AN APPROXIMATE MESSAGE PASSING ALGORITHM FOR RAPID PARAMETER-FREE COMPRESSED SENSING MRI

Charles Millard[⋆†‡]*, Aaron T Hess*[†]*, Boris Mailhe*[‡] *and Jared Tanner*[⋆]

[⋆]Mathematical Institute, University of Oxford, UK
[†]Oxford Centre for Clinical Magnetic Resonance, University of Oxford, UK
[‡]Siemens Healthineers, Princeton, NJ, USA

ABSTRACT

For certain sensing matrices, the Approximate Message Passing (AMP) algorithm efficiently reconstructs undersampled signals. However, in Magnetic Resonance Imaging (MRI), where Fourier coefficients of a natural image are sampled with variable density, AMP encounters convergence problems. In response we present an algorithm based on Orthogonal AMP constructed specifically for variable density partial Fourier sensing matrices. For the first time in this setting a state evolution has been observed. A practical advantage of state evolution is that Stein's Unbiased Risk Estimate (SURE) can be effectively implemented, yielding an algorithm with no free parameters. We empirically evaluate the effectiveness of the parameter-free algorithm on simulated data and find that it converges over 5x faster and to a lower mean-squared error solution than Fast Iterative Shrinkage-Thresholding (FISTA).

Index Terms— Approximate Message Passing, Compressed Sensing, Stein's Unbiased Risk Estimate, Magnetic Resonance Imaging

1. INTRODUCTION

We consider a complex linear regression problem, where complex data vector $\boldsymbol{y} \in \mathbb{C}^n$ is formed of noisy linear measurements of a signal of interest $\boldsymbol{x}_0 \in \mathbb{C}^N$:

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}_0 + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{n\times N}$ and $\boldsymbol{\varepsilon} \backsim \mathcal{CN}(\boldsymbol{0}, \sigma_\varepsilon^2 \mathbb{1}_n)$, where $\mathbb{1}_n$ is the $n \times n$ identity matrix. Here, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}^2)$ denotes the complex normal distribution with mean $\boldsymbol{\mu}$, covariance $\boldsymbol{\Sigma}^2$ and white phase. A well-studied approach is to seek a solution of

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{C}^N}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + f(\boldsymbol{x}) \tag{2}$$

where $f(\boldsymbol{x})$ is a model-based penalty function. Compressed sensing [1, 2] concerns the reconstruction of $\boldsymbol{x}_0$ from underdetermined measurements $n < N$. Commonly sparsity in $\hat{\boldsymbol{x}}$ is promoted by solving (2) with $f(\boldsymbol{x}) = \lambda\|\boldsymbol{\Psi x}\|_1$ for sparse weighting $\lambda > 0$ and sparsifying transform $\boldsymbol{\Psi}$.

The Approximate Message Passing (AMP) algorithm [3] is an iterative method that estimates $\boldsymbol{x}_0$ in linear regression problems such as (1). At iteration k AMP implements a denoiser $\boldsymbol{g}(\boldsymbol{r}_k; \tau_k)$ on $\boldsymbol{x}_0$ estimate $\boldsymbol{r}_k$ with mean-squared error estimate τ_k. For instance, for problems of the form of (2), $\boldsymbol{g}(\boldsymbol{r}_k; \tau_k)$ is the proximal operator associated with $f(\boldsymbol{x})$:

$$\boldsymbol{g}(\boldsymbol{r}_k; \tau_k) = \underset{\boldsymbol{x}\in\mathbb{C}^N}{\operatorname{argmin}} \frac{1}{2\tau_k}\|\boldsymbol{r}_k - \boldsymbol{x}\|_2^2 + f(\boldsymbol{x}), \tag{3}$$

which is equal to soft thresholding in the case of $f(\boldsymbol{x}) = \lambda\|\boldsymbol{\Psi x}\|_1$ and orthogonal $\boldsymbol{\Psi}$. For certain sensing matrices and given mild conditions on $f(\boldsymbol{x})$, AMP's state evolution guarantees that in the large system limit $n, N \to \infty$, $n/N \to \delta \in (0, 1)$, vector $\boldsymbol{r}_k$ behaves as the original signal corrupted by zero-mean white Gaussian noise:

$$\boldsymbol{r}_k = \boldsymbol{x}_0 + \mathcal{CN}(\boldsymbol{0}, \sigma_k^2 \mathbb{1}_N) \tag{4}$$

