

# End of Year Report

Tom Kealy

September 16, 2015

## 1 Summary

- Extending the DADMM algorithm to composite functions. We extended the framework of [5] to functions of the form  $l(x) + r(x)$  by applying making use of dummy variables at each node. This led to a considerably faster algorithm, and allowed many results from centralised algorithms to be applied directly.
- Providing new, clearer, proofs of the edge decomposition of the Lagrangian from [5].
- De-noising OFCOM spectrum data. We have applied centralised ADMM to data supplied by OFCOM, and found that Haar wavelets did make an appropriate model for the signal.
- Extending the Modulated Wideband Converter to model signals composed of rectangles. This has not been a straightforward success - we have used various transforms to sprasify the signal (including Haar Wavelet and discrete derivative). we have had only limited success with the convergence of DADMM to Haar wavelets, but it does perform well with discrete derivatives and spline coefficients.

## 2 Results

We have chosen a distributed model, based upon the Modulated Wideband Converter [4] [6],

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w} \quad (2.0.1)$$

where  $\mathbf{y}$  contains the output of the measurement process, and  $\mathbf{A}$  is a product matrix of the mixing functions, their Fourier coefficients, a partial Fourier Matrix, and a matrix of channel coefficients.  $\mathbf{x}$  is the vector of unknown samples of  $x(t)$ .

i.e.  $\mathbf{A}$  can be written:

$$\mathbf{A}^{m \times L} = \mathbf{S}^{m \times L} \mathbf{F}^{L \times L} \mathbf{D}^{L \times L} \mathbf{H}^{L \times L} \quad (2.0.2)$$

The system 2.0.1 can then be solved (in the sense of finding the sparse vector  $\mathbf{x}$  by convex optimisation via minimising the objective function:

$$\frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (2.0.3)$$

where  $\lambda$  is a parameter chosen to promote sparsity. Larger  $\lambda$  means sparser  $\mathbf{x}$ .

We model the network of sensors as an undirected graph  $G = (V, E)$ , where  $V = \{1 \dots J\}$  is the set of vertices, and  $E = V \times V$  is the set of edges. An edge between nodes  $i$  and  $j$  implies that the two sensors can communicate. The set of nodes that node  $i$  can communicate with is written  $\mathcal{N}_i$  and the degree of node  $i$  is  $D_i = |\mathcal{N}_i|$ .

Individually nodes make the following measurements (as discussed in section ??):

$$\mathbf{y}_p = \mathbf{A}_p \mathbf{x} + \mathbf{n}_p \quad (2.0.4)$$

where  $\mathbf{A}_p$  is the  $p^{th}$  row of the sensing matrix from (2.0.1), and the system (2.0.1) is formed by concatenating the individual nodes' measurements together.

We assume that a proper colouring of the graph is available: that is, each node is assigned a number from a set  $C = \{1 \dots c\}$ , and no node shares a colour with any neighbour. This is so that nodes may communicate in colour order, as opposed to communicating individually thus reducing the total number of communication rounds required.

To find the minimum of 2.0.3, we used the following program:

**Theorem 1.**

$$x_j^{k+1} := (A_j^T A_j + (\rho D_j + 1)I)^{-1} (A_j^T y_j + z^k - \nu^{kT}) \quad (2.0.5)$$

$$z_j^{k+1} := S_{\beta/\rho} (x_j^{k+1}) \quad (2.0.6)$$

$$\theta_j^{k+1} := \theta_j^k + \rho (x_j^{k+1} - z_j^{k+1}) \quad (2.0.7)$$

$$\eta_j^{k+1} := \eta_j^k + \rho \left( \sum_{m \in N_j} z_m^k - z_j^k \right) \quad (2.0.8)$$

We simulated (2.0.1) with a wideband signal of 201 channels and a network of 50 nodes (i.e. the signal will be sampled at a 1/4 of rate predicted by Nyquist theory). The mixing patterns were generated from iid Gaussian sources (i.e the matrix  $S$  had each entry drawn from an iid Gaussian source). Monte Carlo simulations were performed at SNR values ranging from 5 to 20, and the expected Mean Squared Error (MSE) of solutions of a centralised solver (spgl1) and a distributed solver (ADMM) were calculated over 10 simulations per SNR value. The results can be seen in fig (.2).

The MSE was calculated as follows:

$$\frac{\|Z^k - Z^*\|}{\|Z^*\|} \quad (2.0.9)$$

where  $Z^k$  is the result of the algorithm at iteration  $k$ , and  $Z^*$  is the optimal solution.

