AUTOCALIBRATED SIGNAL RECONSTRUCTION FROM LINEAR MEASUREMENTS USING ADAPTIVE GAMP

Ulugbek S. Kamilov, Aurélien Bourquard, Emrah Bostan, and Michael Unser

Biomedical Imaging Group, EPFL, Switzerland

ABSTRACT

In this paper, we reconstruct signals from underdetermined linear measurements where the componentwise gains of the measurement system are unknown a priori. The reconstruction is performed through an adaptation of the message-passing algorithm called adaptive GAMP that enables joint gain calibration and signal estimation. To evaluate our approach, we apply it to the problem of sparse recovery and compare it against an ℓ_1-based approach. We numerically show that adaptive GAMP yields excellent results even for a moderate amount of data. It approaches the performance of oracle GAMP where the gains are perfectly known asymptotically.

***Index Terms*—** Sparsity, inverse problems, approximate message passing, blind learning, MRI

1. INTRODUCTION

We consider a linear model where an unknown signal $\mathbf{x} \in \mathbb{C}^N$ is represented by measurements

$$\mathbf{y} = \mathbf{F}\mathbf{S}\mathbf{x} + \mathbf{w}. \tag{1}$$

The matrix $\mathbf{F} \in \mathbb{C}^{M \times N}$ with $M \leq N$, which models the effect of the measurement system, is assumed to be known. The unknown component of the system is a diagonal matrix $\mathbf{S} \in \mathbb{R}_+^{N \times N}$ that contains a series of gains applied to $\mathbf{x}$. The vector $\mathbf{w} \in \mathbb{C}^M$ represents complex additive white Gaussian noise (AWGN).

In this work, our task is to jointly estimate $\mathbf{S}$ and $\mathbf{x}$. In particular, we are interested in a *blind* calibration scenario [1] where, given L measurement vectors assembled in the matrix $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_L]$, we estimate the gains $\mathbf{S}$ while simultaneously recovering the unknown signals $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_L]$. This joint estimation problem arises in several applications. In parallel MRI, for instance, the matrix $\mathbf{F}$ represents a Fourier encoding, while $\mathbf{S}$ corresponds to the sensitivity profile associated to a receiver coil [2]. Another potential application is ptychographic nanotomography [3], which involves Fourier measurements and a constant illumination profile that can be modeled by a diagonal positive-semidefinite matrix.

This work was supported by the European Commission under Grant ERC-2010-AdG 267439-FUN-SP.

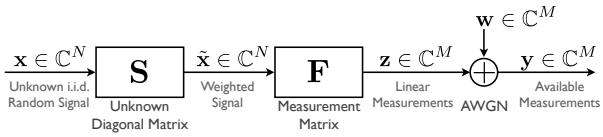

Fig. 1. Joint estimation problem considered in this work.

Although the joint recovery of the signals $\mathbf{X}$ and the gain matrix $\mathbf{S}$ can be formulated as a variational problem, we concentrate on the alternative approach based on message-passing algorithms [4–6]. The major advantage of these algorithms are their generality, low computational cost, and reconstruction performance. In particular, *generalized approximate message passing* (GAMP) was shown to successfully solve (1) when the problem is formulated in Bayesian terms under the assumption that $\mathbf{S}$ as well as the signal and noise distributions are perfectly known [6]. Recently, a generalized version of GAMP called *adaptive GAMP* was proposed for solving inverse problems where the signal and noise distributions have parametric uncertainties [7]. Adaptive GAMP generalizes several prior works [8, 9], and provably yields asymptotically consistent estimates of the unknown parameters. In the following, we demonstrate that the adaptive GAMP framework can be extended in order to solve (1). This makes it potentially applicable to a large class of calibration problems arising in several applications.

2. BAYESIAN FORMULATION

We start by introducing the Bayesian formulation of the estimation problem illustrated in Figure 1. The input signal $\mathbf{x}$ is random with a separable distribution

$$p_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\lambda}) = \prod_{i=1}^{N} p_x(x_i; \boldsymbol{\lambda}), \tag{2}$$

where $\boldsymbol{\lambda}$ represents the parameters of the prior. The noiseless measurement vector $\mathbf{z} = \mathbf{y} - \mathbf{w}$ is obtained via $\mathbf{z} = \mathbf{F}\mathbf{S}\mathbf{x}$. Each component z_i is then corrupted by complex i.i.d. Gaussian noise $w_i \sim \mathcal{CN}(0, v_w)$ to yield the vector of measurements $\mathbf{y}$. Note that, in practice, the noise variance v_w as well

as the parameters of the prior $\boldsymbol{\lambda}$ are unknown and must be estimated from the data.

