

SAMPLE ALLOCATION FOR STATISTICAL MULTIREOLUTION COMPRESSED SENSING

Chunli Guo, Mike E. Davies

School of Engineering and Electronics, The University of Edinburgh

ABSTRACT

We model the compressible signal with the two states Gaussian mixture distribution and consider the sample distortion function for the recently proposed Bayesian optimal AMP decoder. By leveraging the rigorous analysis of the AMP algorithm, we are able to derive the theoretical SD function and a sample allocation scheme for multi-resolution statistical image model. We then adopt the "turbo" message passing method to integrate the bandwise sample allocation with the exploitation of the hidden Markov tree structure of wavelet coefficients. Experiments on natural image show that the combination outperforms either of them working alone.

Index Terms— Sample distortion function, Bayesian optimal AMP, Turbo decoding, Sample allocation

1. INTRODUCTION

There has recently been much attention on analysing the compressed sensing (CS) problem from the statistical perspective. In the stochastic setting, the signals are drawn from a probability distribution that is in some sense compressible. In many situations, it is reasonable to assume the statistical property is known. Such Bayesian compressed sensing framework has been considered in terms of signal estimation and measurement matrix design in [1], [2], [3]. In [4] we introduced the notion of sample distortion (SD) function and consider the SD function for the Gaussian encoder- ℓ_1 decoder. By extending the SD function to a multi-resolution statistical image model, we are able to derive the optimal bandwise sample allocation scheme and predict the approximation performance for a given undersampling ratio.

In this paper we adopted the same SD framework [4] and consider the SD function for the recently proposed approximate message passing (AMP) [3] decoder and Gaussian mixture (GM) distribution for the signal model. AMP admits a rigorous analysis in the large-limit system which naturally provides us with the theoretical SD function given the signal distribution [5]. We further modify the AMP decoder by exploiting the convexity property of the SD function and achieve a better SD curve lies below the one for the conventional AMP decoder. The convexified SD function enables us to perform a greedy sample allocation to optimize the bandwise sampling.

Because the marginal statistics of the wavelet coefficients are not significantly compressible [6], we should incorporate other dependencies to maximise the SD performance. The second part of the paper considers leveraging Som and Schniter's recently proposed TurboAMP [7] algorithm to exploit both the sparsity and the persistence across scale (PSA) property [8]. The advantage of modelling the wavelet coefficients with the GM distribution is that the hidden

states form a quad-tree structure [9]. As in [7], we use the "turbo" message passing scheme to alternate between the AMP decoding and tree structure decoding. But unlike their work, we use an optimized block diagonal sensing matrix derived from the theoretical SD function instead of distributing samples uniformly. By integrating the bandwise sampling and the turbo decoding scheme, we expect the accurate message acquired from the coarse-scale wavelet bands will propagate to the fine-scale bands through the hidden Markov tree HMT structure and thus enhance the reconstruction quality. The natural image simulation confirms the analytical improvement.

2. SAMPLE DISTORTION FRAMEWORK

Suppose $X \in \mathbb{R}^n$ is a realization of a random vector, $X := [x_1, \dots, x_n]^T$, *i.i.d* $\sim p(x)$. In compressed sensing, we observe the linear combination of the source $Y = \Phi X$ through the measurement matrix, or the encoder $\Phi \in \mathbb{R}^{m \times n}$, $m < n$. Let the decoder $\Delta : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be some Lipschitz mapping that estimates the original signal from the observation. The squared error distortion between X and $\Delta(\Phi X)$ at undersampling ratio δ is

$$D_\Delta(\delta) = \frac{1}{n} \mathbb{E} \|X - \Delta(\Phi_\delta X)\|_2^2 \quad (1)$$

Theorem 1. Define the Sample Distortion (SD) function for the source X as the minimum achievable distortion over all encoder-decoder pairs (Φ, Δ) for a fixed undersampling ratio δ . Then the SD function is convex.

