Robust compressed sensing with side information based on Laplace mixtures models

Chiara Ravazzi
Institute of Electronics, Computer and Telecommunication Engineering,
National Research Council of Italy, Torino, Italy
Email: chiara.ravazzi@ieiit.cnr.it

Enrico Magli
Department of Electronics and Telecommunications,
Politecnico di Torino, Italy
Email: enrico.magli@polito.it

***Abstract*—In this paper, we propose a new method for the recovery of a sparse signal from few linear measurements using a reference signal as side information. Modeling the signal coefficients with a double Laplace mixture model, and assuming that the distribution of the components of the prior information differs slightly from the unknown signal, the problem is formulated as a weighted ℓ_1 minimization problem.**

We derive sufficient conditions for perfect recovery and we show that our method is able to reduce significantly the number of measurements required for reconstruction. Numerical experiments demonstrate that the proposed approach outperforms the best algorithms for compressed sensing with prior information and is robust in imperfect scenarios.

I. Sparse recovery with side information

In this paper, we consider the problem of compressed sensing with side information as addressed in [1]. More precisely, we are interested in recovering the high dimensional signal $x^\star$ from $y = Ax^\star \in \mathbb{R}^m$ ($m \ll n$), with the additional information that $x^\star$ is sparse (i.e., it has at most $k \ll n$ nonzero entries) and is similar to a reference signal w. This problem arises in several situations, as in compressive image sampling [2], [3], where the spatial and temporal correlation within image/video is exploited. Also in sensor/camera networks [4], [5], [6], the signals acquired by close sensors are similar and can be used as prior information. In [1], the authors propose to solve the following optimization problem

$$\min_{x\in\mathbb{R}^n} \|x\|_1 + \gamma\|x-w\|_p^p \quad \text{s.t. } y = Ax \tag{1}$$

with $p \in \{1,2\}$ and $\gamma > 0$, referred as ℓ_1-ℓ_1 minimization, and ℓ_1-ℓ_2 minimization, respectively. Moreover, it is shown that the number of measurements required by ℓ_1-ℓ_1 minimization is much smaller than that obtained using classical CS. The use of prior information as a tool to reduce the number of measurements required for signal reconstruction has appeared in CS literature [1], [7], [8] also with different assumptions. In [7], the authors employ as prior information an estimate T of the support of $x^\star$ and propose a truncated ℓ_1-minimization problem, i.e. the minimization of

$$\min \|x_{T^c}\|_1 \quad \text{s.t. } y = Ax^\star. \tag{2}$$

It should be noticed that (2) can be adapted to our problem using $T = \mathrm{supp}(w)$ (Mod-CS, [7]). Another piece of literature [8], [9] considers a weighted ℓ_1-minimization with weights $w_i = -\log p_i$ where p_i is the probability that $x_i^\star = 0$. It should be remarked that in our setting p_i is not available.

II. Proposed method

Here, we propose a new weighted ℓ_1 minimization, which we call 2LMM-CS. The fundamental idea is to use a *good generative model* for sparse and compressible vectors [10], [11]. For this purpose, we use a Laplace mixture model (2LMM) as the parametric representation of the prior distribution of the signal coefficients:

$$x_i^\star = z_i u_i + (1-z_i)v_i, \quad i \in \{1,\dots,n\}$$

where u_i are identically and independently distributed (i.i.d.) according to $\mathsf{Laplace}(0,\alpha)$, v_i are i.i.d. as $\mathsf{Laplace}(0,\beta)$ and z_i are i.i.d. Bernoulli random variables with probability mass function $\mathbb{P}(z_i = 1) = 1-p$, with $p = K/n < 1/2$, $0 < \alpha \ll \beta$, and $K \geq k$ is an estimate of the signal sparsity. Then, we cast the estimation problem as a non convex optimization problem that incorporates the parametric representation of the signal. However, the optimization problem turns out to be computationally hard. Assuming that the distribution of the signal coefficients of w is similar to that of $x^\star$, the estimation problem is simplified to

$$\min_{x\in\mathbb{R}^n} \sum_{i\in T}\omega|x_i| + \sum_{i\in T^c}(\pi_i + (1-\pi_i)\omega)\,|x_i|, \quad \text{s.t. } Ax = y, \tag{3}$$

with $\omega = \alpha/\beta$ and

$$\pi_i = \pi_i(w) = \mathbb{P}(z_i = 1|w,\alpha,\beta) = \left(1 + \frac{\alpha}{\beta}\frac{p}{1-p}\mathrm{e}^{|w_i|\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)}\right)^{-1}, \tag{4}$$

