Bayesian Compressed Sensing with Unknown Measurement Noise Level

Hansen, Thomas Lundgaard; Jørgensen, Peter Bjørn; Pedersen, Niels Lovmand; Manchón, Carles Navarro; Fleury, Bernard Henri

Published in:
Proc. 47th Asilomar Conference on Signals, Systems and Computers

DOI (link to publication from Publisher):
10.1109/ACSSC.2013.6810248

Publication date:
2013

Document Version
Early version, also known as pre-print

Bayesian Compressed Sensing with Unknown Measurement Noise Level

Thomas L. Hansen, Peter B. Jørgensen, Niels L. Pedersen,
Carles N. Manchón and Bernard H. Fleury,
Dept. of Electronic Systems, Aalborg University, Denmark

Introduction

- In Bayesian compressed sensing (BCS) we apply approximate Bayesian inference to estimate a sparse vector $\mathbf{w} \in \mathbb{C}^M$ from noisy measurements $\mathbf{y} \in \mathbb{C}^N$ taken as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{w} + \mathbf{n},$$

with dictionary $\mathbf{\Phi} \in \mathbb{C}^{N\times M}$ and white Gaussian noise $\mathbf{n} \in \mathbb{C}^N$.

- In most BCS literature it is not tractable to estimate the noise precision λ (inverse variance) as an integral part of the Bayesian inference. Heuristic methods are instead employed resulting in increased computational complexity.
- In this work we propose to modify the three-layer hierarchical prior model in [1] such that the estimation of the noise precision can be included in the inference scheme with no penalty in terms of complexity.

Graphical Models

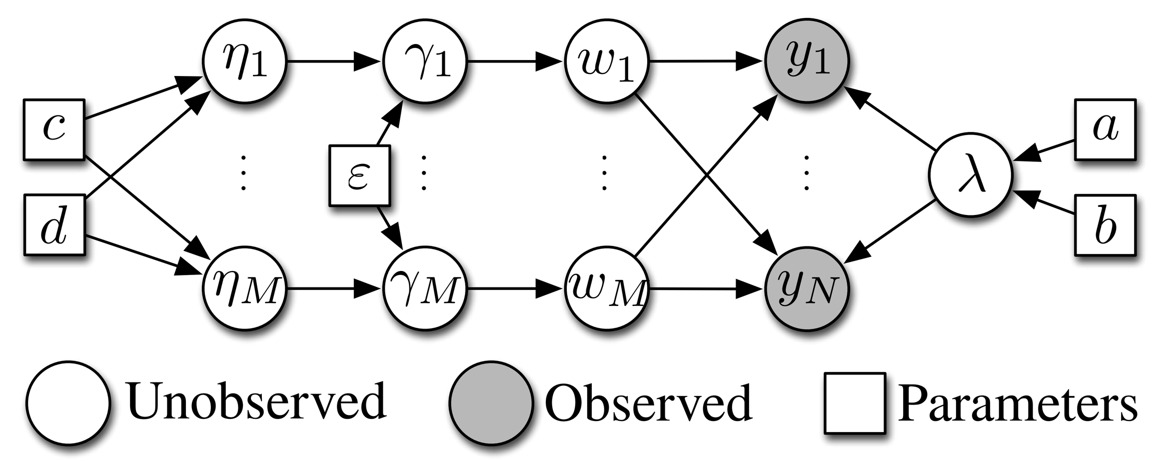


Fig 1: Hierarchical prior model as proposed by Lovmand et. al. [1].

