Compressive Imaging via Approximate Message Passing with Wavelet-Based Image Denoising

Jin Tan, Yanting Ma, and Dror Baron
Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695, USA
Email: {jtan,yma7,barondror}@ncsu.edu

***Abstract*—We consider compressive imaging problems, where images are reconstructed from a reduced number of linear measurements. Our objective is to improve over current state of the art compressive imaging algorithms in terms of both reconstruction error and runtime. To pursue our objective, we propose a compressive imaging algorithm that employs the approximate message passing (AMP) framework. AMP is an iterative signal reconstruction algorithm that performs scalar denoising of noisy signals. In this work, we apply an adaptive Wiener filter, which is a wavelet-based image denoiser, within AMP. Numerical results show that the proposed algorithm improves over the state of the art in both reconstruction error and runtime.**

***Index Terms*—approximate message passing, compressive imaging, image denoising, wavelet transform.**

I. Introduction

A. Motivation

Compressed sensing (CS) [1, 2] has sparked a tremendous amount of research activity in recent years, because it performs signal acquisition and processing using far fewer samples than required by the Nyquist rate. Breakthroughs in CS have the potential to greatly reduce the sampling rates in numerous signal processing applications such as cameras [3], medical scanners, fast analog to digital converters [4], and high speed radar [5].

Compressed sensing has been used in compressive imaging, where the input signal is an image, and the goal is to acquire the image using as few measurements as possible. Acquiring images in a compressive manner requires less sampling time than conventional imaging technologies. Applications of compressive imaging appear in medical imaging [6–8], seismic imaging [9], and hyperspectral imaging [10, 11].

B. Related work

Many compressive imaging algorithms have been proposed in the literature. For example, Som and Schniter [12] modeled the structure of the wavelet coefficients by a hidden Markov tree (HMT), and applied a turbo scheme that alternates between inference on the HMT structure with standard belief propagation and inference on the compressed sensing measurement structure with the generalized approximate message passing algorithm. He and Carin [13] proposed a hierarchical Bayesian approach with Markov chain Monte Carlo (MCMC) for natural image reconstruction. Soni and Haupt [14] exploited a hierarchical dictionary learning method [15] and assumed that projecting images onto the learned dictionary will yield tree-sparsity, and therefore the nonzero supports of the dictionary can be identified and estimated accurately by setting an appropriate threshold.

However, existing compressive imaging algorithms may either not achieve good reconstruction quality or not be fast enough. Therefore, in this paper, we focus on a variation of a fast and effective algorithm called approximate message passing (AMP) [16] to improve over the prior art. AMP is an iterative signal reconstruction algorithm that performs scalar denoising within each iteration, and proper selection of the denoising function used within AMP is needed to obtain better reconstruction quality.

There are two popular classes of image denoisers: wavelet-based methods and dictionary learning methods. A typical wavelet-based image denoiser proceeds as follows: (*i*) apply a wavelet transform to the image and obtain wavelet coefficients; (*ii*) denoise the wavelet coefficients; and (*iii*) apply an inverse wavelet transform to the denoised wavelet coefficients, yielding a denoised image. Two popular examples of denoisers that can be applied to the wavelet coefficients are hard thresholding and soft thresholding, which were proposed by Donoho and Johnstone [17]. Other wavelet-based methods have been proposed by Mıhçak et al. [18, 19] and Figueiredo and Nowak [20].

Dictionary learning methods [21–23] generally achieve lower reconstruction error than wavelet-based methods. However, the learning procedure requires a large amount of training images, and may involve manual tuning. Owing to these limitations, our main focus in this paper is to integrate wavelet-based image denoisers into compressive imaging reconstruction algorithms.

C. Contributions

The objective of this paper is to develop compressive imaging algorithms that are fast and reconstruct well. In order to obtain a fast implementation, we utilize wavelet-based image denoisers within the AMP framework. We apply an existing image denoiser, an adaptive Wiener filter [18], within AMP. Our numerical results are promising, showing that the proposed algorithm outperforms the state of the art algorithms in both reconstruction quality and runtime. Although we only

This work was supported in part by the National Science Foundation under Grant CCF-1217749 and in part by the U.S. Army Research Office under Grant W911NF-04-D-0003.

present results for one wavelet-based image denoiser within AMP, we have tested another denoiser [20] and the results are favorable; we believe that other image denoisers would also work well within the AMP framework, and those denoisers need not be wavelet-based.