These results indicate that for both centralised and distributed solvers, adding noise to the system results in a degrading of performance. Interestingly note, that the distributed solver seems to (slightly) outperform the centralised solver at all SNRs. This is counter-intuitive, as it would be expected that centralised solvers knowing *all* the available information would outperform distributed solutions. We conjecture that the updates described in section (??), take into account differences in noise across the network. The distributed averaging steps, which form the new prior for each node, then penalise updates from relatively more noisy observations. This corroborates observations from [1].

This observation is (partially) confirmed in figure (??), which plots the progress of the centralised and distributed solvers (as a function of iterations) towards the optimum solution. The SNR is 0.5 (i.e the signal is twice as strong as the noise). Note that after around 300 iterations, the MSE of the distributed solver is consistently below that of the centralised solver.

We have also extended the framework of [4] to linear transforms  $z = Tx$ , where  $T$  is a Wavelet transform (Haar), Spline, discrete derivative

### 3 Conclusions

We have demonstrated an alternating direction algorithm for distributed optimisation with closed forms for the computation at each step, and discussed the statistical properties of the estimation.

We have simulated the performance of this distributed algorithm for the distributed estimation of frequency spectra, in the presence of additive (white, Gaussian) and multiplicative (frequency flat) noise. We have shown that the algorithm is robust to a variety of SNRs and converges to the same solution as an equivalent centralised algorithm (in relative mean-squared-error).

We plan to work on larger, more detailed, models for the frequency spectra and to accelerate the convergence via Nesterov type methods to smooth the convergence of the distributed algorithm [3]. Specifically, we seek to dampen the ringing seen in Figure .9

### 4 Further Work

- Inference via non-reconstruction of signals. Following [2], we aim to make inferences about the radio spectrum from compressive measurements without ever performing the inversion. This is done by constructing an appropriate projection matrix and then computing the inner product of the result with a known dictionary. We aim to do this inference over a network, with distributed measurements as in 2.0.1 in the Results section. The main difficulty will be in constructing an appropriate projection matrix, and an appropriate dictionary for the spectrum sensing problem.

- Referees comments. There are still some comments from the Referees report from our submission to GLOBECOM which need to be addressed.

## Appendices

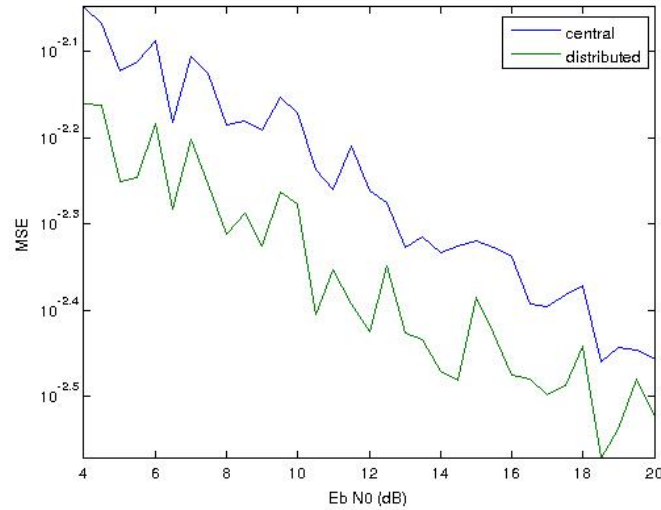


Figure .1: Mse vs SNR for the sensing model, with AWGN only, showing the performance of distributed and centralised solvers

## References

- [1] G B Bazerque, J.A Giannakis. Distributed spectrum sensing for cognitive radios by exploiting sparsity. *Proc. of 42nd Asilomar Conf. on Signals, Systems, and Computers*, 2008.
- [2] Mark Davenport, Petros T Boufounos, Michael B Wakin, Richard G Baraniuk, et al. Signal processing with compressive measurements. *Selected Topics in Signal Processing, IEEE Journal of*, 4(2):445–460, 2010.
- [3] Tom Goldstein, Brendan O’Donoghue, Simon Setzer, and Richard Baraniuk. Fast alternating direction optimization methods. *SIAM Journal on Imaging Sciences*, 7(3):1588–1623, 2014.
- [4] Moshe Mishali and Yonina C Eldar. From theory to practice: Sub-nyquist sampling of sparse wideband analog signals. *Selected Topics in Signal Processing, IEEE Journal of*, 4(2):375–391, 2010.

