The Empirical Effect of Gaussian Noise in Undersampled MRI Reconstruction

Patrick Virtue and Michael Lustig

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA

Corresponding Author:
Patrick Virtue, BS
Electrical Engineering and Computer Sciences, University of California,
Berkeley, 253 Cory Hall, Berkeley, CA 94720;
E-mail: virtue@eecs.berkeley.edu

In Fourier-based medical imaging, sampling below the Nyquist rate results in an underdetermined system, in which a linear reconstruction will exhibit artifacts. Another consequence is lower signal-to-noise ratio (SNR) because of fewer acquired measurements. Even if one could obtain information to perfectly disambiguate the underdetermined system, the reconstructed image could still have lower image quality than a corresponding fully sampled acquisition because of reduced measurement time. The coupled effects of low SNR and underdetermined system during reconstruction makes it difficult to isolate the impact of low SNR on image quality. To this end, we present an image quality prediction process that reconstructs fully sampled, fully determined data with noise added to simulate the SNR loss induced by a given undersampling pattern. The resulting prediction image empirically shows the effects of noise in undersampled image reconstruction without any effect from an underdetermined system. We discuss how our image quality prediction process simulates the distribution of noise for a given undersampling pattern, including variable density sampling that produces colored noise in the measurement data. An interesting consequence of our prediction model is that recovery from an underdetermined nonuniform sampling is equivalent to a weighted least squares optimization that accounts for heterogeneous noise levels across measurements. Through experiments with synthetic and in vivo datasets, we demonstrate the efficacy of the image quality prediction process and show that it provides a better estimation of reconstruction image quality than the corresponding fully sampled reference image.

INTRODUCTION

Undersampling in Fourier-based medical imaging provides a variety of clinical benefits including shorter exam times, reduced motion artifacts, and the ability to capture fast-moving dynamics, such as cardiac motion. Undersampling reduces acquisition time by collecting fewer measurements in the frequency domain than required by the Nyquist rate. However, undersampling causes two specific challenges for the reconstruction system, namely, an **underdetermined system**[1] of linear equations and **lower SNR** (signal-to-noise ratio) because of reduced measurement time. When reconstruction algorithms are able to overcome these challenges, undersampling can benefit a variety of Fourier-based imaging modalities, including magnetic resonance imaging (MRI) with parallel imaging or compressed sensing (1, 2), computed tomography (CT) with reduced

or gated acquisition views (3, 4), and positron emission tomography (PET) with multiplexed or missing detectors (5, 6). Undersampling for acceleration is becoming the mainstream approach for fast imaging. In fact, this year, two of the major MRI manufacturers have announced products which leverage undersampling and a compressed sensing reconstruction that have been approved by the Food and Drug Administration (FDA). While the tools and analysis discussed in this paper apply generally to Fourier-based medical imaging with Gaussian noise, we will direct our numerical modeling, examples, and experiments to the application of compressed sensing MRI.

When designing an undersampled reconstruction system, the primary concern is often focused on compensating for the underdetermined system caused by sub-Nyquist sampling, for example choosing a sparse representation for compressed sensing. However, we should not overlook the fact that collecting fewer measurements in practice leads to overall lower SNR in the acquired data. If the measurements are too noisy, the low SNR will lead to poor reconstructed image quality even if the reconstruction system were fully determined. On the other hand, with high SNR measurements, the resulting image quality will be

---

[1] In the context of this paper, we specify fully determined and underdetermined as follows: for a fixed Cartesian *k*-space (frequency space) grid with predefined field of view and spatial resolution parameters, fully determined means having at least one measured sample for each *k*-space grid location, and underdetermined means at least one *k*-space location has zero samples, in which case we have more unknowns (image pixels) than equations (one per acquired *k*-space location).

Figure 1. Prediction of image quality: The process to add the proper amount of noise to fully sampled reference k-space and reconstruct an image affected by lower SNR because of reduced acquisition time but not affected by an underdetermined system. The expected measurement time at each k-space location, τ_k, associated with the given sampling pattern is used to calculate the amount of noise (zero mean, complex Gaussian with variance $\sigma^2_{add,k}$) to add to each position in the reference k-space. This k-space with added noise is then processed by the reconstruction algorithm to produce the prediction image.
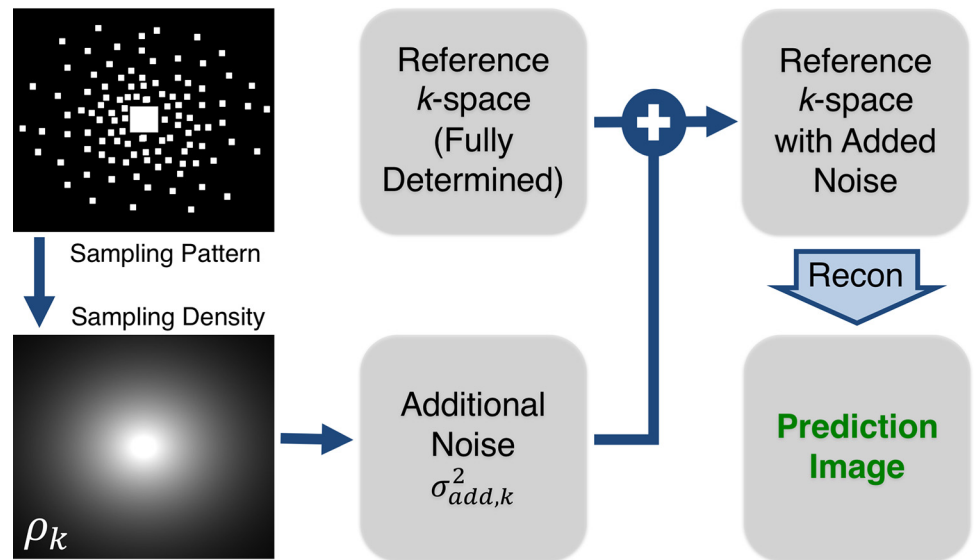
limited by how well the reconstruction can constrain the underdetermined system. The effects of the underdetermined system and the lower SNR are coupled during the reconstruction process, making it difficult to analyze one without the other. It is important, however, to analyze how both issues impact the reconstruction system to determine the empirical limits of undersampling and gain insight on how to improve undersampled acquisition and reconstruction when targeting specific applications.

Compressed sensing theory has provided us with extensive analysis on the bounds for the successful signal recovery from undersampled data. Candès (7) describes a bound on the squared error of the recovered signal limited by the undersampling rate and the sparsity of the data. He also shows that this bound scales linearly with the variance of the noise in the measured data. Candès and Plan (8) provide a more general compressed sensing theory that addresses a combination of practical concerns. For instance, they derive the bounds on squared error of the recovered signal for systems with Fourier encoding matrices, noise measurements, and approximately sparse signals. Unfortunately, while squared error is an important tool in measuring similarity between signals, it often fails to provide a good measure of perceptual image quality. Wainwright (9) improves upon the squared error definition of success by studying the undersampling rates and sparsity levels for which there is a high probability of successfully recovering the support of the sparse signal.