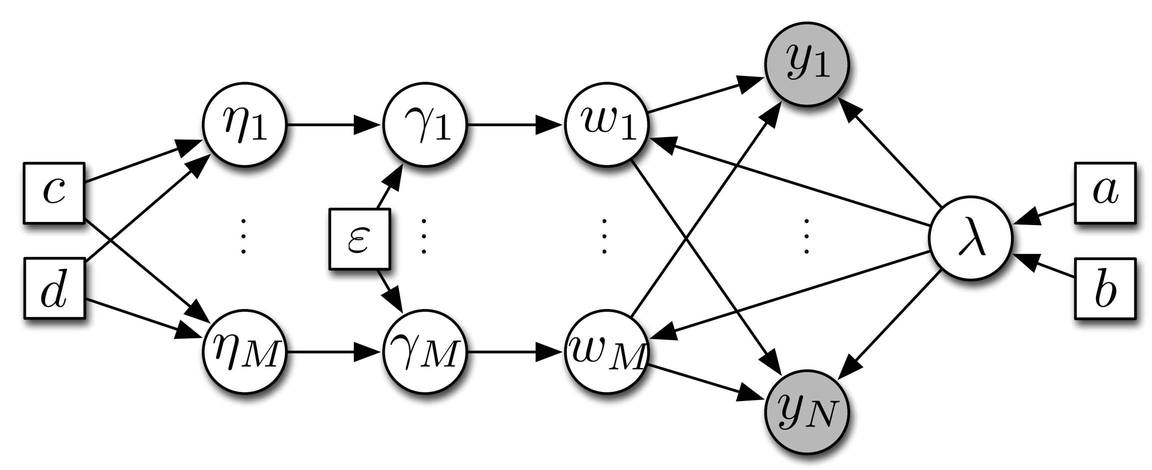


Fig 2: Proposed model with the modification inspired by [2].

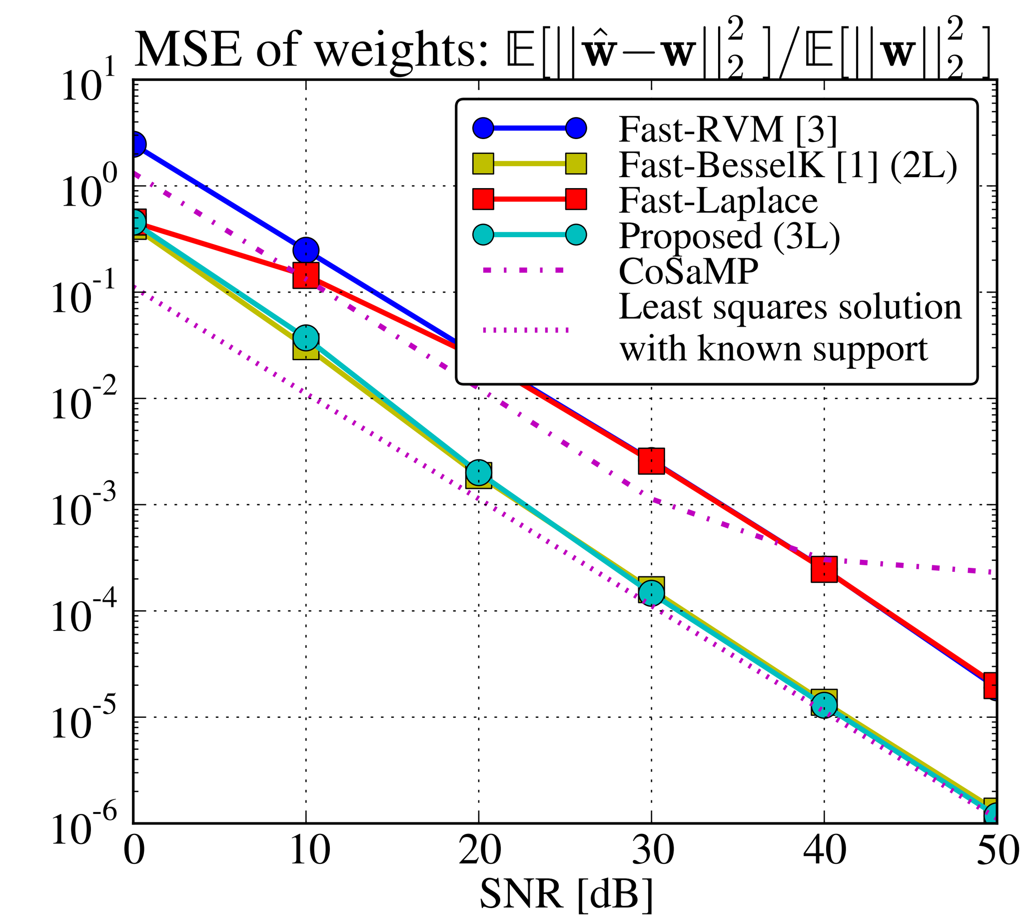
The relevance vector machine (RVM), its fast variant [3] and other algorithms proposed in the literature can be derived from special cases of the model in Fig. 1.