The remainder of the paper is arranged as follows. We review AMP in Section II, and describe an adaptive Wiener filter that is used within AMP in Section III. Numerical results are presented in Section IV, and the paper concludes with a discussion in Section V.

II. Review of Approximate Message Passing

A. Problem setting

Before reviewing AMP, let us first model how measurements are obtained in a compressive imaging system.

Matrix channels: We rearrange the input image $\mathbf{x}$, which is comprised of N pixels, as a column vector of length N (for example, for a 512×512 image, $N = 512^2$). Then, we multiply $\mathbf{x}$ by a known measurement matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, which has M rows (typically $M < N$). Finally, the measurements are corrupted by independent and identically distributed (i.i.d.) zero-mean Gaussian noise $\mathbf{z}$,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z}. \tag{1}$$

We observe the noisy vector $\mathbf{y} \in \mathbb{R}^M$, and want to estimate and reconstruct the original input signal $\mathbf{x}$ from $\mathbf{y}$ and $\mathbf{A}$.

Scalar channels: In AMP, the matrix channel is converted to scalar channels, where the noisy observations obey

$$q_i = x_i + v_i, \tag{2}$$

$i \in \{1, 2, \ldots, N\}$, the subscript $(\cdot)_i$ denotes the i-th component of a vector, and $\mathbf{x}, \mathbf{q} \in \mathbb{R}^N$ are the input signal and the noisy observations, respectively. The noise $\mathbf{v}$ is i.i.d. Gaussian, $v_i \sim \mathcal{N}(0, \sigma^2)$. Note that we use different notations for the noise and observations in matrix channels and scalar channels. The main difference between the two types of channels is that the observations $\mathbf{y}$ in the matrix channel contain linear combinations of the entries of $\mathbf{x}$.

B. Algorithmic framework

We are now ready to describe AMP [16], which is an iterative signal reconstruction algorithm in matrix channels. Consider a matrix channel model (1) where the signal distribution follows $x_i \sim f_X$ and the noise is i.i.d. Gaussian. The entries of the measurement matrix $\mathbf{A}$ are i.i.d. $\mathcal{N}(0, \frac{1}{M})$ distributed, and thus the columns of the matrix have unit ℓ_2-norm, on average. AMP [16] proceeds iteratively according to

$$\mathbf{x}^{t+1} = \eta_t(\mathbf{A}^T\mathbf{r}^t + \mathbf{x}^t), \tag{3}$$

$$\mathbf{r}^t = \mathbf{y} - \mathbf{A}\mathbf{x}^t + \frac{1}{R}\mathbf{r}^{t-1}\langle\eta'_{t-1}(\mathbf{A}^T\mathbf{r}^{t-1} + \mathbf{x}^{t-1})\rangle, \tag{4}$$

where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$, $R = M/N$ represents the measurement rate, $\eta_t(\cdot)$ is a denoising function, $\eta'_t(\mathbf{s}) = \frac{\partial}{\partial \mathbf{s}}\eta_t(\mathbf{s})$, and $\langle\mathbf{u}\rangle = \frac{1}{N}\sum_{i=1}^N u_i$ for some vector $\mathbf{u} = (u_1, u_2, \ldots, u_N)$. In the t-th iteration, we obtain the vectors $\mathbf{x}^t \in \mathbb{R}^N$ and $\mathbf{r}^t \in \mathbb{R}^M$. We highlight that the vector $\mathbf{A}^T\mathbf{r}^t + \mathbf{x}^t \in \mathbb{R}^N$ in (3) can be regarded as noisy measurements of $\mathbf{x}$ in the t-th iteration with noise variance σ_t^2, and therefore the denoising function $\eta_t(\cdot)$ is performed on a scalar channel (2). Let us denote the equivalent scalar channel at iteration t by

$$\mathbf{q}^t = \mathbf{A}^T\mathbf{r}^t + \mathbf{x}^t = \mathbf{x} + \mathbf{v}^t, \tag{5}$$

where $v_i^t \sim \mathcal{N}(0, \sigma_t^2)$. The asymptotic performance of AMP can be characterized by a state evolution (SE) formalism:

$$\sigma_{t+1}^2 = \sigma_z^2 + \frac{1}{R}E\left[\left(\eta_t\left(X + \sigma_t W\right) - X\right)^2\right]$$

where the random variables $W \sim \mathcal{N}(0, 1)$ and $X \sim f_X$. Formal statements about SE appear in Bayati and Montanari [24].