Although it is important to have a theory showing that reconstruction techniques are mathematically founded, when testing a reconstruction algorithm on a new undersampled clinical dataset shows unacceptable image results, it is difficult to leverage the theoretical bounds to understand the cause of the failure. Conversely, when an undersampled reconstruction is

successful at a certain undersampling rate, it is natural to then ask, how much further can we push undersampling? In this case, it is difficult to translate theoretic analyses, such as time constants for polylogarithmic bounds (8), into practice. Our goal in this paper is to provide the tools to empirically analyze the effects of lower SNR from reduced measurement time using a reconstruction system that is fully determined, rather than underdetermined. To this end, we present the **image quality prediction** process (Figure 1). The image quality prediction process takes a Nyquist-sampled (fully determined) reference dataset and adds the proper amount of noise to mimic the lower SNR produced by a given undersampling pattern. By reconstructing this noisy, but still Nyquist-sampled dataset, we have a **prediction image** that has been affected by lower SNR from reduced measurement time but not by artifacts from an underdetermined reconstruction. The image quality prediction process give us the following three benefits:

- Comparing the prediction image to the reference reconstruction allows us to see the impact of lower SNR from reduced measurement time on the reconstruction system.
- Comparing the prediction image to the underdetermined reconstruction, we are able to assess the added effect of the underdetermined system on the reconstructed image.
- The prediction image provides a better estimate of undersampled image quality than overoptimistically comparing an underdetermined reconstruction to a fully sampled reference reconstruction.

As exemplified in Figure 2, for a given clinical application and undersampling pattern, pulse sequence and reconstruction developers can use the image quality prediction process to determine if low SNR, rather than the underdetermined system, is the limiting factor for a successful reconstruction. Specifi-
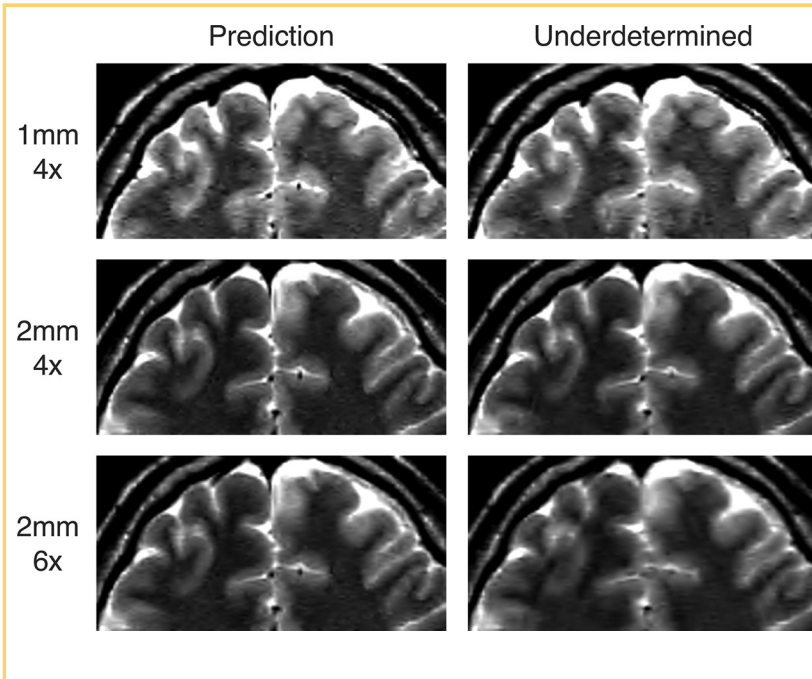
Figure 2. Using the image quality prediction process to adjust scan parameters. This 2D fast spin echo acquisition with 1 mm slice thickness and 4× undersampling produces poor reconstruction image quality (top right). The corresponding prediction image (top left) also has poor image quality, indicating that noise is the limiting factor. Increasing to 2 mm slice thickness (center row) reduces the noise and produces higher image quality in both the prediction and the underdetermined reconstruction. Further accelerating the scan with 6× undersampling (bottom row), the prediction image quality is significantly higher than the reconstruction image quality, indicating that the underdetermined system is the limiting factor for those scan parameters.

cally, an unsatisfactory prediction image indicates that the undersampled acquisition contains more noise than the reconstruction can handle. On the other hand, a high quality prediction image and poor results from the underdetermined reconstruction indicate that the constraints on the underdetermined system are not adequate for the limited number of samples acquired. Once developers understand the limiting factor in a given undersampling application, they can then recommend changes to the acquisition protocol to adjust the measurement SNR or the undersampling rate. Developers can also appropriately focus their efforts on improving the reconstruction algorithm to better account for the noise distribution or to improve the reconstruction constraints, such as the sparsity model.

Before describing the details of the image quality prediction process, we first specify how measurement time affects SNR, specifically when undersampling. We complete this section by introducing a weighted least squares optimization that generalizes the reconstruction process for both undersampled data and the fully determined prediction data.

Measurement Time and SNR

For MRI reconstruction, we can model the signal s as the discrete Fourier transform of the unknown target image object m:

$$s_k = (Fm)_k \tag{1}$$

where F is the multidimensional discrete Fourier transform operator and k is the k-th location in k-space. However, each measurement s_k comes with an associated noise. We can model the noisy measurement Y_k as:

$$Y_k \sim \mathcal{N}(Re(s_k), \sigma^2_{acq}/\tau_k) + i\mathcal{N}(Im(s_k), \sigma^2_{acq}/\tau_k) \tag{2}$$

where Y_k is a random variable drawn from a complex-valued Gaussian distribution with mean s_k and variance defined by the system noise variance, σ^2_{acq}, scaled by one over the measurement time, τ_k, as described in (10). With this definition, we assume that the signal is deterministic based on our model, the signal is

independent of the noise, and that the noise is independent and identically distributed. In cases where these assumptions do not hold, additional care may be taken to adjust the data to this model, for example, prewhitening coil channels in parallel imaging or accounting for echo time variation in fast spin echo acquisitions.

Again following (10), we define SNR as the signal intensity divided by the standard deviation of the noise and note that from equation (2) we see that the SNR for measured data at the k-th location in k-space scales with $1/\sqrt{\tau_k}$ is:

$$SNR = \frac{signal}{\sqrt{noise\ var.}} = \frac{s_k}{\sqrt{\sigma^2_{acq}/\tau_k}} \tag{3}$$

As an example, if we double measurement time at each k-space location (e.g., acquire two samples rather than one), the modeled signal remains the same, the noise variance is reduced by a factor of 2, and the SNR increases by a factor of $\sqrt{2}$.

We model the measurement time at the k-th location in k-space, τ_k, as the acquisition time per sample times the number of samples:

$$\tau_k = \tau_{acq} n_k \tag{4}$$

Without loss of generality, we assume a fixed acquisition time for every sample, τ_{acq}, defined by the acquisition parameters and the number of samples, n_k, that may vary across k-space locations.