where σ_k is an iteration-dependant standard deviation. The state evolution of AMP has been proven for real i.i.d. Gaussian measurements in [4] and i.i.d. sub-Gaussian measurements in [5]. It has also been empirically shown that state evolution holds for uniformly undersampled Fourier measurements of a random artificial signal [3]. When state evolution holds, AMP is known to exhibit very fast convergence. However, for generic $\boldsymbol{\Phi}$, the behavior of AMP is not well understood and it has been noted by a number of authors [6–8] that it can encounter convergence problems. The recent Vector Approximate Message Passing (VAMP) [9] algorithm and related Orthogonal AMP (OAMP) [10] obey (4) for a broader class of measurement matrices $\boldsymbol{\Phi}$, and were found to perform very well on certain reconstruction tasks. For VAMP, state evolution was proven for sensing matrices that are 'right-orthogonally invariant': see [9] for details.

In compressed sensing MRI [11], measurements are formed of undersampled Fourier coefficients, so that $\boldsymbol{\Phi}$ =

This research was supported by an EPSRC Industrial CASE studentship with Siemens Healthineers, voucher number 17000051, and through the Alan Turing Institute under the EPSRC grant EP/N510129/1. The concepts and information presented in this paper are based on research results that are not commercially available.

$M_\Omega F$, where F is a 2D or 3D discrete Fourier transform and $M_\Omega \in \mathbb{R}^{n\times N}$ is a undersampling mask that selects the jth row of F if $j \in \Omega$ for sampling set Ω. The signal of interest x_0 is a natural image, so typically has a highly non-uniform spectral density that is concentrated at low frequencies. Accordingly, the sampling set Ω is usually generated such that there is a higher probability of sampling lower frequencies. This work considers an Ω with elements drawn independently from a Bernoulli distribution with non-uniform probability, such that $\mathrm{Prob}(j \in \Omega) = p_j \in [0, 1]$. In this variable density setting there are no theoretical guarantees for AMP, VAMP and OAMP and in practice the behavior of (4) is not observed and the algorithms typically perform poorly. VAMP for Image Recovery (VAMPire) [12] is an adaption of VAMP for variable density Fourier sampled images that "whitens" the signal in the wavelet domain with a hand-tuned prediction of the wavelet energy, but like the aforementioned algorithms does not obey (4), as discussed in Section 2.

Herein we present a method for undersampled signal reconstruction based on OAMP [10] which we term the Variable Density Approximate Message Passing (VDAMP) algorithm. For Fourier coefficients of a realistic image randomly sampled with generic variable density and orthogonal wavelet Ψ we have empirical evidence that a state evolution occurs. Unlike the white effective noise of (4), the r_k of VDAMP behaves as the ground truth corrupted by zero-mean Gaussian noise with a separate variance for each wavelet subband, such that

$$r_k = w_0 + \mathcal{CN}(\mathbf{0}, \Sigma_k^2), \tag{5}$$

where $w_0 := \Psi x_0$ and the effective noise covariance Σ_k^2 is diagonal so that for a Ψ with s decomposition scales

$$\Sigma_k^2 = \begin{bmatrix} \sigma_{k,1}^2 \mathbb{1}_{N_1} & 0 & \dots & 0 \\ 0 & \sigma_{k,2}^2 \mathbb{1}_{N_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{k,1+3s}^2 \mathbb{1}_{N_{1+3s}} \end{bmatrix}, \tag{6}$$

where $\sigma_{k,j}^2$ and N_j refer to the variance and dimension of the jth subband respectively. We refer to (5) as the *colored* state evolution of VDAMP.

Selecting appropriate regularisation parameters such as λ is a notable challenge in real-world compressed sensing MRI applications. We present an approach to parameter-free compressed sensing reconstruction using Stein's Unbiased Risk Estimate (SURE) [13] in conjunction with VDAMP, building on the work on AMP introduced in [14]. A strength of automatic parameter tuning via SURE is that the it is possible to have a richer regularizer than would usually be feasible for a hand-tuned $f(x)$ [15]. We implement a variation on the SureShrink denoiser [16], using a iteration-dependant regularizer that has a separate sparse weighting per subband:

$$\lambda_k = \begin{bmatrix} \lambda_{k,1} \mathbf{1}_{N_1}^H & \lambda_{k,2} \mathbf{1}_{N_2}^H & \dots & \lambda_{k,1+3s} \mathbf{1}_{N_{1+3s}}^H \end{bmatrix}^H, \tag{7}$$

Algorithm 1 VDAMP

Require: Sensing matrix Φ, orthogonal wavelet Ψ, probability matrix P, measurements y, denoiser $g(r_k; \tau_k)$ and number of iterations K_{it}.