Finally, we need to estimate the effective Gaussian noise σ_t^2 in each AMP iteration. The estimated noise variance $\widehat{\sigma}_t^2$ can be calculated as [25]:

$$\widehat{\sigma}_t^2 = \frac{1}{M}\sum_{i=1}^M (r_i^t)^2,$$

where $\mathbf{r}^t$ is defined in (4).

III. Wavelet-Based Image Denoising within AMP

In this section, we describe how wavelet-based image denoisers are applied within AMP, and outline the adaptive Wiener filter image denoiser, which was proposed by Mıhçak et al. [18].

A. Wavelet transforms in AMP

In image processing, one often computes the wavelet coefficients [26] of images, applies some signal processing technique to the wavelet coefficients, and finally applies the inverse wavelet transform to the processed coefficients to obtain processed images. We now show how image denoising can be performed within AMP in the wavelet domain. Let us denote the wavelet transform by $\mathcal{W}$ and the inverse wavelet transform by $\mathcal{W}^{-1}$. By applying the wavelet transform to a vectorized image signal $\mathbf{x}$ (a 2-dimensional wavelet transform is used), we obtain the wavelet coefficient vector $\theta_{\mathbf{x}} = \mathcal{W}\mathbf{x}$. Conversely, $\mathbf{x} = \mathcal{W}^{-1}\theta_{\mathbf{x}}$. Therefore, the matrix channel (1) becomes $\mathbf{y} = \mathbf{A}\mathcal{W}^{-1}\theta_{\mathbf{x}} + \mathbf{z}$, where $\mathbf{A}\mathcal{W}^{-1}$ can be regarded as a new matrix in the matrix channel (1) and $\theta_{\mathbf{x}}$ as the corresponding input signal.

Let us express the AMP iterations (3, 4) for settings where the matrix is $\mathbf{A}\mathcal{W}^{-1}$,

$$\begin{aligned}\theta_{\mathbf{x}}^{t+1} &= \eta_t((\mathbf{A}\mathcal{W}^{-1})^T\mathbf{r}^t + \theta_{\mathbf{x}}^t), &(6)\\ \mathbf{r}^t &= \mathbf{y} - (\mathbf{A}\mathcal{W}^{-1})\theta_{\mathbf{x}}^t \\ &\quad + \frac{1}{R}\mathbf{r}^{t-1}\langle\eta'_{t-1}((\mathbf{A}\mathcal{W}^{-1})^T\mathbf{r}^{t-1} + \theta_{\mathbf{x}}^{t-1})\rangle \\ &= \mathbf{y} - \mathbf{A}\mathbf{x}^t \\ &\quad + \frac{1}{R}\mathbf{r}^{t-1}\langle\eta'_{t-1}((\mathbf{A}\mathcal{W}^{-1})^T\mathbf{r}^{t-1} + \theta_{\mathbf{x}}^{t-1})\rangle. &(7)\end{aligned}$$

Because the wavelet transform matrix is orthogonal, i.e., $\mathcal{W}\mathcal{W}^T = \mathbf{I} = \mathcal{W}\mathcal{W}^{-1}$, it can be shown that $(\mathbf{A}\mathcal{W}^{-1})^T = \mathcal{W}\mathbf{A}^T$. Therefore, the input of the denoiser $\eta_t(\cdot)$ (6) becomes $(\mathbf{A}\mathcal{W}^{-1})^T\mathbf{r}^t + \theta_{\mathbf{x}}^t = \mathcal{W}\mathbf{q}^t$, where $\mathbf{q}^t$ (5) is the noisy image at

iteration t, and $\mathcal{W}\mathbf{q}^t$ is the wavelet transform applied to the noisy image.