The measurement time τ_k is not necessarily equal for all k-space locations. Variable density sampling across k-space can be a natural effect of certain acquisition techniques, such as radial sampling. Variable density sampling may also be used to take advantage of the higher energy in the low frequency regions to improve SNR (similar to Weiner filtering) or to account for asymptotic incoherence (11). A variable density distribution of measurement time generates a corresponding distribution of expected noise variance across k-space. Lower sampling density at the high frequency k-space locations results in higher vari-
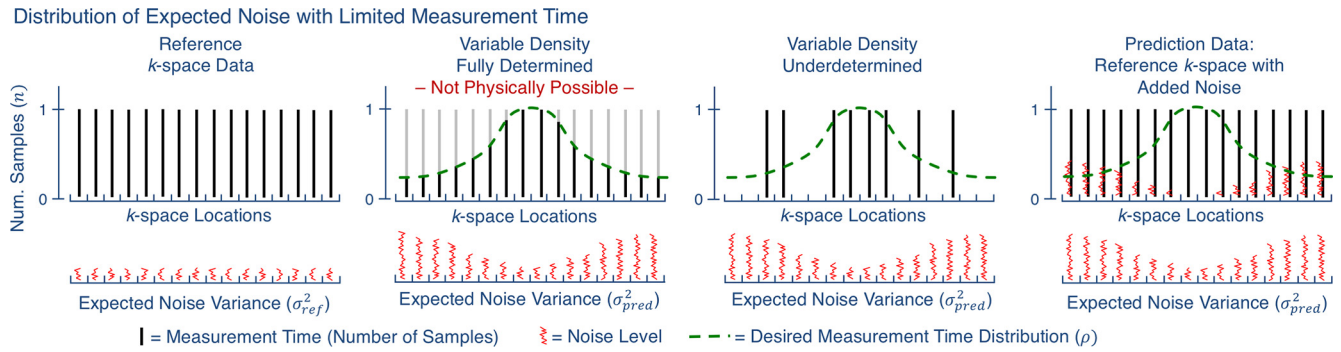
Figure 3. With limited measurement time, the sampling density distribution τ (dashed green line) may fall below one unit of measurement time. For systems with a minimum measurement time, fractional samples (second column) are not possible/do not contribute to reduction in scan time, and we are forced to sample below the Nyquist rate (third column) to meet the required measurement time limit. To simulate the infeasible Nyquist-sampled, fully determined acquisition (second column), the image quality prediction process adds noise to a fully determined reference acquisition (fourth column). Note that all three of these datasets have the same distribution of expected noise variance across k-space (bottom row).

ance at these locations, generating a colored (blue) noise distribution, rather than the white noise associated with a uniform acquisition time distribution. It is this colored noise that is coupled with the underdetermined system's effects during image reconstruction.

Note that in this paper, we are not attempting to determine the optimal sampling density, but rather providing a tool to help analyze the effects of the chosen sampling density as well as other acquisition and reconstruction parameters.

Undersampling and Expected Measurement Time

Fast and/or short acquisitions require a limit on the total measurement time. Unfortunately, some systems and applications have constraints on the minimum measurement time at a single k-space location. In this case, it is not feasible to sample k-space such that the reconstruction system is fully determined (Figure 3, second column). Undersampling is required to meet the measurement time constraints without sacrificing other scan requirements, such as spatial resolution, that are defined by the desired measurement time distribution (Figure 3, dashed green curve). Without loss of generality, we will define the system's minimum measurement time to be one sample of duration τ_{acq} and any shorter acquisition times, $\tau_k < \tau_{acq}$, are infeasible.

Undersampling (Figure 3, third column) avoids acquiring fractional samples by measuring either one or zero samples at each k-space location. A binary undersampling pattern can be constructed to fit the desired sampling density, whether it be uniform or variable density. This technique of constructing a continuous output with discrete inputs is analogous to pulse-width modulation in digital signal generation and to digital halftoning in computer graphics.

At first, it may appear that the SNR using these binary undersampling patterns is the same as a fully sampled acquisition because at the k-space locations where we collect a measurement, it has the same variance, σ^2_{acq}, as any fully sampled measurement. Also, at locations where we don't measure any

signal, we also don't collect any noise. However, the Fourier transform effectively averages the measured k-space locations with the zeros from the missing measurements, scaling the SNR by the square root of the sampling density equation (3). To model this averaging effect based on the density of binary sampling patterns, we consider the expected measurement time at each k-space location, τ_k.

We model the expected measurement time for random undersampling patterns by considering the generation of a random sampling pattern. The binary value for each location in the pattern may be determined by drawing a random sample from a Bernoulli distribution. To generate a pattern with a particular sampling density, the mean parameter of each Bernoulli distribution is set to the desired fractional measurement time, ρ_k, for that location. Specifically, let us model T_k as a Bernoulli random variable representing the measurement time at a single location in k-space. The expected value of T_k is $\tau_{pred,k}$:

$$Sample_k \sim Bern(\rho_k) \tag{5}$$

$$T_k = \tau_{acq} n_k Sample_k \tag{6}$$

$$\tau_{pred,k} = \mathbb{E}[T_k] = \tau_{acq} n_k \rho_k \tag{7}$$

$\tau_{pred,k}$ gives us the expected measurement time per k-space location, which in turn leads us to the expected noise variance per k-space location, $\sigma^2_{pred,k} = \sigma^2_{acq}/\tau_{pred,k}$.

Image Quality Prediction

Using the expected measurement time described in the previous section, the image quality prediction process generates an image that shows the empirical effects of reduced measurement time without any effects of an underdetermined system caused by undersampling. This process, as depicted in Figure 1, creates the prediction image by adding noise (based on the expected measurement time of a specific undersampling pattern) to a fully sampled reference k-space dataset and then passing that adjusted k-space through the regularized weighted least squares reconstruction algorithm described in the following section.
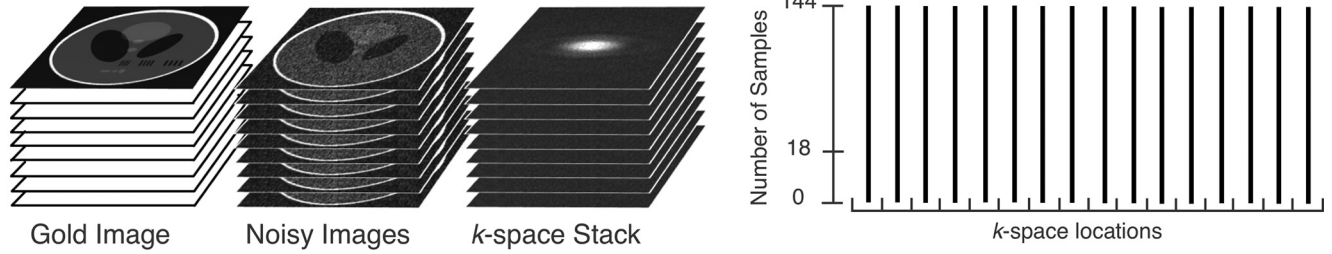
Figure 4. Experimental setup allowing us to choose the number of acquisition samples (from 0 to 144) at each *k*-space location. Noise is added to 144 copies of the input gold image. These noisy images are then Fourier transformed to create a stack of *k*-space images with 144 samples available at each *k*-space location. Note that for the tomato dataset, there is no gold image and the *k*-space stack comes directly from the 144 scanner acquisitions of the tomato.