The conditional probability distribution of the signal $\mathbf{x}$ given the measurements $\mathbf{y}$ is given by

$$\begin{aligned} p_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) &\propto p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) p_{\mathbf{x}}(\mathbf{x}) \\ &\propto \prod_{i=1}^{M} \mathcal{G}(y_i - z_i; v_w) \prod_{j=1}^{N} p_x(x_j; \boldsymbol{\lambda}), \end{aligned} \tag{3}$$

where $\propto$ denotes identity after normalization to unity and $\mathcal{G}$ is the Gaussian distribution

$$\mathcal{G}(x-\mu; v_w) = \frac{1}{(\sqrt{2\pi(v_w/2)})^2} \mathrm{e}^{-\frac{|x-\mu|^2}{2(v_w/2)}}, \tag{4}$$

where $x, \mu \in \mathbb{C}$. The distribution $p_{\mathbf{x}|\mathbf{y}}$ in (3) implicitly depends on $\mathbf{S}$, $\boldsymbol{\lambda}$, and v_w. When all the parameters are known, it provides a complete statistical characterization of the problem. In particular, the MAP and MMSE estimators of $\mathbf{x}$ are specified by

$$\widehat{\mathbf{x}}_{\mathrm{MAP}} = \underset{\mathbf{x}\in\mathbb{C}^N}{\operatorname{argmax}} \left\{ p_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) \right\}, \tag{5a}$$

$$\widehat{\mathbf{x}}_{\mathrm{MMSE}} = \mathbb{E}\left[\mathbf{x} \mid \mathbf{y}\right]. \tag{5b}$$

In the sequel, we develop computational approximations to these estimators that also allow to recover $\mathbf{S}$.

3. JOINT VARIATIONAL ESTIMATION

To estimate the unknowns, we first propose a variational approach based on the MAP estimator (5a). Assuming a Laplace-type prior on each signal $\mathbf{x}_l$, the solution is derived as the minimum of a cost functional that involves ℓ_1-type regularization. The Laplace prior has been extensively studied as a model of signal sparsity and applied to several types of problems such as compressed sensing [10, 11]. In the case of our model (1), the recovered quantities are expressed as

$$\begin{aligned} \{\widehat{\mathbf{X}}, \hat{\mathbf{S}}\} \quad = \quad & \underset{\mathbf{X},\mathbf{S}}{\operatorname{argmin}} \left\{ \sum_{l=1}^{L} \|\mathbf{F}\mathbf{S}\mathbf{x}_l - \mathbf{y}_l\|_2^2 + \tau \|\mathbf{x}_l\|_1 \right\} \\ & \text{s.t. } \mathrm{Tr}(\mathbf{S}) = N,\ \mathbf{S} \in \mathbb{R}_+^{N\times N}. \end{aligned} \tag{6}$$

Unless already known, the gains $\mathbf{S}$ are to be jointly estimated with the signals $\mathbf{X}$. In our algorithm, we thus proceed through alternate optimizations, keeping either quantity fixed in each step. According to our formulation, each update of $\widehat{\mathbf{X}}$ amounts to solving a linear program (for which excellent convex-optimization approaches are available and described in the literature [12]), while each gain estimation corresponds to the minimization of a penalized-least-squares-type expression. The trace condition appearing in (6) guarantees the well posedness of the problem and avoids possible scaling ambiguities between $\mathbf{S}$ and each $\mathbf{x}_l$.

Note that a similar ℓ_1-based approach has been proposed and applied to joint calibration and to the recovery of sparse signals by Gribonval *et al.* [1]. Their work, however, addresses a distinct problem where $\mathbf{S}$ and $\mathbf{F}$ act on the unknowns in reverse order. At any rate, our main point in this section is to specify a reference method against which to compare our new message-passing algorithm.

4. CALIBRATION WITH ADAPTIVE GAMP

The GAMP algorithm developed in [6] is an iterative approach to efficiently approximate the MMSE estimator (5b). The algorithm relies on the Central Limit Theorem to simplify loopy belief propagation by replacing continuous-domain convolutions with matrix-vector multiplications followed by pointwise nonlinearities [6,13]. Although the basic GAMP algorithm requires perfect knowledge of $\mathbf{S}$, $\boldsymbol{\lambda}$, and v_w for reconstruction, it was recently extended to incorporate parameter learning [7–9]. In particular, a recently introduced adaptive GAMP method combines maximum-likelihood (ML) estimation with the standard GAMP updates [7]. Adaptive GAMP was shown to yield consistent parameter estimates asymptotically, which corroborates the performance of reconstruction that is addressed in Section 5. Details and complete analysis of adaptive GAMP can be found in [7]. Here, we present the algorithm applied to the specific problem of blind learning from L available measurement vectors.