1: Set $\widetilde{r}_0 = \mathbf{0}$ and compute $S = |\Phi \Psi^H|^2$
2: **for** $k = 0, 1, \dots, K_{it} - 1$ **do**
3: $\quad z_k = y - \Phi \Psi^H \widetilde{r}_k$
4: $\quad r_k = \widetilde{r}_k + \Psi \Phi^H P^{-1} z_k$
5: $\quad \tau_k = S^H P^{-1} [(P^{-1} - \mathbb{1}_n)|z_k|^2 + \sigma_\varepsilon^2 \mathbf{1}_n]$
6: $\quad \hat{w}_k = g(r_k; \tau_k)$
7: $\quad \alpha_k = \langle g'(r_k; \tau_k) \rangle_{\text{sband}}$
8: $\quad \widetilde{r}_{k+1} = (\hat{w}_k - \alpha_k \odot r_k) \oslash (\mathbf{1} - \alpha_k)$
9: **end for**
10: **return** $\hat{x} = \Psi^H w_k + \Phi^H (y - \Phi \Psi^H w_k)$

where $\mathbf{1}_M$ is the M-dimensional column vector of ones. SURE has previously been employed for parameter-free compressed sensing MRI in [17], where the Fast Iterative Shrinking-Thresholding Algorithm (FISTA) algorithm [18] was used with SureShrink in place of the usual shrinkage step. This algorithm is herein referred to as SURE-IT. The effective noise of SURE-IT is highly non-Gaussian, so deviates from a proper theoretical basis for using SURE for threshold selection. To our knowledge, VDAMP is the first algorithm for variable density Fourier sampling of natural images where a state evolution has been observed.

2. DESCRIPTION OF ALGORITHM

For AMP, VAMP and OAMP, (4) states that the effective noise $r_k - x_0$ is white, so can be fully characterised by a scalar τ_k. This is appropriate for the kind of uniform measurement matrices and separable, identical sparse signal models $f(x)$ that are often encountered in abstract compressed sensing problems. However, when Fourier coefficients of a natural image are sampled the effective noise is colored, so is poorly represented by a scalar [19]. While VAMPire attempts to whiten the effective noise, we propose allowing the effective noise to be colored, and present a method for modelling the color with a vector τ_k that has one real number per wavelet subband.

The VDAMP algorithm is shown in Algorithm 1. Here, $P \in \mathbb{R}^{n\times n}$ is the diagonal matrix formed of sampling probabilities p_j for $j \in \Omega$. The function $g(r_k; \tau_k)$ refers to some denoiser with a colored effective noise model. The notation $\langle g'(r_k; \tau_k) \rangle_{\text{sband}}$ in line 7 refers to the (sub)-gradient of the denoiser averaged over subbands, so that for s decomposition scales α_k has $1+3s$ unique entries, having the same structure as the λ_k of (7). In line 8, the notation $\odot$ is used for entry-wise multiplication and $\oslash$ for entry-wise division. $|\cdot|$ refers to the entry-wise absolute magnitude of a vector or matrix.

To ensure that r_k is an unbiased estimate of x_0, the sensing matrix must be correctly normalized. In VDAMP this is manifest in the gradient step of lines 3-4, which features a