The first step in the prediction process is to determine the expected measurement time at each *k*-space location, $\tau_{pred,k}$, for the given undersampling pattern. For random sampling patterns, this sampling density distribution is readily available, as it is the same distribution that generated the sampling pattern. When the sampling density is not explicitly or analytically available, the measurement time distribution may be approximated from the sampling pattern with local averaging, Voronoi diagrams, or other techniques used in sampling density compensation.

From the measurement time distribution, we calculated how much noise needs to be added to the fully sampled (fully determined) reference *k*-space dataset to match the equivalent statistical noise produced by the given undersampling pattern. To simulate an undersampled acquisition with Gaussian noise variance $\sigma^2_{pred,k} = \sigma^2_{acq}/\tau_{pred,k}$, we simply added complex-valued Gaussian noise to the reference *k*-space based on the expected measurement time distribution, $\tau_{pred,k}$ from equation (7), (Figure 3, right). Given that $\sigma^2_{ref} = \sigma^2_{acq}/\tau_{ref}$ is the Gaussian noise variance measured from the reference data, we can calculate the variance of the complex Gaussian noise, $\sigma^2_{add,k}$, to add to location k in the reference *k*-space:

$$\sigma^2_{pred,k} = \sigma^2_{ref} + \sigma^2_{add,k} \qquad (8)$$

$$\sigma^2_{add,k} = \sigma^2_{pred,k} - \sigma^2_{ref} \qquad (9)$$

$$= \left(\frac{\tau_{ref}}{\tau_{pred,k}} - 1 \right) \sigma^2_{ref} \qquad (10)$$

where $\tau_{ref} = \tau_{acq} n_{ref}$ in equation (4) and n_{ref} is the number of samples acquired in the reference data. The detailed derivation between equations (9) and (10) may be found in the online supplementary Appendix. Often $n_{ref} = 1$; however, the reference data may be acquired using many samples, for example the number of averages might equal two or, in the case of our first two experiments, $n_{ref} = 144$ (Figure 4).

Note that with variable density sampling patterns, $\tau_{pred,k}$, is not constant across *k*-space, and thus, the variance of the added noise, $\sigma^2_{add,k}$, will also vary across *k*-space.

The noise to add at each point in *k*-space is drawn from a complex-valued, zero-mean Gaussian distribution with vari-ance equal to $\sigma^2_{add,k}$ for that *k*-space location. This noise is simply added to the reference *k*-space to produce fully determined *k*-space with the noise distribution matching that of the under-sampled data (Figure 3, right).

The final step in the image quality prediction process is to pass the reference *k*-space with added noise through the regularized weighted least squares reconstruction algorithm described in the following section, producing a prediction image that gives an estimate of the reconstruction image quality assuming no effect from an underdetermined reconstruction system.

Weighted Least Squares Reconstruction

We require a consistent reconstruction formulation that supports standard fully sampled and undersampled data as well as the prediction data. To this end, we use a maximum *a posteriori* (MAP) formulation of MRI reconstruction that leads, in general, to a regularized weighted least squares optimization. Equations (1) and (2) combine to give us a Gaussian likelihood probability of measuring a signal y_k given an image object m:

$$P(y_k|m) = \frac{1}{\sqrt{2\pi\sigma^2_{acq}/\tau_k}} exp\left(\frac{-|y_k - (Fm)_k|^2}{2\sigma^2_{acq}/\tau_k} \right) \qquad (11)$$

With this Gaussian likelihood and assuming a general prior probability on our image data $P(m)$, the resulting MAP formulation leads to a weighted-least squares optimization:

$$m^* = \underset{m}{\text{argmax }} P(m|y) \qquad (12)$$

$$= \underset{m}{\text{argmax }} P(y|m)P(m) \qquad (13)$$

$$= \underset{m}{\text{argmin }} \frac{1}{2}\sum_{k=1}^{N_P} \tau_k|y_k - (Fm)_k|^2 - \log P(m) \qquad (14)$$

$$= \underset{m}{\text{argmin }} \frac{1}{2} \|Wy - WFm\|^2_2 - \log P(m) \qquad (15)$$

where m is the vectorized image with N_P number of pixels; y is the vectorized acquired *k*-space locations with N_P number of elements; F is the N_PxN_P multidimensional discrete Fourier transform operator; and W is an N_PxN_P diagonal matrix with

$W_{k,k} = \sqrt{\tau_k}$ values along the diagonal. The detailed derivation between equations (13) and (14) may be found in the online supplementary Appendix.

In this paper, we will use a Laplacian-based prior to promote sparsity, $(-\log P(m) = \lambda \|\Psi(m)\|_1)$, where Ψ is a sparsity transform function and λ is the Laplace prior parameter. This ℓ_1 regularized weighted least squares (WLS) optimization does not have an analytic solution, and finding the solution requires a nonlinear reconstruction algorithm. In general, we can solve this optimization using an iterative algorithm, such as fast iterative shrinkage-thresholding algorithm (FISTA) (12) or alternating direction method of multipliers (ADMM) (13).

This optimization framework, given the proper weight values described below, generalizes the reconstruction of a) fully sampled, b) undersampled, and c) image quality prediction datasets.

a) Fully sampled weights: The least squares weights for a fully sampled dataset (both uniform and variable density sampling) are simply equal to the square root of the measurement time, $W_{k,k} = \sqrt{\tau_k} = \sqrt{\tau_{acq} n_k}$, for the k-th sample. Assuming, again, that the acquisition time per sample is constant across k-space, τ_{acq} may be pulled out of the ℓ_2 norm term, simplifying the weights to be equal to the square root of the number of samples, $W_{k,k} = \sqrt{n_k}$.

Note that with n_k constant across k-space and a uniform prior $P(m)$, the MAP optimization becomes the standard least squares optimization:

$$m^* = \underset{m}{\arg\min} \frac{1}{2} \|y - Fm\|_2^2 \qquad (16)$$

b) Undersampled weights: When undersampling, the weights, $W_{k,k}$, are simply set to one or zero depending on whether or not that k-space location has been sampled (assuming the same measurement time at each sampled location). With these binary weights, the operator W in equation (15) becomes the undersampling operator defined by the binary sampling pattern. With a Laplacian-based prior, the MAP reconstruction becomes the standard Lasso optimization (14) commonly used in compressed sensing. In addition to strictly binary undersampling patterns, the WLS optimization also allows for undersampling patterns that have zero measurement time at certain locations and a range of measurement times across the remaining locations, for example, an acquisition with undersampled high frequencies and oversampled low frequencies.

c) Prediction weights: The prediction data are designed to simulate the noise variance from the expected measurement time for a given sampling density ρ_k, leading us to WLS weights $W_{k,k} = \sqrt{\tau_{pred,k}} = \sqrt{\tau_{acq} n_{ref,k} \rho_k}$, which may be simplified to $W_{k,k} = \sqrt{\rho_k}$ assuming constant sampling time and constant number of samples per location in the fully sampled reference data.

METHODOLOGY

In this institutional review board-approved study, we acquired MRI data by scanning two healthy, adult volunteers.

Effect of Measurement Time Distribution

To better understand effects of reduced measurement time and undersampling and to test our image quality prediction process, we created an experiment that enables us to compare the recon-structions of 1) a fully determined dataset, 2) an underdetermined dataset, and 3) the corresponding prediction data, all using the same total measurement time and sampling density distribution.