Proof: Consider two achievable SD points $(\delta_1, D(\delta_1))$ and $(\delta_2, D(\delta_2))$. To prove the SD function is convex, we only need to show the convex combination of the two points is also achievable. Let $\delta_t = t\delta_1 + (1-t)\delta_2$, $0 \leq t \leq 1$. To sample the source $X \in \mathbb{R}^n$ at the undersampling ratio δ_t , we could split X into two parts $X = [X_1, X_2]^T$, where $X_1 \in \mathbb{R}^{tn}$, $X_2 \in \mathbb{R}^{(1-t)n}$, and apply encoders with undersampling ratio δ_1, δ_2 to X_1, X_2 respectively. Then the reconstruction of X_1 and X_2 has achievable MSE: $tnD(\delta_1)$ and $(1-t)nD(\delta_2)$. So the MSE of the reconstruction of X is:

$$nD(\delta_t) = tnD(\delta_1) + (1-t)nD(\delta_2) \quad (2)$$

Then we have

$$D(t\delta_1 + (1-t)\delta_2) = tD(\delta_1) + (1-t)D(\delta_2) \quad (3)$$

Through a slight abuse of terminology, we will also use the term SD function to refer to the undersampling ratio for which specific encoders and decoders can achieve a certain distortion level. The remainder of the paper considers $p(x)$ as the two states GM model:

$$\begin{aligned} p(x|s) &= p(s=1)\mathcal{N}(x; 0, \sigma_L^2) + p(s=0)\mathcal{N}(x; 0, \sigma_S^2) \\ &= \lambda\mathcal{N}(x; 0, \sigma_L^2) + (1-\lambda)\mathcal{N}(x; 0, \sigma_S^2) \end{aligned} \quad (4)$$

And we will investigate the SD function for a Gaussian encoder-Bayesian optimal AMP (BAMP) decoder pair.

Chunli Guo gratefully acknowledge the Maxwell Advanced Technology Fund at the University of Edinburgh in funding her PhD studies.

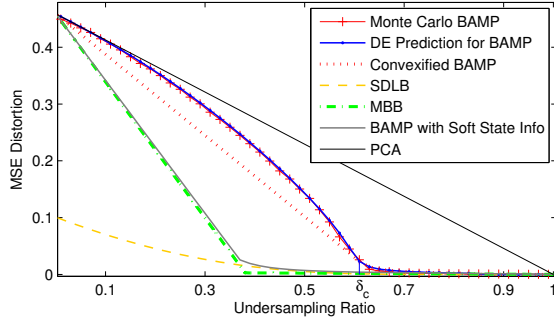


Fig. 1: SD functions for GM data $p(x) = 0.38 \mathcal{N}(0, 1.198) + 0.62 \mathcal{N}(0, 0.004)$ and lower bounds

2.1. SD function of i.i.d Gaussian mixture source

It is shown in [10] that when $p(x)$ is available, the generic AMP algorithm can be tuned optimally by using the conditional expectation as the soft thresholding function and the conditional variance for the threshold level. [11] has also shown that the asymptotic MSE for BAMP is well defined for the Gaussian random sensing matrix through a one variable density evolution (DE) formalism, thus providing a theoretical basis for the SD function. We summarize the DE function for BAMP in Table 1.

To compute the DE prediction (theoretical SD function) for GM distribution, we only need to derive $F(\cdot)$ in the formula. Given the white Gaussian noise model of Y , we have $p(Y|X) = \mathcal{N}(Y; X, c)$. With straightforward calculation, the prior of Y is

$$p(Y) = \lambda \mathcal{N}(Y; 0, c + \sigma_L^2) + (1 - \lambda) \mathcal{N}(Y; 0, c + \sigma_S^2) \quad (5)$$

Combining the definition of the conditional mean with $p(Y)$, it can be shown that [7]

$$F(\xi; c) = \frac{\alpha_L + \alpha_S \tau(\xi)}{1 + \tau(\xi)} \xi \quad (6)$$

$$\tau(\xi) = \epsilon \exp(-\kappa \xi^2) \quad (7)$$

where

$$\alpha_L := \frac{\sigma_L^2}{c + \sigma_L^2} \quad \alpha_S := \frac{\sigma_S^2}{c + \sigma_S^2}$$

$$\epsilon := \frac{1 - \lambda}{\lambda} \sqrt{\frac{c + \sigma_L^2}{c + \sigma_S^2}} \quad \kappa := \frac{\sigma_L^2 - \sigma_S^2}{2(c + \sigma_L^2)(c + \sigma_S^2)}$$