$T^c = \mathrm{supp}(\sigma_{n-K}(\pi(w)))$, and $\sigma_j(v)$ is a thresholding operator which acts on v by keeping the j biggest elements in absolute value and setting the others to zero.

The following theorem shows that in the large system limit, as n is large enough and for a sufficiently small $\alpha \approx 0$ and the prior information has good enough quality, then the number of measurements sufficient for perfect reconstruction of a sparse vector supported on Λ with $|\Lambda| \leq k$ can be significantly reduced.

Theorem 1. *Let $0 \approx \alpha \ll \beta$, $k \leq K$, $\pi = \pi(w)$ as defined in (4) and $A \in \mathbb{R}^{m\times n}$ be a matrix whose entries are i.i.d. Gaussian random variables with zero-mean and unit variance. If $\|w - x^\star\| \leq \frac{1}{2}\min_{i\in\Lambda}|x_i^\star|$ then the weighted ℓ_1-minimization in (3) uniquely recovers $x^\star$ from measurements $y = Ax^\star$ if it holds condition*

$$m \geq K + O\left((K-k)\frac{1}{1+\frac{\alpha}{\beta}\mathrm{e}^{|r(w)_K|(1/\alpha-1/\beta)}}\ln(en/k)\right),$$

where $r(w)$ the non increasing rearrangement of w, i.e. $r(w) = (|w_{i_1}|,\dots,|w_{i_n}|)$ with $|w_{i_\ell}| \geq |w_{i_{\ell+1}}|$ for all $\ell = 1,\dots,n-1$.

In Fig. 1, we show the empirical recovery success rate, averaged over 50 experiments, as a function of the number of measurements m. For a fixed m, we mean the success when the algorithm reconstructs the signal $x^\star$ with a relative error smaller than 10^{-2}. The sensing matrix entries A_{ij} are sampled from the Gaussian distribution with zero mean and variance $1/m$. It should be noticed that the number of measurements required for reconstruction can be significantly reduced compared to the techniques used in literature before. Moreover, the reconstruction is robust against imperfect scenarios (Fig. 2 and 3), achieving excellent performance in several situations.

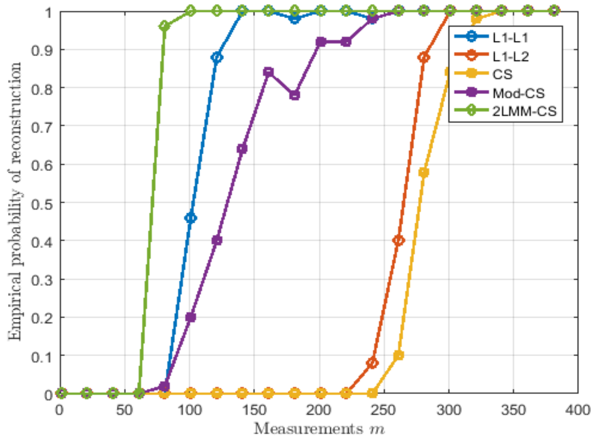

Fig. 1. Empirical probability of reconstruction of classical CS (BP), ℓ_1-ℓ_1 minimization, ℓ_1-ℓ_2 minimization, and 2LMM-CS. A signal $x^\star$ of length $n = 1000$ is generated with sparsity $k = 70$. The nonzero elements of $x^\star$ are drawn from a standard Gaussian distribution. The prior information w is obtained $w = x^\star + z$, where z is a 28-sparse signal, whose nonzero elements are drawn from a Gaussian distribution with standard deviation 0.8. Mod-CS uses as prior information $T = \mathrm{supp}(w)$. The parameters have been set to $\alpha = 10^{-4}$, $\beta = 10$, $K = |\mathrm{supp}(w)| = 76$.