The foundation of this experiment is a "stack" of 144 fully sampled k-space images (Figure 4). Each entry in the k-space stack is a different noisy acquisition of the same object slice. With 144 samples available at each of the N k-space locations, we are able to select a subset of these samples to simulate acquiring a specific number of samples at each k-space location based on a desired measurement time distribution.

We used two different datasets for this experiment. The first dataset was the classic Shepp-Logan digital phantom (15) with a slight modification to add a set of parallel dark bars that will help analyze spatial resolution. This phantom was chosen because it has an explicitly sparse representation (many true zero values) in the finite differences domain (often seen in total variation reconstructions), implying that we can use compressed sensing to find a proper solution to the underdetermined system of equations caused by undersampling. As seen in Candes, Romberg, and Tao (16), the Shepp-Logan phantom, without noise, may be perfectly recovered after severe undersampling. To analyze how noise propagates through the reconstruction system, we generated a different instance of complex-valued, zero-mean, Gaussian noise to add to 144 copies of the k-space for the Shepp-Logan phantom. The second dataset is 144 actual MRI acquisitions of a tomato at a single slice location. These data were acquired on a 3T scanner (Siemens Healthineers, Erlangen, Germany) using a T1-weighted gradient echo sequence with 10 ms echo time (TE), 35 ms repetition time (TR), 12° flip angle (FA), 90 mm field of view (FOV), 2 mm slice thickness, and 192 $\times$ 192 acquisition matrix. Only the body coil was used during acquisition to both simplify the reconstruction model and ensure that each of the 144 acquisitions had relatively low SNR.

For both datasets, we selected a subset of the full stack of k-space samples based on three different sampling distributions, as depicted in the top row of Figure 5: **reference**, using all $144N$ samples (where N is the number of k-space locations); **fully determined**, selecting only $18N$ samples according to either a uniform or variable density sampling distribution across k-space locations; and **underdetermined**, selecting $18N$ samples and following the same density distribution but collecting all 144 samples at $N/8$ randomly chosen k-space locations and collecting zero samples for the remaining locations. We also reconstructed both datasets using the **image quality prediction** process to add noise to the $144N$ reference dataset to simulate the noise level from the $18N$ fully determined dataset.

For all reconstructions, the selected k-space samples were averaged at each k-space location to create a single k-space image to be reconstructed (y, from equations (12) and (15)).

We reconstructed all data using our implementation of ADMM, formulated for the regularized weighted least squares optimization, with the weights equal to the number of measurements acquired at each k-space location, as specified in the weighted least squares reconstruction section of the introduction. For the digital phantom dataset, we used isotropic total variation as the sparsity model. For the single-channel MRI
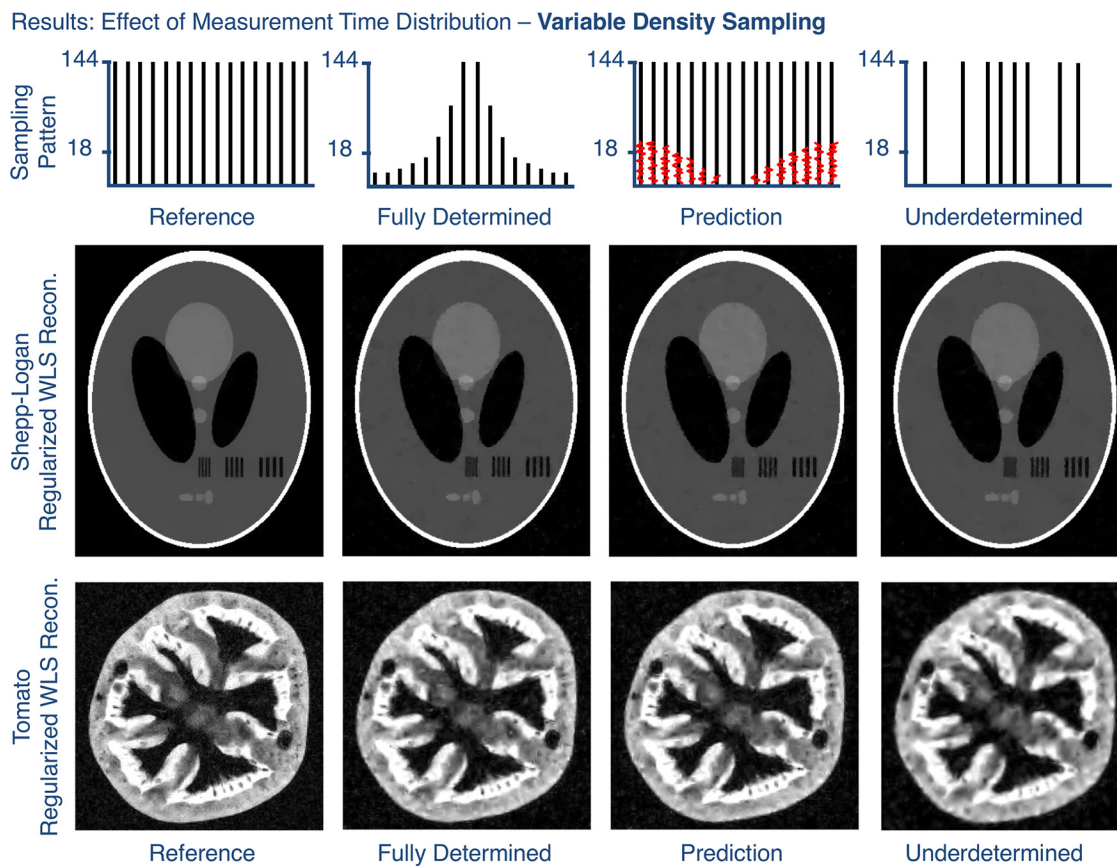
Results: Effect of Measurement Time Distribution – **Variable Density Sampling**

Figure 5. Results from the effect of the measurement time distribution experiment using variable density sampling. Each column uses a different set of *k*-space samples; from left to right: **oversampled reference**, using all 144 samples at each *k*-space location; **variable density, fully determined**, using 1/8 of the total samples following a variable density distribution; **prediction data**, using fully sampled *k*-space with noise added to simulate the variable density, fully determined dataset; **variable density (randomly sampled), underdetermined**, using 1/8 of the total samples and following the same variable density distribution, but only using either 144 or zero samples at each location. *Row 1*: Illustration of how measurement time is distributed across *k*-space. *Row 2*: WLS reconstruction of the Shepp-Logan data, regularized with total variation. *Row 3*: WLS reconstruction of the tomato *k*-space data, regularized with wavelets.

acquired data, we used Daubechies-4 wavelets with translation invariant cycle spinning (17) as the sparsity model.

Effects of Measurement Noise and Undersampling Rate

Given enough acquisition time, we can satisfy a given sampling density distribution by either Nyquist sampling *k*-space or by undersampling. Both of these sampling patterns produce similar distributions of expected noise variance in our data, but undersampling incurs an additional cost from having an underdetermined system of equations. In this experiment, we will extend the oversampled stack experiment above to take a closer look at the effect of measurement noise and undersampling rate on reconstruction image quality. We accomplish this by varying both the measurement noise level and the undersampling rate and then comparing the mean squared error (MSE) images reconstructed from variable density **fully determined** data and from variable density **underdetermined** data.