With the above analysis of the modified AMP (6, 7), we formulate a compressive imaging procedure as follows. Let us denote the the wavelet transform of the scalar channel (5) by

$$\theta^t_{\mathbf{q}} = \theta_{\mathbf{x}} + \theta^t_{\mathbf{v}}, \tag{8}$$

where $\theta^t_{\mathbf{q}} = \mathcal{W}\mathbf{q}^t$, $\theta_{\mathbf{x}} = \mathcal{W}\mathbf{x}$, and $\theta^t_{\mathbf{v}} = \mathcal{W}\mathbf{v}^t$. First, $\mathbf{r}^t$ and $\theta^t_{\mathbf{x}}$ are initialized to all-zero vectors. Then, at iteration t the algorithm proceeds as follows,

1) Calculate the residual term $\mathbf{r}^t$.
2) Calculate the noisy image $\mathbf{q}^t = (\mathbf{A}\mathcal{W}^{-1})^T\mathbf{r}^t + \theta^t_{\mathbf{x}}$, and apply the wavelet transform $\mathcal{W}$ to the noisy image $\mathbf{q}^t$ to obtain wavelet coefficients $\theta^t_{\mathbf{q}}$, which are the inputs of the scalar denoiser $\eta_t(\cdot)$ in (6).
3) Apply the denoiser $\eta_t(\cdot)$ to the wavelet coefficients $\theta^t_{\mathbf{q}}$, and obtain denoised coefficients $\theta^{t+1}_{\mathbf{x}}$.
4) Apply the inverse wavelet transform $\mathcal{W}^{-1}$ to the coefficients $\theta^{t+1}_{\mathbf{x}}$ to obtain the estimated image $\mathbf{x}^{t+1}$, which is used to compute the residual term in the next iteration.

B. Adaptive Wiener filter

We choose to denoise the wavelet coefficients using a scalar denoiser proposed by Mıhçak et al. [18], because this denoiser is simple to implement while producing promising numerical results (see Section IV).

Mıhçak et al. [18] proposed a method to estimate the variances of the wavelet coefficients, and then apply the corresponding Wiener filter to each wavelet coefficient. The variance of the noisy wavelet coefficient $\theta^t_{\mathbf{q},i}$ is estimated from its neighboring coefficients. More specifically, a set of 3×3 or 5×5 neighboring coefficients $\mathcal{N}_i$ that is centered at $\theta^t_{\mathbf{q},i}$ is considered, and the variance of $\theta^t_{\mathbf{q},i}$ is estimated by averaging the sum of $(\theta^t_{\mathbf{q},k})^2$ where $k \in \mathcal{N}_i$. This method of averaging the neighboring coefficients can be regarded as first convolving a 3×3 or 5×5 mask of all 1's with the matrix of squared wavelet coefficients $\theta^t_{\mathbf{q}}$, and then dividing by the normalizing constant 9 (for a 3×3 mask) or 25 (for a 5×5 mask). Other masks can be applied to produce different and possibly better denoising results. For example, we have found that the mask

$$\begin{array}{ccccc} & 1 & 1 & 1 & \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ & 1 & 1 & 1 & \end{array}$$

obtains lower mean square error (MSE) than the 5×5 mask of all 1's. Recall the scalar channel defined in (8) where the noise variance is $\widehat{\sigma}^2_t$, we estimate the variance of a noisy wavelet coefficient $\theta^t_{\mathbf{q},i}$ by $\widehat{\sigma}^2_i$, and the variance of the true wavelet coefficient $\theta^t_{\mathbf{x},i}$ by $\widehat{\sigma}^2_i - \sigma^2_t$.[1] Therefore, the adaptive Wiener filter being used as the denoising function can be expressed as follows,

$$\theta^{t+1}_{\mathbf{x},i} = \eta_t(\theta^t_{\mathbf{q},i}) = \frac{\widehat{\sigma}^2_i - \sigma^2_t}{\widehat{\sigma}^2_i}\theta^t_{\mathbf{q},i}. \tag{9}$$

[1]We use $\max\{\widehat{\sigma}^2_i - \sigma^2_t, 0\}$ to restrict the variance to be non-negative.