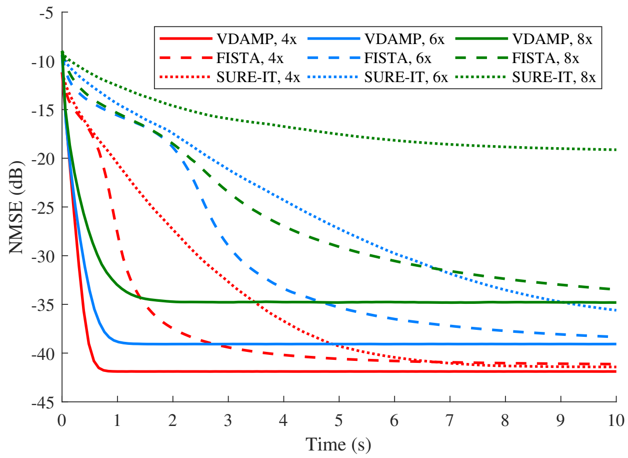
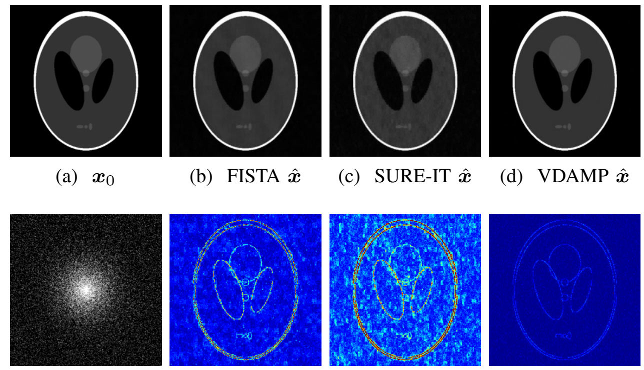


Fig. 1. NMSE of VDAMP, FISTA and SURE-IT versus compute time for undersampling factors 4, 6 and 8.



(a) $\boldsymbol{x}_0$ (b) FISTA $\hat{\boldsymbol{x}}$ (c) SURE-IT $\hat{\boldsymbol{x}}$ (d) VDAMP $\hat{\boldsymbol{x}}$

(e) Ω (f) FISTA $|\hat{\boldsymbol{x}} - \boldsymbol{x}_0|$ (g) SURE-IT $|\hat{\boldsymbol{x}} - \boldsymbol{x}_0|$ (h) VDAMP $|\hat{\boldsymbol{x}} - \boldsymbol{x}_0|$

Fig. 2. Reconstruction of a 8x undersampled Shepp-Logan after 2 seconds with (b) FISTA (NMSE = -19.3dB) (c) SURE-IT (NMSE = -16.6dB) and (d) VDAMP (NMSE -34.9dB), where absolute values are shown.

crucial weighting by $\boldsymbol{P}^{-1}$. As an intuitive example, consider $k = 0$, where $\boldsymbol{r}_0 = \boldsymbol{\Psi}\boldsymbol{\Phi}^H \boldsymbol{P}^{-1}\boldsymbol{y}$. Since $\mathbb{E}_{\Omega,\varepsilon}(\boldsymbol{P}^{-1}\boldsymbol{y}) = \boldsymbol{y}_0$, it follows that $\boldsymbol{r}_0$ is unbiased: $\mathbb{E}(\boldsymbol{r}_0) = \boldsymbol{w}_0$. Such a rescaling is referred to as "density compensation" in the MRI literature [20], and was used in the original compressed sensing MRI paper with zero-filling to generate a unregularized, non-iterative baseline [11]. VDAMP's density compensation contrasts with VAMPire's approach, where the sensing matrix is right-multiplied by a diagonal matrix of predicted wavelet coefficient energies. This leads to a biased $\boldsymbol{r}_k$ update, $\mathbb{E}(\boldsymbol{r}_k) \neq \boldsymbol{w}_0$, and consequently it does not obey a state evolution.

Line 5 of Algorithm 1 computes an estimate of the colored effective noise variance $|\boldsymbol{w}_0 - \boldsymbol{r}_k|^2$. The mean-squared error estimate $\boldsymbol{\tau}_k$ is an unbiased estimate of the expected entry-wise squared error, $\mathbb{E}_{\Omega,\boldsymbol{\varepsilon}}(\boldsymbol{\tau}_k) = \mathbb{E}_{\Omega,\boldsymbol{\varepsilon}}(|\boldsymbol{w}_0 - \boldsymbol{r}_k|^2)$, for any $\widetilde{\boldsymbol{r}}_k$. We assume this estimator concentrates around its expectation, and leave the study of the constraints this imposes on $\boldsymbol{P}$ for future works. Note that $\boldsymbol{S}$ has $1 + 3s$ unique columns, so for fixed s the complexity of VDAMP is governed by $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$, whose fast implementations have complexity $O(NlogN)$. Lines 6-8 are the model-based regularization phase from OAMP and VAMP, but with a colored effective noise model. Line 10 implements an unweighted gradient step that enforces exact consistency of the image estimate with the measurements.