As in the measurement time experiment above, we have a stack of *k*-space data, and we generate an output image by

reconstructing a subset of *k*-space samples, selected according to a either a fully determined variable density sampling pattern or an underdetermined pattern following the same measurement time distribution.

In this experiment, the *k*-space stack is generated from copies of a single relatively high SNR (31.3 dB) in vivo head acquisition. Similar to the Shepp-Logan *k*-space stack, we added to *k*-space a sample of complex-valued, Gaussian noise with zero mean and a given standard deviation. We executed the experiment using three different values for the added noise standard deviation (1, 5, 8) and four undersampling rates (2×, 4×, 8×, and 12× undersampled).

The head dataset for this experiment is an axial slice of a three-dimensional (3D) fully sampled, spoiled gradient echo dataset acquired on a 1.5T scanner (GE Healthcare, Waukesha, WI) with 8 receive channels, 5 ms TE, 12 ms TR, 20° FA, 184 × 230 mm FOV, 1 mm slice thickness, and 256 × 256 acquisition matrix. This multichannel dataset was preprocessed, using ES-
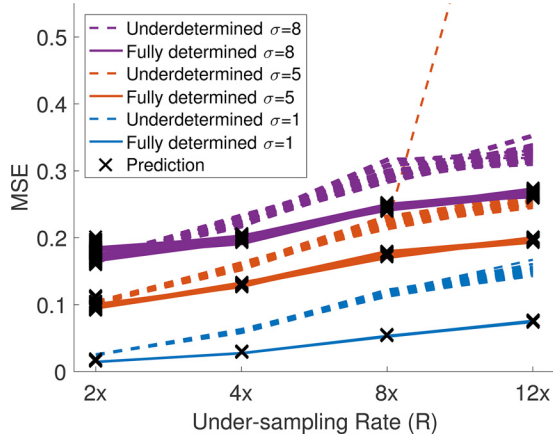
Figure 6. Results of our experiment to compare fully determined and underdetermined reconstructions with the same total measurement time across four different undersampling rates (2×, 4×, 8×, 12×) for three different noise levels (added noise standard deviations 1.0, 5.0, 8.0). Mean squared error (MSE) values are plotted for each of the 100 repetitions of the same experiment. Note that for the 2× undersampling rate, the fully determined and underdetermined reconstructions have essentially the same MSE. As the undersampling rate increases, the underdetermined system produces an increasingly worse MSE than the fully determined system. Note that one of the 100 underdetermined reconstructions at $\sigma = 5$ and $R = 12\times$ failed to converge. This outlier is consistent with compressed sensing theory and practice where the reconstruction may fail to converge at higher undersampling rates.

PIRiT coil sensitivity maps ([18]), to combine the data into a single channel, allowing us to use a simpler reconstruction model for this experiment. This head dataset has relatively high SNR, so we were able to experiment with very low noise and subsequently experiment with higher noise levels by adding Gaussian noise to the k-space stack for this dataset. This head dataset also provides a real example of an image that is only approximately sparse in the wavelet transform domain.

We repeated these 12 experiments (three noise levels by four undersampling rates) 100 times, each time reconstructing the fully determined data and the underdetermined data, as well as the corresponding prediction data. We then plotted the resulting MSE values (relative to the original head image) (Figure 6).

Image Quality Prediction

We demonstrate the image quality prediction process by comparing the output of the actual undersampled reconstruction to both the generated prediction image and the fully determined reference image. We executed this experiment for two in vivo fully sampled MRI datasets using increasingly aggressive retrospective undersampling rates.

In vivo Knee

The in vivo knee dataset is an axial slice of a 3D fully-sampled, fast spin echo dataset acquired on a 3T scanner (GE Healthcare, Waukesha, WI) with 8 receive channels, 25 ms TE, 1550 ms TR, echo train length of 40, 160 mm FOV, 0.6 mm slice thickness, and 320 × 320 acquisition matrix. This dataset was collected by Epperson et al. ([19]) and is available at ([20]).

The two retrospective undersampling patterns used were 4× and 12× undersampled, variable density Poisson disc. Both patterns fully sampled the center of k-space to allow for ESPIRiT auto-calibration ([18]). Neither the reference data nor the undersampling patterns included the corners of k-space, a common acquisition acceleration.

The optimization equation for this parallel imaging, compressed sensing reconstruction is an extension of equation ([15]), modified to include parallel imaging and a Laplacian prior:

$$\min_{m} \frac{1}{2} \| Wy - WFSm \|_2^2 + \lambda \| \Psi m \|_1 \qquad (17)$$

where m is the vectorized image with N_P number of pixels; y is the vectorized acquired multichannel k-space data with $N_C N_P$ number of elements (N_P is the number of pixels, N_C is the number of coils); F is the $N_C N_P \times N_C N_P$ two-dimensional (2D) Fourier transform operator for each coil independently; S is the $N_C N_P \times N_P$ block diagonal sensitivity maps generated with ESPIRiT calibration; Ψ is the sparsity transform; λ is the regularization parameter; and W is the $N_C N_P \times N_C N_P$ diagonal weight matrix.

Note that the reconstruction process now includes the parallel imaging coil combination operator S^H. With the addition of parallel imaging, the undersampled reconstruction system is now both ill-conditioned and underdetermined. Previous works have provided tools to empirically analyze the noise propagation through the ill-conditioned parallel imaging system, for example by computing the geometry-factor ([21]) or with Monte Carlo simulations with added noise ([22]). The image quality prediction process will empirically show the effect of lower SNR because of reduced measurement time on the compressed sensing and parallel imaging reconstruction without any effect from an underdetermined or ill-conditioned system. The actual underdetermined reconstruction will then produce an image affected by similar lower SNR as well as the effects from the ill-conditioned and underdetermined parallel imaging and compressed sensing system.

The sparsity filter (associated with Ψ) used within the reconstruction was wavelet soft-thresholding using Daubechies-4 with translation invariant cycle spinning ([17]).