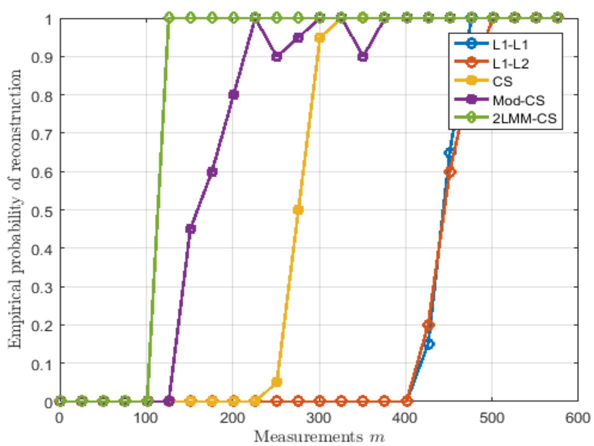

Fig. 2. Empirical probability of reconstruction of classical CS (BP), ℓ_1-ℓ_1 minimization, ℓ_1-ℓ_2 minimization, and 2LMM-CS in imperfect scenarios. We consider signal $x^\star$ and z generated as in the previous experiment. The prior information w is obtained by $w = x^\star + z + \eta$, where η is a gaussian noise with standard deviation 10^{-3}. Mod-CS uses as prior information the set T of the 123 largest components in absolute value of vector w. For CS-2LMM the mixture parameters have been set as follows: $\alpha = 10^{-4}$, $\beta = 10$, and $K = 123$.

Acknowledgments

This work is supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement n. 279848.

References

[1] J. F. C. Mota, N. Deligiannis, and M. R. D. Rodrigues, "Compressed sensing with prior information: Optimal strategies, geometry, and bounds," preprint: http://dx.doi.org/10.1109/18.382009.

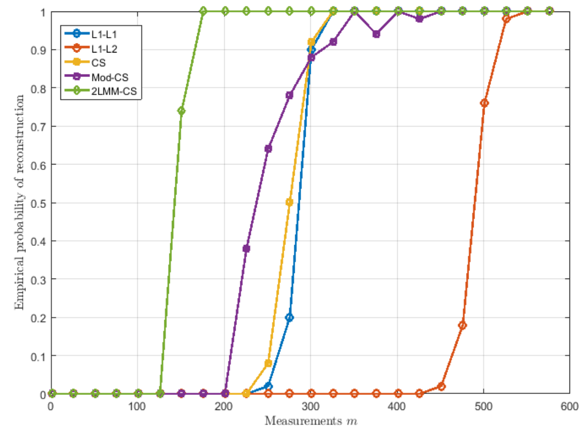

Fig. 3. Empirical probability of reconstruction of classical CS (BP), ℓ_1-ℓ_1 minimization, ℓ_1-ℓ_2 minimization, and 2LMM-CS in imperfect scenarios. $x^\star$ is generated as in the previous experiment and the prior information is a blurred version of $x^\star$. More precisely, w is obtained from $x^\star$ by applying a blur filter of order 3: $w_i = \left(x^\star_{i-1} + x^\star_i + x^\star_{i+1}\right)/3$.

[2] L. W. Kang and C. S. Lu, "Distributed compressive video sensing," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2009, pp. 1169–1172.

[3] V. Stanković, L. Stanković, and S. Cheng, "Compressive image sampling with side information," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, Nov 2009, pp. 3037–3040.

[4] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 92–101, March 2008.

[5] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, *Compressive Sensing for Background Subtraction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 155–168.

[6] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," 2005 (revised 2009), http://arxiv.org/abs/0901.3403.

[7] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4595–4607, 2010.

[8] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, "Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models," 2011.

[9] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing Sparsity by Reweighted L1 Minimization," Tech. Rep.

[10] R. Gribonval, V. Cevher, and M. E. Davies, "Compressible distributions for high-dimensional statistics," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5016–5034, Aug 2012.

[11] C. Ravazzi and E. Magli, "Bayesian tuning for support detection and sparse signal estimation via iterative shrinkage-thresholding," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 4628–4632.