Finally, the derivative of this denoising function with respect to $\theta^t_{\mathbf{q},i}$ is simply the scaling factor $\frac{\widehat{\sigma}^2_i - \sigma^2_t}{\widehat{\sigma}^2_i}$ of the Wiener filter, and so the derivative term $\langle \eta'_{t-1}(\cdot) \rangle$ within the AMP iteration (4) can be obtained efficiently.

In standard AMP [16], the denoising function $\eta_t(\cdot)$ is separable, meaning that $\theta^{t+1}_{\mathbf{x},i}$ only depends on its corresponding noisy wavelet coefficient $\theta^t_{\mathbf{q},i}$. In this adaptive Wiener filter, however, the estimated variance $\widehat{\sigma}^2_i$ of each noisy wavelet coefficient depends on the neighboring coefficients of $\theta^t_{\mathbf{q},i}$, and so the denoising function in (9) implicitly depends on the neighboring coefficients of $\theta^t_{\mathbf{q},i}$. Therefore, the adaptive Wiener filter in (9) is not a strictly separable denoising function.

IV. Numerical Results

Having described the AMP algorithm [16] and the image denoiser [18], in this section we present the numerical results of applying this image denoiser within AMP.

We call the algorithm where the adaptive Wiener filter [18] is utilized "AMP-Wiener." Let us compare AMP-Wiener with three prior art compressive imaging algorithms, (*i*) Turbo-BG proposed by Som and Schniter [12]; (*ii*) Turbo-GM, also by Som and Schniter [12]; and (*iii*) a Markov chain Monte Carlo (MCMC) method by He and Carin [13]. Both Turbo-BG and Turbo-GM are message passing based algorithms. However, these two algorithms are more complex than AMP-Wiener, because they include two message passing procedures; the first procedure solves for dependencies between the wavelet coefficients and the second procedure is AMP. The performance metrics that we use to compare the algorithms are runtime and normalized MSE (NMSE), $\|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2/\|\mathbf{x}\|_2^2$, where $\widehat{\mathbf{x}}$ is the estimate of the vectorized input image $\mathbf{x}$.

Let us begin by contrasting the three prior art compressive imaging algorithms. Turbo-BG and Turbo-GM have similar runtimes; the NMSE of Turbo-GM is typically 0.5 dB better (lower) than the NMSE of Turbo-BG. At the same time, the NMSE of the MCMC algorithm [13] is comparable to those of Turbo-BG and Turbo-GM, but MCMC is 30 times slower than the Turbo approaches of Som and Schniter [12]. Other algorithms have also been considered for compressive imaging. For example, compressive sampling matching pursuit (CoSaMP) [27] requires only half the runtime of Turbo-GM, but its NMSE is roughly 4 dB worse than that of Turbo-GM; and model-based CS [28] is twice slower than Turbo-GM and its NMSE is also roughly 4 dB worse. Therefore, we provide numerical results for Turbo-BG, Turbo-GM, MCMC, and our proposed AMP-Wiener approach.

We have performed simulations for two numerical settings.

Numerical setting 1: We downloaded 591 images from "pixel-wise labeled image database v2" at http://research.microsoft.com/en-us/projects/objectclassrecognition. For each image, a 192×192 patch is extracted from the upper left corner of each image, and then the patch is resized to 128×128. This resizing approach was used in Som and Schniter [12]; we have verified that similar results can be obtained for 192×192 patches without resizing, and omit the details for brevity.