The regularized weighted least squares optimization for both prediction and underdetermined reconstruction used our implementation of the ADMM algorithm. The only difference between the two reconstructions was the appropriate change to the weights as specified in the weighted least squares section of the introduction. Specifically, the weights for the prediction reconstruction were the square root of the sampling density, ρ_k, at each k-space location, and the actual underdetermined reconstruction weights were binary with ones for acquired locations and zeros elsewhere.
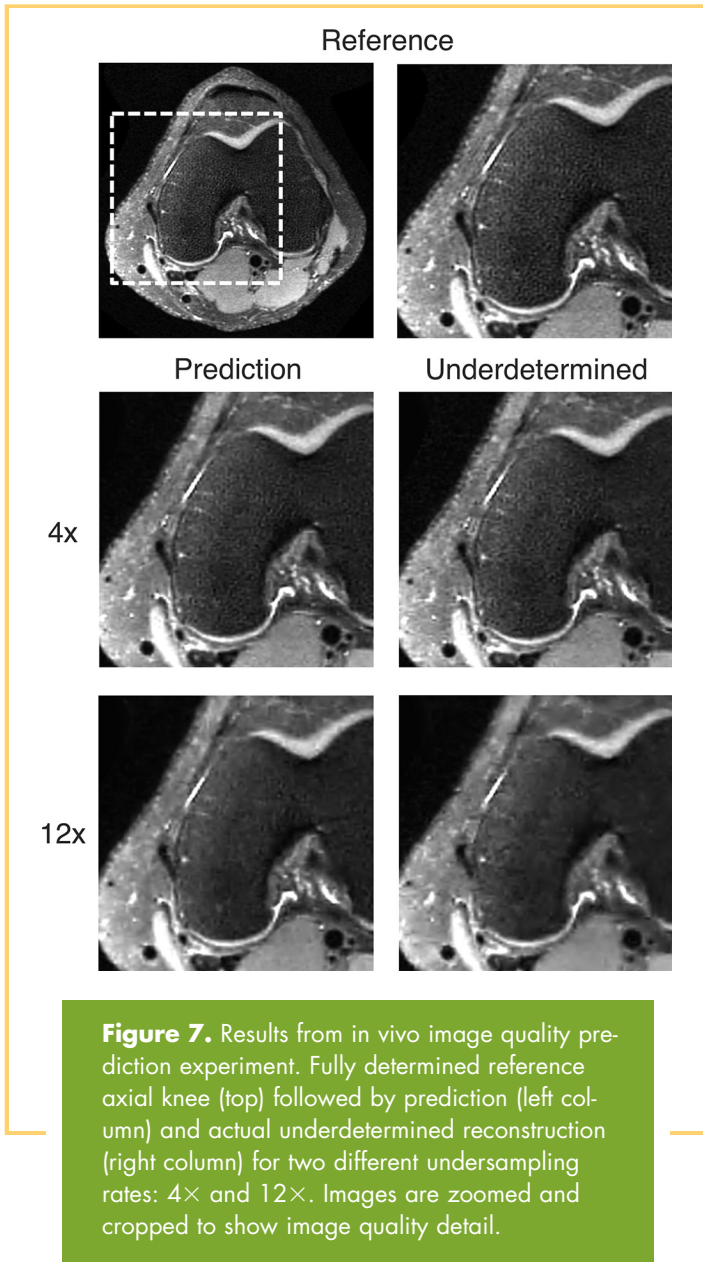
Figure 7. Results from in vivo image quality prediction experiment. Fully determined reference axial knee (top) followed by prediction (left column) and actual underdetermined reconstruction (right column) for two different undersampling rates: 4× and 12×. Images are zoomed and cropped to show image quality detail.

The image quality prediction process requires an understanding of the existing noise level in the fully sampled reference data (σ_{full}^2). Ideally, this noise level could be obtained from an explicit measurement of the received signal using the coils on the same scanner, prior to the actual exam. In our experiments, we measured the noise level from the reference data directly by Fourier transforming the (multichannel) k-space data and measuring the variance of the values from a 11×11 background patch in each coil image. The noise level was measured and applied independently for each coil channel.

A direct inverse 2D Fourier transform followed by coil combination ($m_{ref} = S^H F^{-1} y_{ref}$) was used on the reference k-space to generate the fully sampled reference image for comparison (Figure 7, top).

In vivo Head
The in vivo head dataset is an axial 2D fast spin echo dataset acquired on a 3T scanner (Siemens Healthineers, Erlangen, Ger-

many) with 12 receive channels, 91 ms TE, 6000 ms TR, echo train length of 11, 195 × 220 mm FOV, and 286 × 320 acquisition matrix. The 12 coil channels were reduced to 4 channels with Siemens coil compression. Multiple slices were acquired at slice thicknesses of 1 mm and 2 mm. The phase encode lines were retrospectively undersampled at 4× and 6× acceleration using a one-dimensional (1D) variable density Poisson disc sampling with the central 24 lines fully sampled. This dataset was processed in the same manner as the in vivo knee dataset above.

RESULTS
The following three results are shared across all of our experiments:

1. the **prediction** image has equivalent or worse image quality than the **reference** image,
2. the **undersampled** reconstruction image has equivalent or worse image quality than the **prediction** image,
3. the **prediction** image for a given sampling density has equivalent image quality to the **fully determined** image with the same sampling density.

Effect of Measurement Time Distribution
Figure 5 shows the results of our experiment to test the effect of various measurement time distributions on reconstruction image quality. For both the Shepp-Logan digital phantom and the MRI acquisition of the tomato, the fully determined images with reduced measurement time show lower image quality than the images reconstructed from the reference acquisition data. As seen specifically in the blurred spatial vertical bars, the fully determined images did not fully recover from the limited acquisition time despite not having any corruption from an underdetermined systems of equations.

The variable density underdetermined Shepp-Logan reconstruction (Figure 5, third row, right) was successful and has nearly identically image quality to the fully determined reconstruction but still lower image quality than the reference reconstruction. This indicates that the underdetermined reconstruction recovered well from the underdetermined system, but still could not completely recover from the lower SNR because of reduced measurement time. For the acquired tomato dataset, however, the underdetermined image quality (Figure 5, bottom, right) is lower than the prediction and fully determined image quality, indicating that the reconstruction could not completely recover from the underdetermined system. This is not a surprising result because the tomato image is not sufficiently sparse in the wavelet transform domain, especially when compared to the explicit sparsity of the Shepp-Logan phantom in the finite differences domain.

Figure 5 also shows the results of the image quality prediction process for the same two datasets and sampling distributions. The second and third columns in this figure show that the fully determined reconstructions have essentially identical image quality to their corresponding prediction images. This verifies that the image quality prediction process closely simulates the noise level and reconstructed image quality of the associated fully determined acquisitions.

Similar results from the same experiment with uniform density sampling have been included in the online supplementary Appendix.

Effects of Measurement Noise and Undersampling Rate

By varying the input noise level and the undersampling rate, we see the differences in the resulting MSE for the reconstructions of the fully determined data and underdetermined data, both with the same measurement time distribution. Figure 6 shows that for a fixed noise level and increasing undersampling rate, the MSE of the fully determined images increases, showing that the reduced measurement time affects image quality despite no undersampling. Also, as we increase the undersampling rate, the MSE of the underdetermined images increases significantly faster than the fully determined images. This gap in image quality shows the degrading effect of the underdetermined reconstruction increasing as the undersampling rate increases and the sparsity transform can no longer adequately model the image in a sufficiently sparse representation.

As seen in Figure 6, the MSE of the prediction images matches the MSE of the fully determined reconstructions for all noise levels and undersampling rates, indicating that the image quality prediction process is consistently simulating the expected noise level for the given sampling density.

The results from this experiment help us to see that when the image quality of the prediction image is unacceptable, the actual undersampled reconstruction will also be unacceptable (i.e., higher MSE). In this situation, the low SNR of the acquisition is the limiting factor in the reconstruction, not the artifacts because of the underdetermined system. To improve the reconstruction in this case, steps should be taken to adjust the acquisition parameters to increase the SNR or better handle the expected noise levels (e.g., reducing spatial resolution, decreasing undersampling rate, or improving the image prior $P(\boldsymbol{m})$).