Given the measurements $\mathbf{Y} \in \mathbb{C}^{M\times L}$, the measurement matrix $\mathbf{F} \in \mathbb{C}^{M\times N}$, and the initial values $\widehat{\boldsymbol{\lambda}}^0$, $\widehat{v}_w^0$, and $\widehat{\mathbf{S}}^0$, the proposed adaptive GAMP-based calibration is given next.

1. *Initialization:* Set $t = 0$ and evaluate

$$\widehat{\mathbf{X}}^0 = \mathbb{E}\left[\mathbf{X}\right], \mathbf{V}_x^0 = \mathrm{var}\left[\mathbf{X}\right], \mathbf{U}^0 = 0, \tag{7}$$

where the expected value and variance are with respect to $p_{\mathbf{x}}(\mathbf{x})$ in (2).

2. *Measurement Update:* First, compute the linear step

$$\mathbf{V}_p^{t+1} = |\mathbf{F} \odot \mathbf{F}||\widehat{\mathbf{S}}^t|^2 \mathbf{V}_x^t, \tag{8a}$$

$$\mathbf{P}^{t+1} = \mathbf{F}\widehat{\mathbf{S}}^t\widehat{\mathbf{X}}^t - \mathbf{V}_p^{t+1} \odot \mathbf{U}^t, \tag{8b}$$

where the modulus operator $|\cdot|$ is applied componentwise and $\odot$ denotes the Hadamard (pointwise) product. Then, estimate the AWGN variance $\widehat{v}_w^{t+1}$ by solving

$$\widehat{v}_w^{t+1} = \underset{v_w\in\mathbb{R}_+}{\operatorname{argmax}} \left\{ \sum_{l=1}^{L} \log\left(p_{\mathbf{y}}(\mathbf{y}_l|v_w)\right) \right\}, \tag{9}$$

where

$$p_{\mathbf{y}}(\mathbf{y}|v_w) = \prod_{i=1}^{N} \mathcal{G}(y_i - p_i^{t+1}; v_{pi}^{t+1} + v_w). \tag{10}$$

Finally, evaluate the nonlinear step

$$\mathbf{U}^{t+1} = \mathrm{F}_1(\mathbf{Y}, \mathbf{P}^{t+1}, \mathbf{V}_p^{t+1}), \tag{11a}$$

$$\mathbf{V}_u^{t+1} = \mathrm{F}_2(\mathbf{Y}, \mathbf{P}^{t+1}, \mathbf{V}_p^{t+1}), \tag{11b}$$

where the scalar functions F_1 and F_2 are applied componentwise and given by

$$\mathrm{F}_1(y, p, v_p) = \frac{1}{v_p}\left(\mathbb{E}\left[z \mid y\right] - p\right),$$

$$\mathrm{F}_2(y, p, v_p) = \frac{1}{v_p}\left(1 - \frac{\mathrm{var}[z \mid y]}{v_p}\right).$$

The expected value and variance implicitly depend on $\widehat{v}_w^{t+1}$ and are evaluated with respect to $z \sim \mathcal{CN}(p, v_p)$.

3. *Estimation Update:* First, compute the linear step

$$\mathbf{V}_r^{t+1} = \left(|\mathbf{F} \odot \mathbf{F}|^* \mathbf{V}_u\right)^{-1}, \tag{12a}$$

$$\mathbf{R}^{t+1} = \widehat{\mathbf{X}}^t + \mathbf{V}_r^{t+1} \odot \left(\mathbf{F}^* \mathbf{U}^{t+1}\right), \tag{12b}$$

where the inversion is componentwise and $\mathbf{F}^*$ denotes the Hermitian transpose of $\mathbf{F}$. Then, estimate $\boldsymbol{\lambda}$ and $\mathbf{S}$ by solving

$$\left\{\widehat{\mathbf{S}}^{t+1}, \widehat{\boldsymbol{\lambda}}^{t+1}\right\} = \underset{\mathbf{S}, \boldsymbol{\lambda}}{\mathrm{argmax}} \left\{\mathcal{L}\left(\mathbf{R}^{t+1} | \mathbf{S}, \boldsymbol{\lambda}\right)\right\}$$

$$\text{s.t. } \mathrm{Tr}(\mathbf{S}) = N, \ \mathbf{S} \in \mathbb{R}_+^{N \times N}, \tag{13a}$$

where the log-likelihood function $\mathcal{L}$ is given by

$$\mathcal{L}\left(\mathbf{R}^{t+1} | \mathbf{S}, \boldsymbol{\lambda}\right) = \sum_{l=1}^{L} \sum_{j=1}^{N} \log\left(p_r(r_{lj}^{t+1} | s_{jj}, \boldsymbol{\lambda})\right).$$