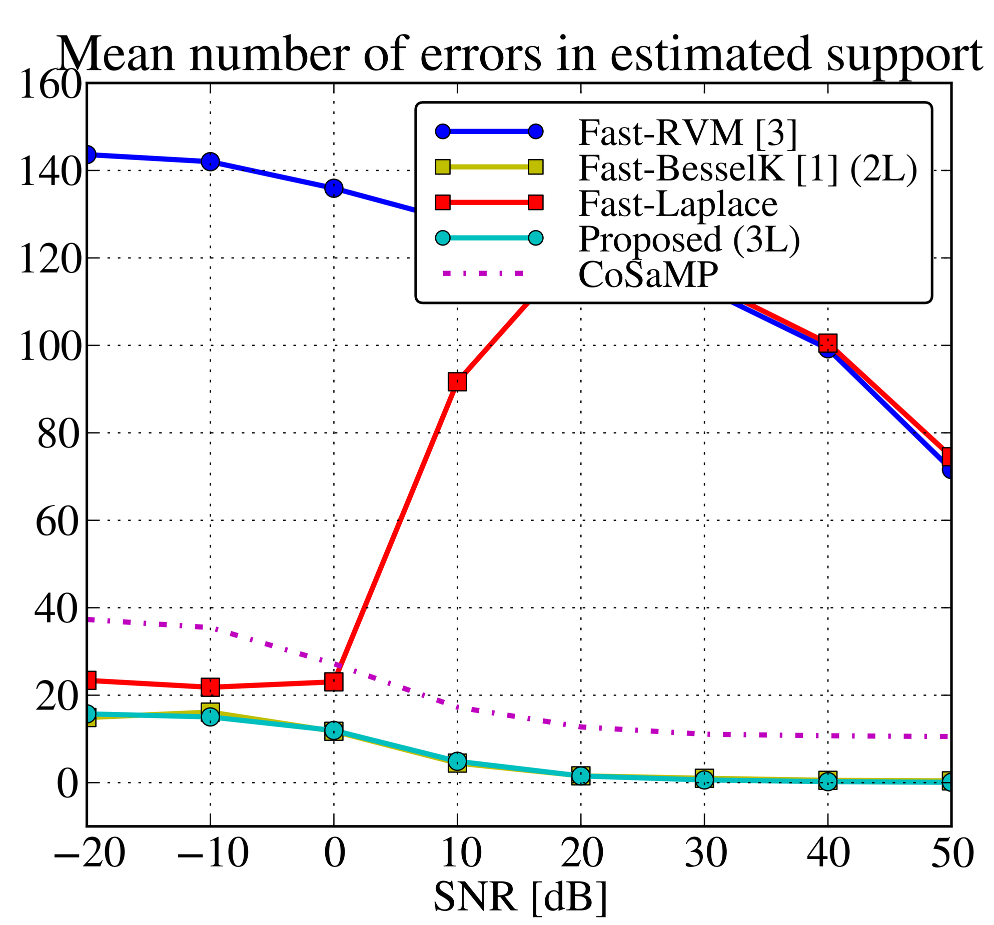
Hierarchical Prior Modelling

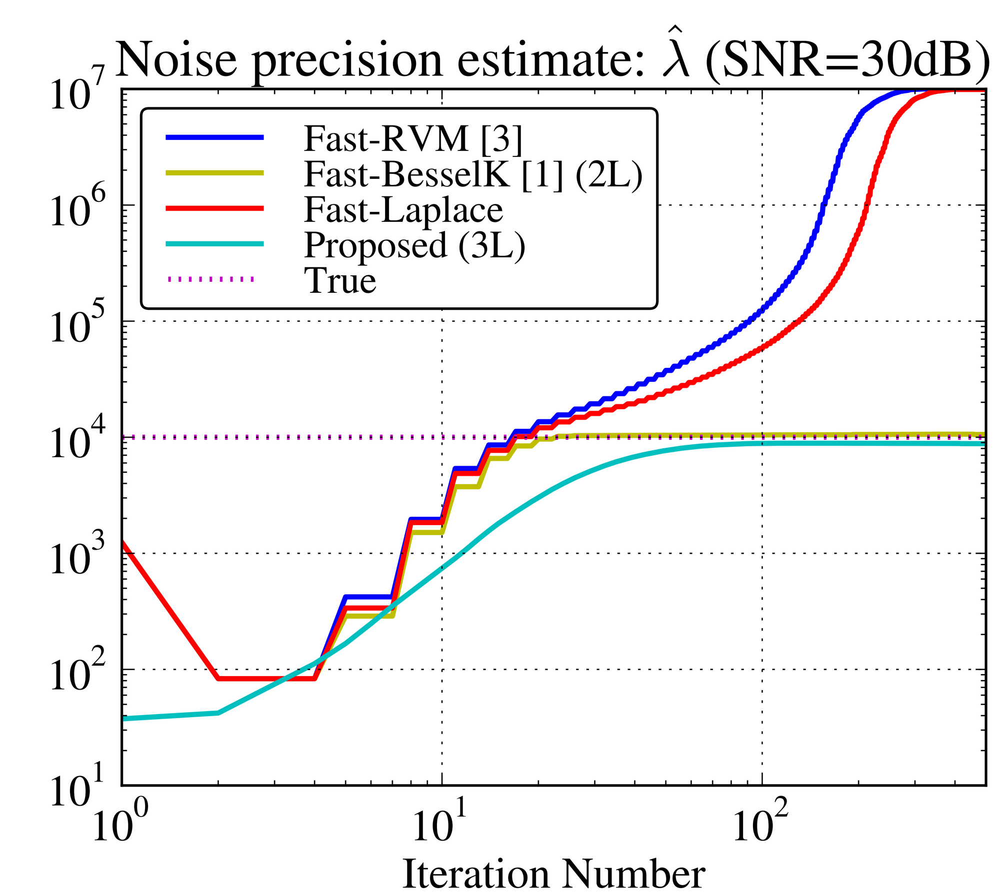
With $\mathbf{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$, the probability distributions in the prior models are:

Density	Model in [1]	Proposed changes
Observation model $p(\mathbf{y}\|\mathbf{w},\lambda)$	$\mathrm{N}(\mathbf{y}\|\mathbf{\Phi}\mathbf{w},\lambda^{-1}\mathbf{I})$	-
Prior on λ, $p(\lambda)$	$\mathrm{Ga}(\lambda\|a,b)$	-
Layer 1 on weights, $p(\mathbf{w}\|\boldsymbol{\gamma})$	$\mathrm{N}(\mathbf{w}\|\mathbf{0},\mathbf{\Gamma})$	$\mathrm{N}(\mathbf{w}\|\mathbf{0},\mathbf{\Gamma}\lambda^{-1})$
Layer 2 on weights, $p(\boldsymbol{\gamma}\|\boldsymbol{\eta})$	$\prod_{i=1}^{M}\mathrm{Ga}(\gamma_i\|\varepsilon,\eta_i)$	-
Layer 3 on weights, $p(\boldsymbol{\eta})$	$\prod_{i=1}^{M}\mathrm{Ga}(\eta_i\|c,d)$	-

Numerical Results







Computational Complexity

- The fast implementation of the RVM [3] can only use its computationally efficient matrix-vector updates when the noise precision estimate is held fixed.
- Using our proposed model, these updates become independent of the noise precision estimate.
- We assume $S \leq N \leq M$, with N compressed measurements, M basis vectors in the dictionary and S nonzero entries in $\mathbf{w}$.
- Computational complexity per iteration:

	Model in [1] & [3]	Proposed
Fixed $\hat{\lambda}$	$\mathcal{O}(MN)$	$\mathcal{O}(MN)$
Updating $\hat{\lambda}$	$\mathcal{O}(MNS)$	$\mathcal{O}(MN)$

References

[1] N. L. Pedersen, D. Shutin, C. N. Manchón, and B. H. Fleury, "Sparse estimation using bayesian hierarchical prior modeling for real and complex models," submitted to *IEEE Transactions on Signal Processing*, 2013, arXiv:1108.4324.

[2] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 92–106, Jan. 2009.

[3] M. E. Tipping and A. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, 2003, pp. 3–6.