Plugging in (6), we can formally track the evolution of the mean squared error. It is illustrated in Fig. 1 that the asymptotic SD function matches well with the Monte Carlo simulation for the GM source. Close observation of the SD function for BAMP reveals that for undersampling ratio $\delta < \delta_c$, there is an better convexified SD curve by applying Theorem 1: the convex combination of the trivial decoder ($x = 0$) and the BAMP decoder at the crucial undersampling ratio δ_c . It indicates that instead of sampling the source X with $\Phi \in \mathbb{R}^{\delta N \times N}$, we could simply throw away a portion of the source $(1 - k)X$, $k = \delta/\delta_c$, and sample the rest kX with $\Phi \in \mathbb{R}^{\delta N \times kN}$ to achieve better MSE performance. The resulting convexified SD curve is also shown in Fig. 1.

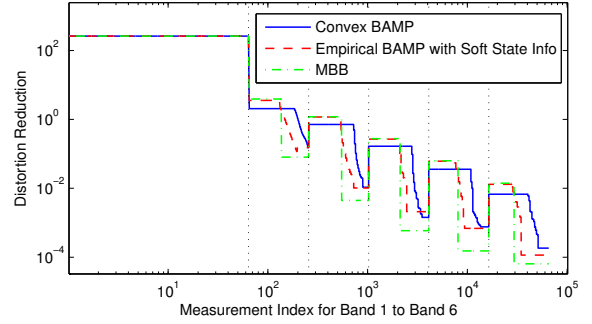


Fig. 2: DR function of 6 bands Haar wavelet decomposition of cameraman image (including the low-pass band)

Definitions: $X \sim p(X)$, $W \sim \mathcal{N}(0, c)$, $Y = X + W$
 $F(\xi; c) = \mathbb{E}(X|Y = \xi)$

Density Evolution: $D_0 = \mathbb{E}(X^2)$, $Z \sim \mathcal{N}(0, 1)$
for $k = 1, 2, 3, \dots$

$$D_{k+1} = D_0 - \mathbb{E}[XF(X + Z\sqrt{\frac{D_k}{\delta}}; \frac{D_k}{\delta})]$$

end

Table 1: The DE equation for Bayesian optimal AMP algorithm [11]

For the GM model it is possible to derive a model based bound (MBB) that is tighter than our previous sample distortion lower bound (SDLB) [4]:

$$D_{MBB}(\delta) = \begin{cases} (1 - \lambda)\sigma_S^2 + (\lambda - \delta)\sigma_L^2 & 0 \leq \delta \leq \lambda \\ (1 - \delta)\sigma_S^2 & \lambda < \delta \leq 1 \end{cases} \quad (8)$$

Assume the hidden states are available for both encoder and decoder. We can partition X into two different Gaussian groups according to S : X_1 i.i.d $\sim \mathcal{N}(0, \sigma_L^2)$ and X_2 i.i.d $\sim \mathcal{N}(0, \sigma_S^2)$. The MBB for GM data consists of two linear parts corresponding to the lower bound of the two Gaussian sources [4]. We plot both D_{SDLB} and D_{MBB} for a GM distribution in Fig. 2.

2.2. SD function and sample allocation for natural images

we begin by introducing a multi-resolution statistical model for natural images. The wavelet decomposition of an image $f(x)$, $x \in [0, 1]^2$ has the form [8]

$$f = \sum_k \mu_{i,k} \phi_{i,k} + \sum_{j \geq i,k} \omega_{j,k} \psi_{j,k} \quad (9)$$

where $\mu_{i,k}$ are the scaling coefficients at scale i and $\omega_{j,k}$ are the wavelet coefficients at scale j . We denote the group of wavelet coefficients at scale $i, i + 1, \dots$ as band 0, band 1 \dots for simplicity. The scaling coefficients are treated as Gaussian since they exhibit no sparsity. We model the wavelet coefficients as mutually independent and impose the GM distribution for each wavelet band. To be specific, $\omega_{j,k}$ at scale j follows

$$p(\omega_{j,k}) = \lambda_{j,k} \mathcal{N}(0, \sigma_{L,j}^2) + (1 - \lambda_{j,k}) \mathcal{N}(0, \sigma_{S,j}^2) \quad (10)$$