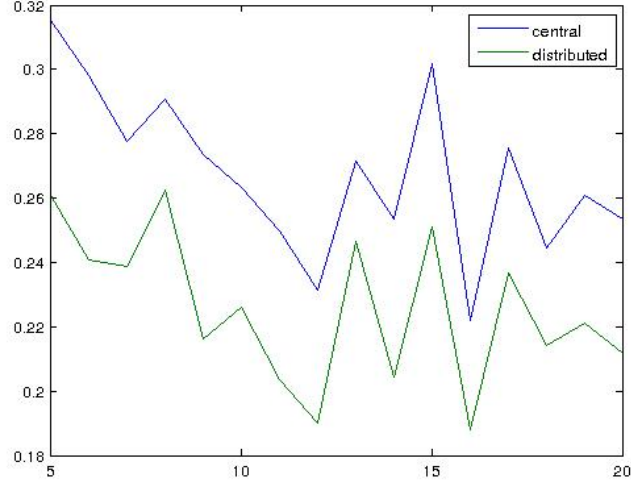


Figure .2: Mse vs SNR for the sensing model, showing the performance of distributed and centralised solvers

- [5] João FC Mota, João MF Xavier, Pedro MQ Aguiar, and Markus Puschel. D-admm: A communication-efficient distributed algorithm for separable optimization. *Signal Processing, IEEE Transactions on*, 61(10):2718–2723.
- [6] Huazi Zhang, Zhaoyang Zhang, and Yuen Chau. Distributed compressed wideband sensing in Cognitive Radio Sensor Networks. In *2011 IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2011*, pages 13–17, 2011.

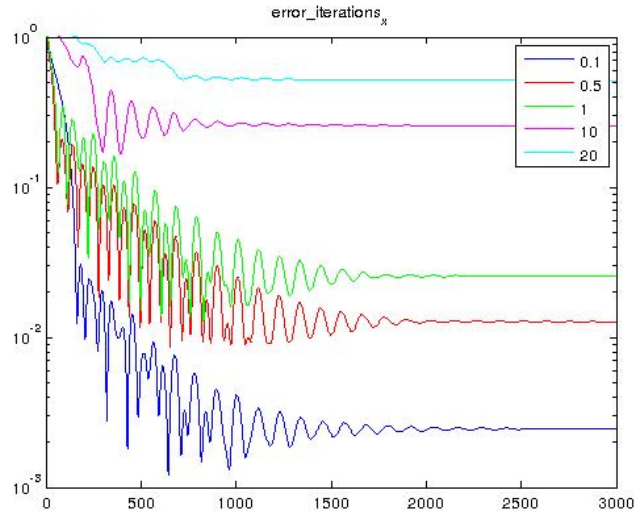


Figure .3: The progress of the distributed solver as a function of the number of iterations, with different values of the regression parameter  $\lambda$

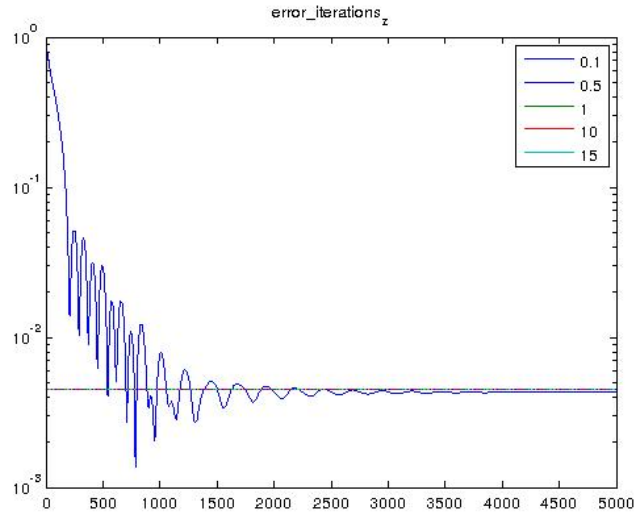


Figure .4: The progress of a distributed (blue) and a centralised (green) solver as a function of the number of iterations. The value of  $\lambda = 0.1$

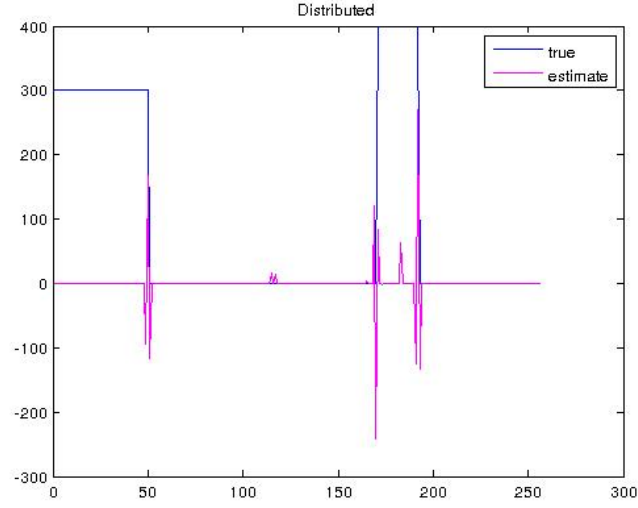


Figure .5: The reconstruction of spline coefficients from compressive measurements (cyan) against the true spectrum.