The measurement matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ is generated with i.i.d. Gaussian entries distributed as $\mathcal{N}(0, \frac{1}{M})$; each column is

Algorithm	NMSE (dB)	runtime (sec)
Turbo-BG [12]	-20.37	12.39
Turbo-GM [12]	-20.72	12.47
MCMC [13]	-20.31	423.15
AMP-Wiener	-21.00	3.34

TABLE I: *NMSE and runtime averaged over 591 images patches: a* 192×192 *patch from the upper left corner of each image is first extracted, and then resized to* 128×128*. The number of measurements* $M = 5,000$*, and the measurements are noiseless.*

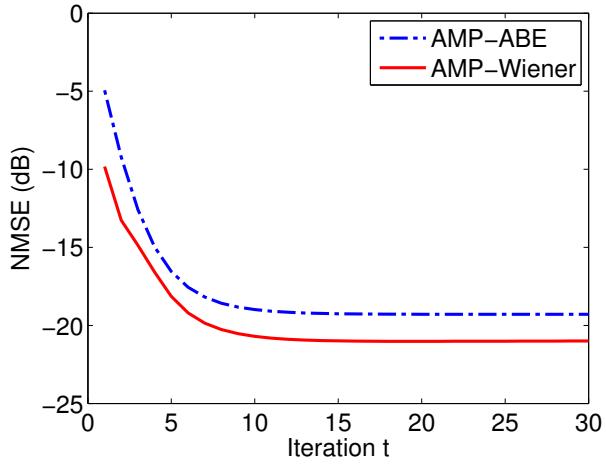

Figure 1: *Average NMSE over 591 images from AMP iteration 1 to iteration 30.*

then normalized to have unit norm. The number of measurements $M = 5,000$, and the measurements $\mathbf{y}$ are noiseless, i.e., $\mathbf{y} = \mathbf{Ax}$. Note that the measurement rate is $R = M/N = 5000/(128 \cdot 128) \approx 0.3$, which offers a reasonable degree of compression in the sensing process. Finally, we set the number of AMP iterations to be 30.

Result 1: Table I shows the NMSE and runtime averaged over the 591 image patches.

It can be seen from Table I that the NMSE of AMP-Wiener is the best (lowest) among all the algorithms compared. At the same time, AMP-Wiener runs approximately 3.5 times faster than the Turbo approaches of Som and Schniter [12], and 120 times faster than MCMC [13].

Figure 1 complements Table I by plotting the average NMSE over 591 images from iteration 1 to iteration 30. The horizontal axis represents iteration numbers, and the vertical axis represents NMSE. The NMSE drops markedly from -10 dB to -21 dB . Note that the average NMSE is approximately -21 dB around iteration 15, which suggests that we could halve the runtime of AMP-Wiener (to approximately 1.7 seconds) by reducing the number of AMP iterations from 30 to 15.

Numerical setting 2: We also evaluate the performance of each algorithm by plotting the NMSE (average NMSE over 591 images) versus the measurement rate $R = M/N$, where R varies from 0.1 to 1. The measurement matrix $\mathbf{A}$ is generated the same way as in numerical setting 1.

Result 2: Figure 2 illustrates how the NMSEs achieved by AMP-Wiener and Turbo-GM vary when the measurement

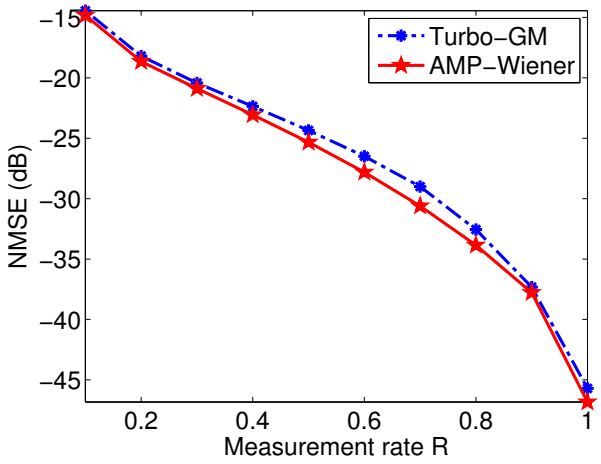

Figure 2: *Average NMSE over 591 images versus measurement rate. Image patches are extracted as follows: a* 192×192 *patch from the upper left corner of each image is first extracted, and then resized to* 128×128*.*

rate R changes from 0.1 to 1, where the horizontal axis represents the measurement rate $R = M/N$, and the vertical axis represents NMSE. It is shown in Figure 2 that AMP-Wiener achieves lower NMSE than that of Turbo-GM for all values of R.