3. NUMERICAL EXPERIMENTS

In the experiments presented in this section the denoiser $\boldsymbol{g}(\boldsymbol{r}_k; \boldsymbol{\tau}_k)$ was the complex soft thresholding operator with an automatically tuned subband-dependant threshold. In other words, we used a complex, colored version of SURE to approximately solve

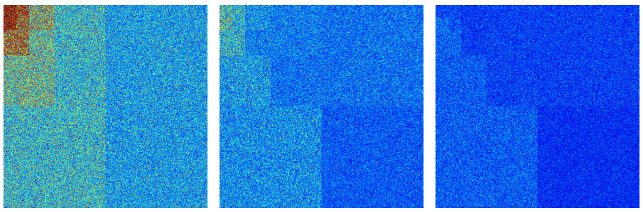
$$\boldsymbol{g}(\boldsymbol{r}_k; \boldsymbol{\tau}_k) \approx \underset{\boldsymbol{w}\in\mathbb{C}^N}{\operatorname{argmin}} \min_{\boldsymbol{\lambda}\in\mathbb{R}^N} \frac{1}{2}\|(\boldsymbol{w} - \boldsymbol{r}_k) \oslash \sqrt{\boldsymbol{\tau}_k}\|_2^2 + \|\boldsymbol{\lambda} \odot \boldsymbol{w}\|_1,$$

where $\sqrt{\boldsymbol{\tau}_k}$ is the entry-wise square root of $\boldsymbol{\tau}_k$ and $\boldsymbol{\lambda}$ is of the form of (7).

We considered the reconstruction of a 512×512 Shepp-Logan artificially corrupted with complex Gaussian noise with white phase so that $\|\boldsymbol{F}\boldsymbol{x}_0\|_2^2/N\sigma_\varepsilon^2 = 40$dB. We assumed that σ_ε^2 was known a priori; in practice it can be well estimated with an empty prescan. All sampling probabilities p_j were generated with polynomial variable density. We considered a Haar wavelet $\boldsymbol{\Psi}$ at $s = 4$ decomposition scales, which are referred to as scales 1-4, where scale 1 is the finest and scale 4 is the coarsest. All experiments were conducted in MATLAB 9.4 and can be reproduced with code available at https://github.com/charlesmillard/VDAMP.

VDAMP was compared with FISTA, which is commonly used for compressed sensing MRI in practise, and SURE-IT, which contrasts our use of SURE. We did not compare with VAMPire as no implementation was available. For FISTA we used a sparse weighting λ tuned with an exhaustive search so that the mean-squared error was minimised after 10 seconds. For SURE-IT the mean-squared error estimate was updated using the ground truth: $\tau_k = \|\boldsymbol{w}_0 - \boldsymbol{r}_k\|_2^2/N$, and (3) with $\boldsymbol{\tau}_k = \tau_k \boldsymbol{1}_N$ was implemented. All algorithms were initialised with a vector of zeros. Three sampling sets Ω were generated at undersampling factors of approximately 4, 6 and 8, and VDAMP, FISTA and SURE-IT were run for 10 seconds.

Fig. 1 shows the NMSE $= \|\hat{\boldsymbol{x}} - \boldsymbol{x}_0\|_2^2/\|\boldsymbol{x}_0\|_2^2$ as a function of time for each algorithm. The mean per-iteration compute time was 0.065s for FISTA, 0.077s for SURE-IT, and 0.091s for VDAMP. Fig. 2 shows the ground truth image, sampling set, and FISTA and VDAMP reconstruction at undersampling factor 8 after 2s, with entry-wise errors $|\hat{\boldsymbol{x}} - \boldsymbol{x}_0|$. In Fig. 3, the absolute value of the residual of $\boldsymbol{r}_k$ is shown for three representative iterations: $k = 0$, $k = 1$ and $k = 2$ for undersampling factor 8. For the same iterations, Fig. 4 shows



(a) $|\boldsymbol{r}_0 - \boldsymbol{w}_0|$ (b) $|\boldsymbol{r}_1 - \boldsymbol{w}_0|$ (c) $|\boldsymbol{r}_2 - \boldsymbol{w}_0|$

Fig. 3. $|\boldsymbol{r}_k - \boldsymbol{w}_0|$ of VDAMP for $k = 0$, $k = 1$ and $k = 2$.