We restrict ourself to a block diagonal encoder which samples different wavelet band separately:

$$\Phi = \text{diag}(\Phi_0, \Phi_1, \Phi_2, \dots) \quad (11)$$

where $\Phi_i \in \mathbb{R}^{m_i \times n_i}$, $m_i \leq n_i$ makes m_i measurements of the n_i wavelet coefficients at the i th band. The equality holds when the i th band is fully sampled with Φ_i being an identity matrix. Otherwise Φ_i is a Gaussian random matrix. To derive the SD function for multi-resolution images, we need to optimize the sample allocation under the sample budget constraint $m = \delta n = \sum_i m_i$, with the aim of minimizing the total reconstruction distortion. We follow the ideas presented in [4] and use a distortion reduction (DR) function for each wavelet band:

$$d^{(i)}(m_i) := n_i[D((m_i - 1)/n_i) - D(m_i/n_i)] \quad (12)$$

where $m_i = 1 \cdots n_i$. Assuming $m_i - 1$ samples have been allocated to the i th band, $d^{(i)}(m_i)$ is the amount of distortion decreased by adding one more sample to that band. The solid line in Fig. 2 is the DR function for a 6 band image model. With the convexified SD function, the optimal bandwise sampling is achieved by progressively allocating samples to the band which provides the greatest distortion reduction. Thus the distortion of the multi-resolution image at undersampling ratio δ is the signal energy subtracting the total distortion reduction achieved by the optimal sample allocation.

3. SAMPLE ALLOCATION WITH TREE STRUCTURE

Until now we have assumed independence of the wavelet bands. In this section we combine Som and Schinter's recently proposed TurboAMP algorithm [7] with our sample allocation to exploit both the sparsity and the intrinsic persistence across wavelet scales. Let $\Omega = [\omega_0, \omega_1, \dots, \omega_L]$ denote the wavelet coefficients vector partitioned into different subsets according to the band index and $S = [s_0, s_1, \dots, s_L]$ the corresponding hidden states vector. Assume $Y = [y_0, y_1, \dots, y_L]$ is the collection of CS observation of the block diagonal sensing matrix. In the Bayesian compressed sensing setting, the reconstruction of Ω from Y is interpreted as computing the expectation of the posterior $p(\Omega|Y)$. To exploit the wavelet dependency across bands, we model S with the HMT structure. Then the posterior pdf has the form:

$$p(\Omega|Y) = \sum_S \prod_{j=0}^L p(\omega_j|y_j, S) p(S) \quad (13)$$

$$= \sum_{j=0}^L \frac{1}{z_j} \sum_S p(S) \prod_{k=1}^{n_j} p(\omega_{j,k}|s_{j,k}) \prod_{t=1}^{m_j} p(y_{j,t}|\omega_j) \quad (14)$$

where z_j ensures the normalization $\int p(\omega_j|y_j) = 1$. This complicated global function can be visualized using a factor graph [12] in Fig. 3. Since exact computation of $p(\Omega|Y)$ is NP hard due to the loopy structure, we split the factor graph along the dashed line into two decoupled subgraphs as in [7], to calculate the marginal posteriors $p(\omega_{j,k}|y_j)$. To be specific, we exchange the local belief of $s_{j,k}$ between AMP decoding and HMT decoding alternately, by treating the likelihood on $s_{j,k}$ from one subgraph as prior for the other subgraph. The key feature of our factor graph is for the AMP decoding part, the sensing procedure is bandwise independent rather than a mixture of all the wavelet coefficients. For the fully sampled wavelet band (e.g. band 0), it is reasonable to assume we also have the correct information of the hidden states so that it could propagate to the partially sampled bands (e.g. band 1, band 2) through the HMT and help with the decoding procedure.