Image Quality Prediction

Figure 2 shows the prediction and underdetermined reconstruction images for the in vivo head experiment. This figure illustrates how the prediction image may be used to gain insight into the causes of poor undersampled image quality and adjust scan parameters, such as slice thickness, as needed.

Figure 7 shows the reference, prediction, and underdetermined reconstruction images for the in vivo experiment using the knee dataset and various undersampling rates. Figure 7 shows the following three qualitative results: 1) the reference image has better image quality than the prediction images; 2) the prediction images have better image quality than the corresponding underdetermined images; and 3) the underdetermined images are more similar in image quality to the prediction images than the reference image. That the reference images look better than the prediction images is expected because the prediction process adds more noise to the fully sampled reference data. That the prediction images look better than the underdetermined image is expected because the underdetermined reconstruction had to find a proper solution to an underdetermined system of equations in addition to recovering from the lower SNR from reduced measurement time. Finally, the prediction image provides a better estimate of reconstruction image quality than the reference image.

With a reasonable amount of undersampling, the 4× underdetermined images only have slightly lower image quality than the prediction images. As undersampling increases to 12×, the image quality gap between the underdetermined and prediction images increases. These results are consistent with our effect of measurement noise and undersampling rate experiment when increasing sampling rate.

DISCUSSION

The presented **image quality prediction** process provides an empirical upper bound on undersampled image quality, which serves as a better metric for evaluating the effectiveness of a reconstruction algorithm than direct comparison to a fully sampled reference reconstruction. The prediction process enables an analysis of a reconstruction algorithm's ability to handle lower SNR because of reduced measurement time without any effect from an underdetermined system. By simulating the effect of lower SNR without any underdetermined effects, the prediction process allows us to determine whether a reconstruction is actually limited by our sparse recovery or simply limited by low acquisition SNR. Comparison of the prediction image to the reference reconstruction provides a means to assess the effects of lower SNR on reconstruction image quality. Comparison of the prediction image to the underdetermined reconstruction enables us to analyze what artifacts are introduced when undersampling is used rather than fully determined following the same measurement time distribution. The image quality prediction results and analysis are consistent with our experiments using our highly oversampled datasets to explicitly compare reconstruction results from variable density fully determined and underdetermined data.

An additional benefit of the prediction process is that it may be used to compare and tune different reconstruction algorithms or parameters, assessing how different reconstruction systems handle the lower SNR because of reduced measurement time in addition to comparing the actual undersampled reconstructions.

A limitation of the image quality prediction process is that it requires a fully sampled reference dataset. Access to a fully sampled acquisition is not always possible, especially in cases with 3D or four-dimensional (4D) dynamic imaging, where long, fully sampled acquisition times are not practical. Also, the image quality prediction process can isolate the effects of low SNR from the effects from an underdetermined system, but it cannot do the contrary, that is, it cannot isolate the effects from an underdetermined system from the effects of low SNR. Future work could investigate the effects of underdetermined systems using in vivo data by reconstructing fully sampled reference datasets that are highly oversampled to have minimal input noise, σ_{ref}. Of course, the effects of the underdetermined system would still be dependent on the image content, which varies significantly across clinical applications.

While developing the image quality prediction process, we use a maximum *a posteriori* formulation to derive a **general weighted least squares optimization framework** that accounts for both uniform and variable density sampling patterns, with

undersampling as a special case using binary weights. This WLS formulation adjusts the standard least squares term to account for the colored noise arising from the distribution of expected measurement time across k-space locations. Future work could develop methods to similarly incorporate the effects of measurement time distribution into the sparsity regularization term, allowing the sparsity filters to better recover from colored noise in addition to incoherent aliasing.

In the spirit of reproducible research, software and datasets to generate the results described in this paper will be available at http://www.eecs.berkeley.edu/~mlustig/Software.html (23).

Supplemental Materials

Supplemental Appendix: http://dx.doi.org/10.18383/j.tom.2017.00019.sup.01

Disclosures: No disclosures to report.

Conflict of Interest: The authors have no conflict of interest to declare.

REFERENCES
1. Sodickson DK, Manning WJ. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays. Magn Reson Med. 1997;38(4):591–603.
2. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med. 2007;58(6):1182–1195.
3. Song J, Liu QH, Johnson GA, Badea CT. Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT. Med Phys. 2007; 34(11):4476–4483.
4. Sidky EY, Kao CM, Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. J Xray Sci Technol. 2006;14(2):119–139.
5. Chinn G, Olcott PD, Levin CS. Improving SNR with a maximum likelihood compressed sensing decoder for multiplexed PET detectors. In: Nuclear Science Symposium Conference Record (NSS/MIC). 2010 IEEE; 2010. 3353–3356.
6. Valiollahzadeh S, Chang T, Clark JW, Mawlawi OR. Image recovery in PET scanners with partial detector rings using compressive sensing. In: Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE; 2012. p. 3036–3039.
7. Candès EJ. The restricted isometry property and its implications for compressed sensing. Comptes Rendus Mathematique. 2008;346(910):589–592.
8. Candès EJ, Plan Y. A probabilistic and RIPless theory of compressed sensing. IEEE Trans Inf Theory. 2011;57(11):7235–7254.
9. Wainwright MJ. Sharp Thresholds for High-Dimensional and Noisy Sparsity Recovery Using ℓ_1-Constrained Quadratic Programming (Lasso). IEEE Trans Inf Theory. 2009 May;55(5):2183–2202.
10. Macovski A. Noise in MRI. Magn Reson Med. 1996;36(3):494–497.
11. Adcock B, Hansen AC, Poon C, Roman B. Breaking the coherence barrier: A new theory for compressed sensing. arXiv preprint arXiv:13020561. Available from: https://arxiv.org/abs/1302.0561.2013.
12. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci. 2009;2(1):183–202.
13. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn. 2011 Jan;3(1):1–122.
14. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol. 1996;58(1):267–288.
15. Shepp LA, Logan BF. The Fourier reconstruction of a head section. IEEE Trans Nucl Sci. 1974;21(3):21–43.
16. Candès EJ, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inf Theory. 2006;52(2):489–509.
17. Coifman RR, Donoho DL. Translation-Invariant De-Noising. In: ANTONIADIS, A. and G. OPPENHEIM, (eds.), Wavelets and Statistics. vol. 103 Lecture Notes in Statistics. New York:Springer; 1995:125–150.
18. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M. ESPIRiT - an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. Magn Reson Med. 2014;71(3):990–1001.
19. Epperson K, Sawyer AM, Lustig M, Alley M, Uecker M, Virtue P, Lai P, Vasanawala S. Creation of fully sampled MR data repository for compressed sensing of the knee. In: 2013 Meeting Proceedings. Section for Magnetic Resonance Technologists; 2013.
20. MRI Data Repository. c2013- [cited 09.02.2017]. Available from: http://mridata.org/.
21. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. Magn Reson Med. 1999;42(5):952–962.
22. Thunberg P, Zetterberg P. Noise distribution in SENSE-and GRAPPA-reconstructed images: a computer simulation study. Magn Reson Imaging. 2007;25(7):1089–1094.
23. Lustig M. Michael Lustig, Software. c2010- [cited 09.02.2017]. Available from: http://www.eecs.berkeley.edu/~mlustig/Software.html