The likelihood function is given by $p_r(r|s, \boldsymbol{\lambda}) \propto \mathbb{E}\left[\mathcal{G}(r - sx; v_r)\right]$, where the expectation is with respect to $p_x(x; \boldsymbol{\lambda})$. Finally, evaluate the nonlinear steps

$$\widehat{\mathbf{X}}^{t+1} = \mathrm{G}_1\left(\mathbf{R}^{t+1}, \mathbf{V}_r^{t+1}\right), \tag{14a}$$

$$\mathbf{V}_x^{t+1} = \mathrm{G}_2\left(\mathbf{R}^{t+1}, \mathbf{V}_r^{t+1}\right), \tag{14b}$$

where the scalar functions G_1 and G_2 are applied componentwise and given by

$$\mathrm{G}_1(r, v_r) = \mathbb{E}\left[x \mid r\right],$$

$$\mathrm{G}_2(r, v_r) = \mathrm{var}\left[x \mid r\right].$$

The expected value and the variance are evaluated with respect to $p_{x|r}(x|r) \propto \mathcal{G}(r - sx; v_r) p_x(x; \boldsymbol{\lambda})$. This is an AWGN denoising problem with gain s and noise $n \sim \mathcal{CN}(0, v_r)$.

4. *Iterate:* Set $t \leftarrow t + 1$ and proceed to Step 2.

For each iteration $t = 0, 1, 2, \ldots$, the proposed update rules produce the estimates $\widehat{\mathbf{X}}^t$ and $\widehat{\mathbf{S}}^t$ of the collection $\mathbf{X}$ of true signals and diagonal matrix $\mathbf{S}$ of gains, respectively. Thus, the algorithm reduces the intractable estimation problem that involves high-dimensional integration to a sequence of simple operations that can be evaluated numerically in an efficient way.

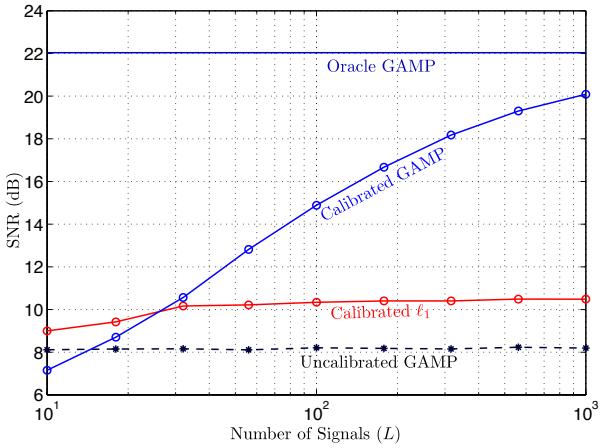

Fig. 2. Reconstruction of sparse signals from undercomplete Fourier measurements. The average reconstruction SNR is plotted against the number of training signals L.

5. EXPERIMENTS

Applications such as MRI suggest that there is considerable value in solving estimation problems of the form (1) in the context of Fourier measurements $\mathbf{F}$ with sparse priors. Although estimators based on ℓ_1-regularization in some suitable transform domain offer a good reconstruction performance, there is often a large performance gap between ℓ_1-based approaches and MMSE estimation [14].

We consider a simple simulation that illustrates how the use of calibration with adaptive GAMP presented in Section 4 outperforms the ℓ_1-approach of Section 3. Specifically, we evaluate the performance of our adaptive-GAMP-based calibration for estimating sparse signals from noisy Fourier measurements. We perform 100 random trials where we generate signals of length $N = 100$ with $\rho = 0.2$ fraction of nonzero components drawn from a zero-mean normal distribution of variance $v_x = 5$. The undersampling ratio is $\beta = M/N = 0.8$ and the noise is AWGN of variance $v_w = 0.01$.