One thing worth noticing here is that the SD function in section 2 is derived by assuming uniform hidden states information $\mathbb{E}[\lambda_j]$ for all the wavelet coefficients in band j . Therefore it will not predict

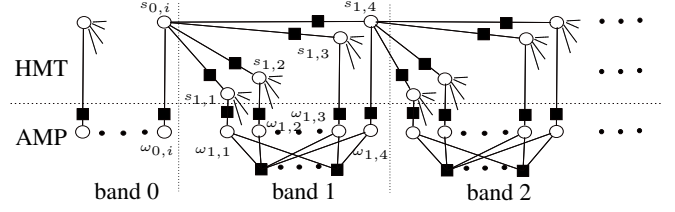


Fig. 3: Factor graph for bandwise sampling with HMT decoding

correctly and guarantee the least distortion when combined with the HMT decoding. Because the purpose of HMT decoding is to provide AMP with a good estimation of $\lambda_{j,k}$ for every wavelet coefficient through the quad-tree structure, thus enhance the reconstruction. We propose an empirical SD curve which takes account of the hidden states. We still use the Gaussian random matrix as the encoder and BAMP as the decoder. Only this time we provide the decoder with extra soft information of the hidden states $\hat{\lambda}_{j,k}$ when we run the DE iterations

$$\hat{\lambda}_{j,k} = \frac{\mathcal{N}(\omega_{j,k}; 0, \sigma_{j,L}^2)}{\mathcal{N}(\omega_{j,k}; 0, \sigma_{j,L}^2) + \mathcal{N}(\omega_{j,k}; 0, \sigma_{j,S}^2)} \quad (15)$$

The empirical SD function given "true" hidden states information is plotted in Fig. 1. The corresponding DR function for multi-resolution image model is shown in Fig. 2. To clarify the notation, we denote the independent model based sample allocation as SA and the sample allocation using the soft information of hidden states as SISA.

We now make a few comments. First, we note that the soft information SD curve lies very close to the MBB but has a smoother transition around $\delta = \mathbb{E}[\lambda]$, which means BAMP is able to almost fully recover the signal with a number of measurements that approaches the theoretical limit providing the full statistical information. Second, both SA and SISA are suboptimal for the turbo scheme. SA tends to put very few or no samples to the last wavelet band considering it contains the least energy when treated independently. It underestimates the effect of the reconstruction quality of the last band on other bands through the HMT structure. In contrast SISA is too optimistic in assuming that the HMT decoding is able to retrieve the true hidden state information thus it tends to put too many samples to the fine-scale band. The optimal sample allocation for turbo decoding must have the merit of both SA and SISA.

4. NATURAL IMAGE EXAMPLE

We take the 256×256 cameraman image as a prototypical example and use the statistical properties of its Haar and db2 wavelet decomposition to define the image model parameters. Table 2 is the estimation of the GM distribution parameters for 6 wavelet bands using EM algorithm. These parameters were used to define the multi-resolution statistical model, generate the SD function as well as the DR function shown in Fig. 2. Let us consider the sample allocation for 4 different undersampling ratios: 10%, 15.26%, 25% and 30% associated with $m = 6554, 10000, 16384, 19661$ noiseless measurements. We evaluate the performance of three different sample allocations: SA, SISA, and an empirically optimised sample allocation, or ESA. Based on the SA result, we manually moved a minimum of 100 samples from other bands to the last band each time to balance the total distortion. The ESA is the best manipulation result for turbo decoding in terms of the reconstruction distortion.

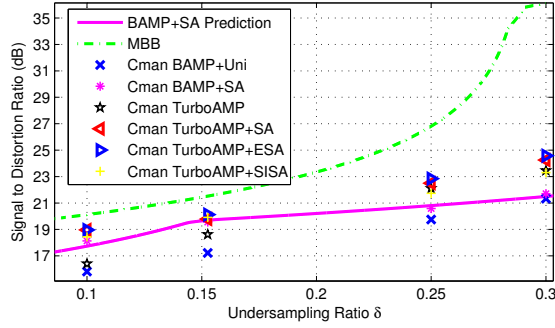


Fig. 4: SDR plot for Haar wavelet representation of cameraman

subband		scaling	band 0	band 1	band 2	band 3	band 4
Haar	λ	1	0.3752	0.3805	0.3559	0.3214	0.2642
	σ_L^2	261.5539	3.9188	1.1984	0.2757	0.0630	0.0141
	σ_S^2		0.0806	0.0044	0.0006	0.0002	0.0001
db2	λ	1	0.4155	0.5309	0.4842	0.3664	0.2792
	σ_L^2	261.4383	4.4215	0.8542	0.1856	0.0453	0.0115
	σ_S^2		0.3331	0.0038	0.0004	0.0002	0.0001