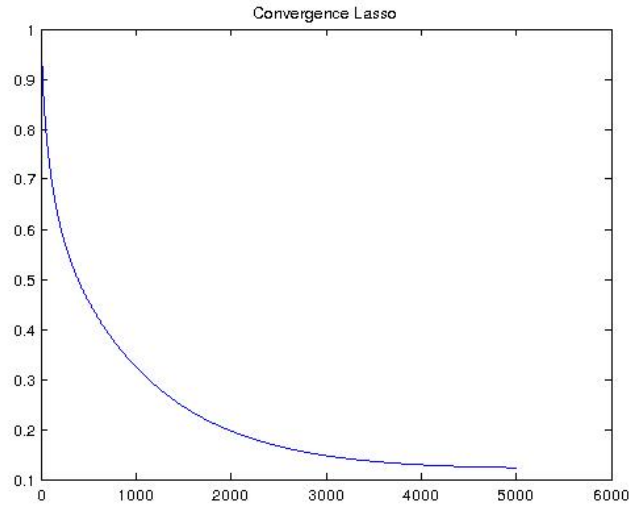


Figure .6: The progress of a distributed solver reconstructing Haar wavelet coefficients as a function of the number of iterations. The value of  $\lambda = \sqrt{40 * \log 256}$

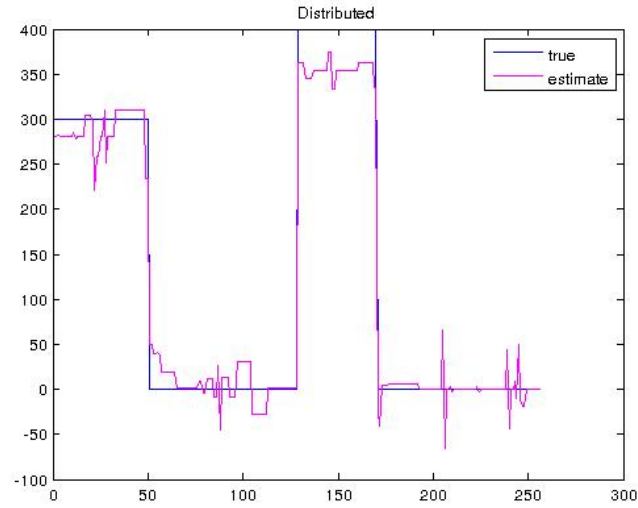


Figure .7: One reconstruction of a simulated spectra from estimated Haar wavelet coefficients.

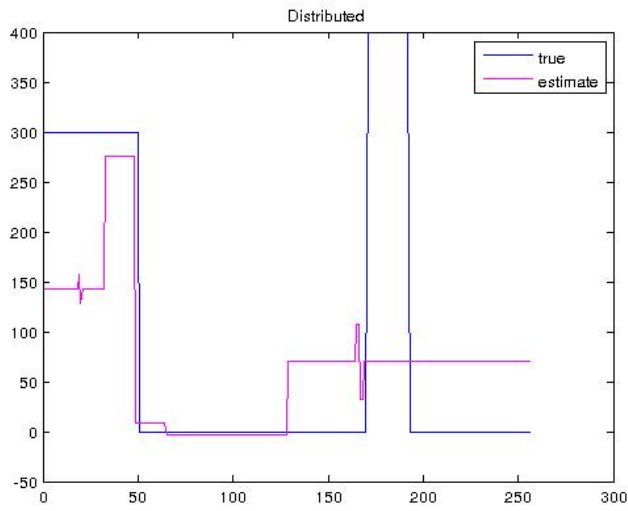


Figure .8: Another reconstruction of a simulated spectra from estimated Haar wavelet coefficients. This spectra does not have a band edge at a power of 2, and so the reconstruction is qualitatively worse.