The reconstruction performance of both ℓ_1 and GAMP-based approaches is compared in terms of signal-to-noise ratio (SNR) of the whole signal stack. The results are shown in Figure 2 as a function of the number L of available signals. *Uncalibrated GAMP* corresponds to the standard adaptive GAMP that learns ρ, v_x, and v_w, but does not account for $\mathbf{S}$. The slight 0.1 dB increase in its performance for higher L is due to improved estimation of these parameters. *Calibrated* ℓ_1 uses the approach in (6) to estimate the signals and recover $\mathbf{S}$. It yields about 2 dB improvement over uncalibrated GAMP and significantly improves (up to 1.5 dB) its own performance when L increases. For the ℓ_1-method, the parameter τ was optimized for best MSE performance; this oracle-aided performance would not be achievable in practice. *Oracle GAMP* corresponds to the basic GAMP algo-

rithm where all parameters, including $\mathbf{S}$, are known. It represents the best achievable GAMP performance. Finally, *calibrated GAMP* attempts to recover all the unknown parameters simultaneously from the data $\mathbf{Y}$, as described in Section 4. The results show that it rapidly approaches the performance of oracle GAMP and outperforms both uncalibrated GAMP and calibrated ℓ_1 for most L.

6. CONCLUSIONS

We have derived an algorithm based on adaptive GAMP that reconstructs signals and simultaneously calibrates the corresponding forward model. The proposed method has been compared to the state of the art through numerical experiments, and the results have shown our calibration approach to yield promising results—starting from moderate amounts of signals—while converging to the oracle performance asymptotically. In specific applications where the parameters to calibrate follow a prior statistical distribution, the potential reconstruction performance would be likely to further improve. This issue is worth addressing in subsequent investigations.

7. RELATION TO PRIOR WORK

The work presented here focuses on a blind calibration scenario. Previously, a distinct calibration setting involving additive uncertainties has been investigated in [15] using an estimation approach that is based on GAMP. Meanwhile, Gribonval *et al.* have studied a calibration framework similar to ours [1] where the gain matrix $\mathbf{S}$ acts on the measurements $\mathbf{Fx}$. They have developed a variational reconstruction algorithm that reconstructs $\mathbf{x}$ while simultaneously calibrating $\mathbf{S}$. In contrast to their work, we propose a statistical estimation method based on the adaptive GAMP algorithm [7] that outperforms the ℓ_1-based approach. We modify adaptive GAMP to calibrate the measurement system by learning $\mathbf{S}$ directly from L measurement vectors. Our work thus extends the scenario considered in [7–9] where $\mathbf{S}$ is assumed to be known.

8. REFERENCES

[1] R. Gribonval, G. Chardon, and L. Daudet, "Blind calibration for compressed sensing by convex optimization," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process.*, Kyoto, Japan, March 25-30, 2012, pp. 2713–2716.

[2] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, November 1999.

[3] P. Thibault and M. Guizar-Sicairos, "Maximum-likelihood refinement for coherent diffractive imaging," *New J. Phys.*, vol. 14, no. 6, pp. 063004, June 2012.

[4] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, November 2009.

[5] 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, February 2011.

[6] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Information Theory*, St. Petersburg, Russia, July 31-August 5, 2011, pp. 2168–2172.

[7] U. S. Kamilov, S. Rangan, A. K. Fletcher, and M. Unser, "Approximate message passing with consistent parameter estimation and applications to sparse learning," in *Proc. NIPS 2012*, Lake Tahoe, NV, USA, December 3-6, 2012, pp. 2447–2455.

[8] J. Vila and P. Schniter, "Expectation-maximization Bernoulli-Gaussian approximate message passing," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, November 6-9, 2011.

[9] F. Krzakala, M. Mézard, F. Sausset, Y. F. Sun, and L. Zdeborová, "Statistical-physics-based reconstruction in compressed sensing," June 2012, arXiv:1109.4424v4 [cond-mat.stat-mech].

[10] E. J. 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, February 2006.

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

[12] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.

[13] U. S. Kamilov, A. Bourquard, A. Amini, and M. Unser, "One-bit measurements with adaptive thresholds," *IEEE Signal Process. Lett.*, vol. 19, no. 10, pp. 607–610, 2012.

[14] U. S. Kamilov, P. Pad, A. Amini, and M. Unser, "MMSE estimation of sparse Lévy processes," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 137–147, January 2013.

[15] J. T. Parker, V. Cevher, and P. Schniter, "Compressive sensing under matrix uncertainties: An approximate message passing approach," in *Conf. Rec. Asilomar Conf. on Signals, Syst. and Computers*, Pacific Grove, CA, USA, November 6-9, 2011, pp. 804–808.