Table 2: Statistics for the Haar and db2 wavelet of cameraman

Fig. 6 is the sample arrangement for the cameraman Haar wavelet coefficients. Note that for all three sample allocations, the scaling coefficients and the root wavelet coefficients (band 0) are always fully sampled before other wavelet bands are partially sampled or not sampled at all. The sample allocation for db2 wavelet follows the same pattern. To implement TurboAMP [7], we initialized the activity rate with the λ reported in Table 2 for each band. The hyperparameters for the transition matrix were set in accordance with the values described in [7]. For turbo scheme with sample allocation, we use the soft information $\hat{\lambda}_{j,k}$ if band j is fully sampled and $\mathbb{E}[\lambda_j]$ if it is partially sampled. For both TurboAMP and turbo with sample allocation, we ran 20 turbo iterations, within which 50 AMP iterations were performed. We compared the numerical results for the following decoding scheme: BAMP with uniform distribution of samples (BAMP+Uni), BAMP+SA, TurboAMP, TurboAMP+SA, TurboAMP+ESA and TurboAMP+SISA. The results are plotted in Fig. 4 and Fig. 5 for Haar wavelet and db2 wavelet respectively. First we observed that the theoretical SD prediction for BAMP+SA is reinforced by the experimental evaluation on the cameraman image. Secondly, it is not surprising to see BAMP+Uni performs the worst since it does not exploit any wavelet property. Between TurboAMP and BAMP+SA, TurboAMP performs relatively poorly under small sampling ratios $\delta = 10\%$, 15% , while beats BAMP+SA when δ equals to 25% and 30% . It highlights the importance of the sample allocation when there is a tight budget of samples. It also confirms that given enough samples, the HMT structure working cooperatively with the AMP decoding benefits the reconstruction.

Unsurprisingly, the combination of turbo scheme and sample allocation generally delivers better reconstruction results than either of the two strategies working alone. The advantage of the cooperation is more obvious for small sampling ratios. For $\delta = 10\%$, TurboAMP+SA has relatively 2.5dB SDR gain over TurboAMP for both Haar wavelet and db2 wavelet. For $\delta = 15\%$, 25% and 30% , there is an average gain of 1dB. We also observed that although the TurboAMP+ESA has the best performance among all, it only slightly

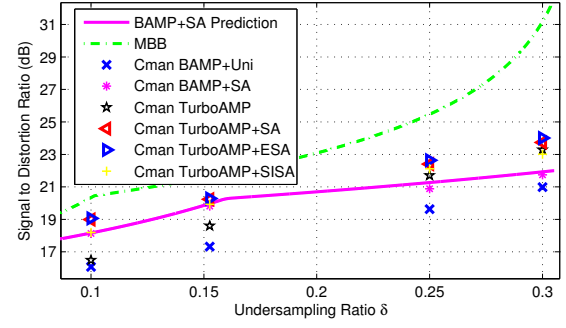


Fig. 5: SDR plot for db2 wavelet representation of cameraman

improves the reconstruction over the TurboAMP+SA. When we are extremely lack of samples (e.g $\delta = 10\%$), we should ensure enough samples to the coarse-scale bands. Any movement of samples to the fine-scale bands will jeopardise the whole reconstruction since the energy of the coarse bands dominates the whole image. Even when we have the luxury of manipulating samples to the last band benefiting from the tree structure decoding, the improvement is limited because of the exponential decay of energy across wavelet scale.

