance in Medicine (ISMRM), Stockholm, Sweden, May 1-7 2010.
- [16] Huajun She, Rong-Rong Chen, Dong Liang, Yuchou Chang, and Leslie Ying, “Image Reconstruction from Phased-Array MRI Data Based on Multichannel Blind Deconvolution”, in IEEE Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Rotterdam, Netherlands, Apr. 14-17 2010.