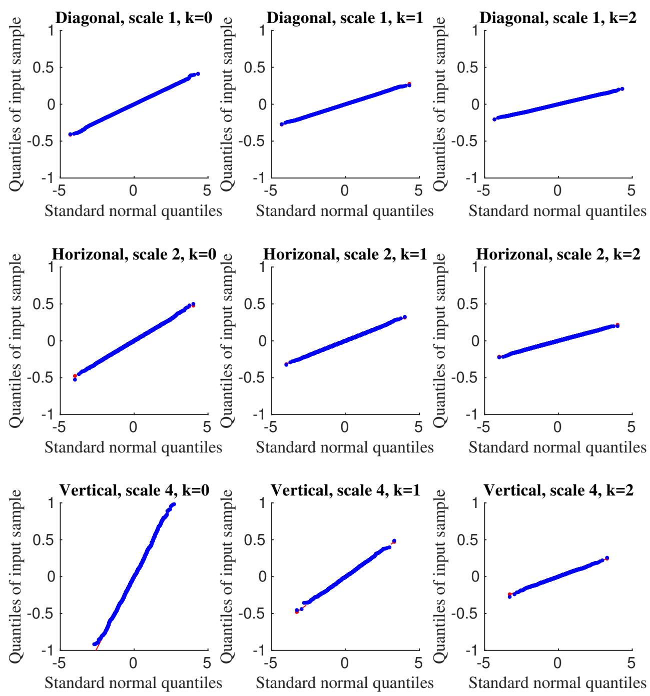


Fig. 4. Normalized quantile-quantile plots against a Gaussian of three subbands of $\boldsymbol{r}_k - \boldsymbol{w}_0$ for $k = 0$, $k = 1$ and $k = 2$ in blue, and points along a straight line in red. The real part is plotted in the top and bottom rows and the imaginary is plotted in the middle row. Linearity of the blue points indicates that that the data comes from a Gaussian distribution. Finite dimensional effects causing small deviations from an exact Gaussian are more apparent at coarse scales, where the dimension is smaller.

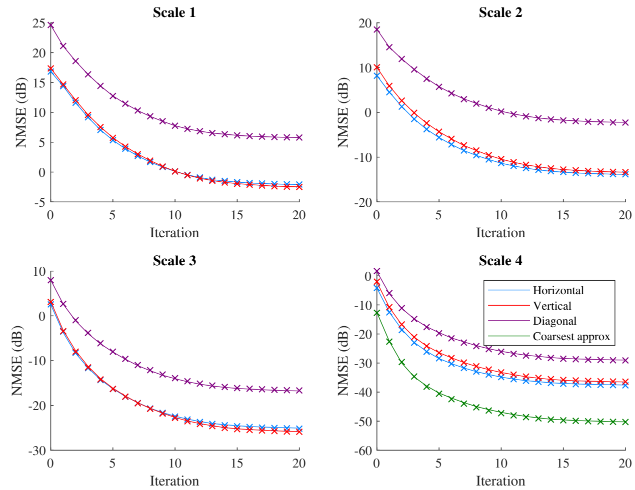


Fig. 5. NMSE versus iteration number of $\boldsymbol{r}_k$ for each subband. Lines show the actual NMSE and crosses show the predictions from $\boldsymbol{\tau}_k$.

quantile-quantile plots of the real parts of three illustrative subbands of $\boldsymbol{r}_k - \boldsymbol{w}_0$: the diagonal detail at scale 1, the horizontal detail at scale 2 and the vertical detail at scale 4. These figures provide empirical evidence that the effective noise of VDAMP evolves as (5).

The efficacy of automatic threshold selection with SURE depends on how accurately the diagonal of $\boldsymbol{\Sigma}_k^2$ from (6) is modelled by $\boldsymbol{\tau}_k$. For $k = 0, 1, \ldots, 20$, Fig. 5 shows the ground truth NMSE of the wavelet coefficients at all four scales and the prediction of VDAMP, where the NMSE is per subband.

4. CONCLUSIONS

VDAMP's state evolution provides an informed and efficient way to tune model parameters via SURE. More degrees of freedom are feasibly allowed in the model, enabling higher order prior information such as anisotropy, variability across scales and structured sparsity, without the need to estimate the structure a priori such as in model-based compressed sensing [21]. Theoretical work is required to establish the generality of the state evolution empirically observed in these experiments.