5. CONCLUSION

In this paper we present the sample distortion framework for GM distribution signal using the Bayesian optimal AMP algorithm. By expanding the concept of SD function to the multi-resolution statistical image model, we derive an optimized bandwise sampling strategy with the aim of minimizing the reconstruction distortion. While the advantage and importance of bandwise sampling of CS has been illustrated by Tsai [13] as the two-gender hybrid CS, we provide a quantitative method with sound theoretical basis to predict and assess the achievable performance. The optimised bandwise sampling is related to the work in [14], where the authors aim to maximise information content. Here, in contrast, we derive a rigorous BAMP based sample allocation to minimize distortion. We build upon the turbo decoding scheme [7] to incorporate the tree structure image model with the bandwise sampling and present improvement in reconstruction quality. Although the sample allocation based on independent model is theoretically suboptimal for turbo decoding, the numerical simulation suggests it is very close to the manually best achievable performance.

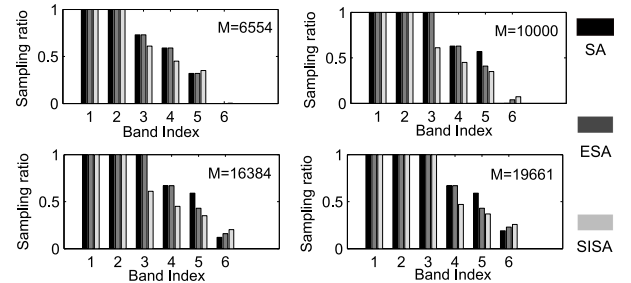


Fig. 6: Sample allocation per band for Haar wavelet.

6. REFERENCES

- [1] Shihao Ji, Ya Xue, and L. Carin, “Bayesian compressive sensing,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2346–2356, june 2008.
- [2] F. Sausset Y. Sun F. Krzakala, M. Mezard and L. Zdeborova, “Statistical physics-based reconstruction in compressed sensing,” arXiv:1109.4424, 2011.
- [3] Andrea Montanari David L. Donoho, Arian Maleki, “Message passing algorithms for compressed sensing,” arXiv:0907.3574, 2009.
- [4] M.E. Davies and Chunli Guo, “Sample-distortion functions for compressed sensing,” in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, sept. 2011, pp. 902–908.
- [5] M. Bayati and A. Montanari, “The lasso risk for gaussian matrices,” *Information Theory, IEEE Transactions on*, vol. 58, no. 4, pp. 1997–2017, april 2012.
- [6] R. Gribonval, V. Cevher, and M.E. Davies, “Compressible distributions for high-dimensional statistics,” *Information Theory, IEEE Transactions on*, vol. 58, no. 8, pp. 5016–5034, aug. 2012.
- [7] P Som, S.; Schniter, “Compressive imaging using approximate message passing and a markov-tree prior,” *Signal Processing, IEEE Transactions on*, vol. 60, pp. 3439–3448, 2012.
- [8] S. Mallat, *A Wavelet Tour of Signal Processing*, San Diego, CA: Academic Press, 1999.
- [9] J.K. Romberg, Hyeokho Choi, and R.G. Baraniuk, “Bayesian tree-structured image modeling using wavelet-domain hidden markov models,” *Image Processing, IEEE Transactions on*, vol. 10, no. 7, pp. 1056–1068, jul 2001.
- [10] D.L. Donoho, A. Maleki, and A. Montanari, “Message passing algorithms for compressed sensing: I. motivation and construction,” in *Information Theory Workshop (ITW), 2010 IEEE*, jan. 2010.
- [11] F. Sausset F. Krzakala, M. Mezard and L. Z. Yifan Sun, “Probabilistic reconstruction in compressed sensing: Algorithms, phase diagrams, and threshold achieving matrices,” arXiv:1206.3953v1 [cond-mat.statmech], 2012.
- [12] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 498–519, feb 2001.
- [13] Y.Tsaig, *Sparse solution of underdetermined linear systems: algorithms and applications*, Ph.D. thesis, Stanford University, 2007.
- [14] Hyun Sung Chang, Y. Weiss, and W.T. Freeman, “Informative sensing of natural images,” in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, nov. 2009, pp. 3025–3028.