V. Discussion

In this paper, we proposed compressive imaging algorithms that apply image denoisers within approximate message passing (AMP). Specifically, we used an adaptive Wiener filter proposed by Mıhçak et al. [18] as the image denoiser within AMP, which we call AMP-Wiener. Numerical results showed that AMP-Wiener achieves the lowest reconstruction error among all competing algorithms in all simulation settings. Moreover, the runtime of AMP-Wiener is significantly lower than those of MCMC [13] and Turbo approaches [12], and Figure 1 suggests that the runtime of AMP-Wiener could be reduced further if we accept a slight deterioration in NMSE.

Besides the adaptive Wiener filter, we have also simulated AMP with an image denoiser called "amplitude-scale-invariant Bayes estimator" (ABE), which was proposed by Figueiredo and Nowak [20]. The results of this AMP-ABE are comparable to those of AMP-Wiener and thus omitted for brevity. Because we have seen that different image denoisers work well within AMP in terms of both reconstruction quality and runtime, we believe that AMP with other image denoisers such as SBSDI [23] or BM3D [21] should also produce promising results for compressive imaging problems, provided that the derivative term $\langle \eta'_{t-1}(\cdot) \rangle$ within the AMP iteration (4) can be obtained.

Acknowledgments

We thank Liyi Dai, Nikhil Krishnan, and Junan Zhu for useful discussions, which greatly helped us improve this manuscript.

REFERENCES

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

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

[3] D. Takhar, J. Laska, M. Wakin, M. Duarte, D. Baron, S. Sarvotham, K. Kelly, and R. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," *IS&T/SPIE Computational Imaging IV*, vol. 6065, Jan. 2006.

[4] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstruction," in *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, vol. 3, Mar. 2006, pp. 872–875.

[5] R. G. Baraniuk, "A lecture on compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, July 2007.

[6] R. Blahut, *Theory of remote image formation*. Cambridge University Press, 2004.

[7] F. Natterer and F. Wubbeling, *Mathematical methods in image reconstruction*. Society for Industrial Mathematics, 2001.

[8] Z. P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. SPIE Optical Engineering Press, 2000.

[9] J. F. Claerbout, *Imaging the earth's interior*. Blackwell Scientific Publications, 1985.

[10] A. Rajwade, D. Kittle, T. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," *SIAM J. on Imag. Sci.*, vol. 6, no. 2, pp. 782–812, Apr. 2013.

[11] R. Willett, M. Duarte, M. Davenport, and R. Baraniuk, "Sparsity and structure in hyperspectral imaging: Sensing, reconstruction, and target detection," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 116–126, Jan. 2014.

[12] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3439–3448, July 2012.

[13] L. He and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3488–3497, Sept. 2009.

[14] A. Soni and J. Haupt, "Learning sparse representations for adaptive compressive sensing," in *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Mar. 2012, pp. 2097–2100.

[15] R. Jenatton, J. Mairal, F. R. Bach, and G. R. Obozinski, "Proximal methods for sparse hierarchical dictionary learning," in *Proc. 27th Int. Conf. Mach. Learning*, June 2010, pp. 487–494.

[16] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[17] D. L. Donoho and J. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sept. 1994.

[18] M. K. Mıhçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Process. Letters*, vol. 6, no. 12, pp. 300–303, Dec. 1999.

[19] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 909–919, Apr. 1999.

[20] M. Figueiredo and R. Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1322–1331, Sept. 2001.

[21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[22] M. Zhou, H. Chen, J. Paisley, L. Ren, L. Li, Z. Xing, D. Dunson, G. Sapiro, and L. Carin, "Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 130–144, Jan. 2012.

[23] L. Fang, S. Li, R. McNabb, Q. Nie, A. Kuo, C. Toth, J. A. Izatt, and S. Farsiu, "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," *IEEE Trans. Med. Imag.*, vol. 32, no. 11, pp. 2034–2049, Nov. 2013.

[24] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.

[25] A. Montanari, "Graphical models concepts in compressed sensing," *Compressed Sensing: Theory and Applications*, pp. 394–438, 2012.

[26] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999.

[27] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harm. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

[28] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, April 2010.