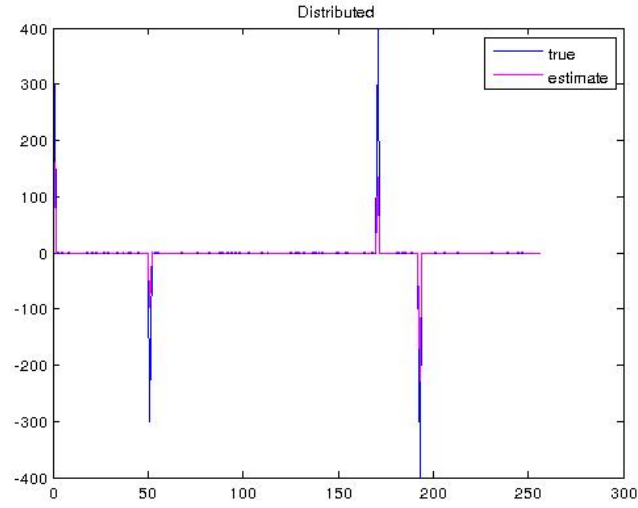


Figure .9: Reconstruction of the discrete derivative of the a simulated signal. Original signal in blue and reconstruction in cyan.

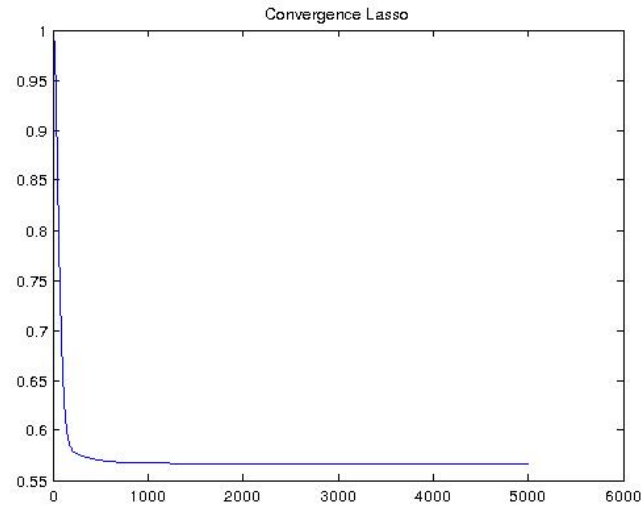


Figure .10: The progress of a distributed solver reconstructing the discrete derivative as a function of the number of iterations. The value of  $\lambda = 10\sqrt{2 * \log 256}$

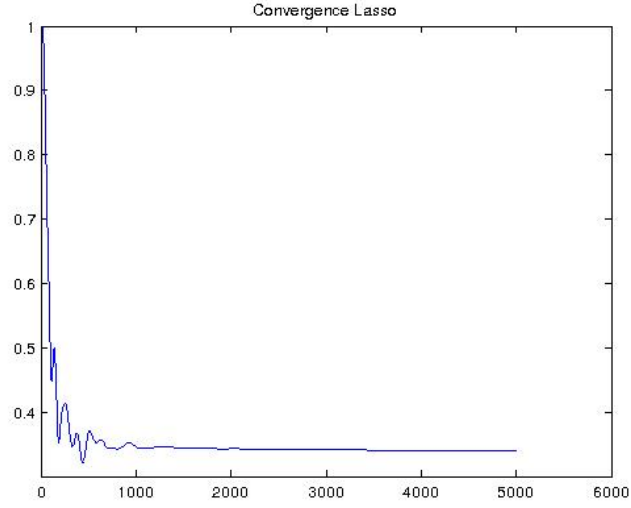


Figure .11: The progress of a distributed solver reconstructing the spline coefficients as a function of the number of iterations. The value of  $\lambda = 2.5\sqrt{2} * \log 256$

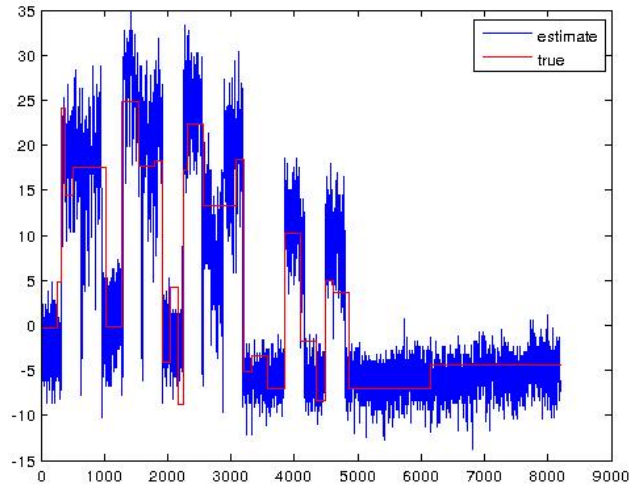


Figure .12: The de-noising of Ofcom spectrum data. In blue is the noisy data, and red the reconstruction from Haar wavelets.