It is known that the state evolution of AMP holds for a wide range of denoisers $\boldsymbol{g}(\boldsymbol{r}_k; \tau_k)$ [22]. In a recent survey [23] a number of standard compressed sensing algorithms that leverage image denoisers designed for Gaussian noise were shown to perform well on MRI reconstruction tasks, despite the mismatch between the effective noise and its model. A sophisticated denoiser equipped to deal with wavelet coefficients corrupted with known colored Gaussian noise would be expected to perform well in conjunction with VDAMP.

5. REFERENCES

[1] David L Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 4 2006.

[2] Emmanuel Candes, Justin Romberg, and Terence Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2 2006.

[3] David L Donoho, Arian Maleki, and Andrea Montanari, "Message-passing algorithms for compressed sensing.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 45, pp. 18914–9, 11 2009.

[4] Mohsen Bayati and Andrea Montanari, "The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 764–785, 2 2011.

[5] Mohsen Bayati, Marc Lelarge, and Andrea Montanari, "Universality in polytope phase transitions and message passing algorithms," *The Annals of Applied Probability*, vol. 25, no. 2, pp. 753–822, 4 2015.

[6] Sundeep Rangan, Philip Schniter, and Alyson Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *2014 IEEE International Symposium on Information Theory*. 6 2014, pp. 236–240, IEEE.

[7] Francesco Caltagirone, Lenka Zdeborova, and Florent Krzakala, "On convergence of approximate message passing," in *2014 IEEE International Symposium on Information Theory*. 6 2014, pp. 1812–1816, IEEE.

[8] Sundeep Rangan, Philip Schniter, Erwin Riegler, Alyson K. Fletcher, and Volkan Cevher, "Fixed Points of Generalized Approximate Message Passing With Arbitrary Matrices," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7464–7474, 12 2016.

[9] Sundeep Rangan, Philip Schniter, and Alyson K. Fletcher, "Vector Approximate Message Passing," *IEEE Transactions on Information Theory*, pp. 1–1, 2019.

[10] Junjie Ma and Li Ping, "Orthogonal AMP," *IEEE Access*, vol. 5, pp. 2020–2033, 2017.

[11] Michael Lustig, David Donoho, and John M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 12 2007.

[12] Philip Schniter, Sundeep Rangan, and Alyson Fletcher, "Plug-and-play Image Recovery using Vector AMP," 2017, BASP Frontiers Workshop 2017.

[13] Charles M. Stein, "Estimation of the Mean of a Multivariate Normal Distribution," *The Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 11 1981.

[14] Ali Mousavi, Arian Maleki, and Richard G. Baraniuk, "Parameterless Optimal Approximate Message Passing," 10 2013.

[15] Charles Alban Deledalle, Samuel Vaiter, Jalal Fadili, and Gabriel Peyré, "Stein Unbiased GrAdient estimator of the Risk (SUGAR) for multiple parameter selection," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2448–2487, 2014.

[16] David L. Donoho and Iain M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, 12 1995.

[17] Kedar Khare, Christopher J. Hardy, Kevin F. King, Patrick A. Turski, and Luca Marinelli, "Accelerated MR imaging using compressive sensing with no free parameters," *Magnetic Resonance in Medicine*, vol. 68, no. 5, pp. 1450–1457, 11 2012.

[18] Amir Beck and Marc Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 1 2009.

[19] Patrick Virtue and Michael Lustig, "The Empirical Effect of Gaussian Noise in Undersampled MRI Reconstruction," *Tomography*, vol. 3, no. 4, pp. 211–221, 12 2017.

[20] James G. Pipe and Padmanabhan Menon, "Sampling density compensation in MRI: Rationale and an iterative numerical solution," *Magnetic Resonance in Medicine*, vol. 41, no. 1, pp. 179–186, 1 1999.

[21] Richard G. Baraniuk, Volkan Cevher, Marco F. Duarte, and Chinmay Hegde, "Model-Based Compressive Sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 4 2010.

[22] Christopher A. Metzler, Arian Maleki, and Richard G. Baraniuk, "From Denoising to Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5117–5144, 9 2016.

[23] Rizwan Ahmad, Charles A. Bouman, Gregery T. Buzzard, Stanley Chan, Edward T. Reehorst, and Philip Schniter, "Plug and play methods for magnetic resonance imaging," 3 2019.