Bachelor project:
Hard thresholding pursuit for sparse approximation

Supervisor: Karin Schnass

A useful concept that allows for efficient signal processing is that the signals in question are sparse in an (overcomplete) dictionary. A dictionary Φ is defined as a collection of K unit norm vectors $\phi_k \in \mathbb{R}^d$ called atoms. The atoms are stacked as columns in a matrix, which by abuse of notation is also referred to as the dictionary, that is $\Phi = (\phi_1, \ldots, \phi_K) \in \mathbb{R}^{d \times K}$. A signal $y \in R^d$ is then called sparse in a dictionary Φ if up to a small approximation error or noise η it can be represented as linear combination of a small (sparse) number of dictionary atoms,

$$y = \sum_{k \in I} \phi_k x_k + \eta = \Phi_I x_I + \eta \quad \text{or} \quad y = \Phi x + \eta \quad \text{with} \quad \|x\|_0 = |I| = S, \tag{1}$$

where $\|\cdot\|_0$ counts the non zero components of a vector or matrix. The index set I storing the non zero entries is called the support with the understanding that for the sparsity level $S = |I|$ we have $S \ll d \leq K$ and that $\|\eta\|_2 \ll \|y\|_2$ or even better $\eta = 0$.
While having an S-sparse approximations is useful for storing signals - store S values and S addresses instead of d values - or for denoising signals - throw away η, looking through $\binom{K}{S}$ possible index sets to find this best S-sparse approximation is certainly not practical. Therefore researchers have developed suboptimal but faster approximation routines, such as thresholding, (Orthogonal) Matching Pursuit and the Basis Pursuit Principle, together with conditions when these routines will find a best approximation. For instance if the dictionary is incoherent, that is

$$\max_{i \neq j} |\langle \phi_i, \phi_j \rangle| = \mu \ll 1, \tag{2}$$

then OMP/BP will recover any best S-sparse approximation as long as $S \lesssim \mu^{-1}$. Unfortunately this means that either S or μ have to be unpractically small. Even more unfortunately this bound can be sharp. One way to go around the coherence bound is to assume that the dictionary is a random matrix, which with high probability satisfies a restricted isometry property. Under this additional assumption many algorithms can be shown to work for $S \lesssim d/\log(K)$, and this concept is widely used for compressed sensing. However, for sparse approximation the dictionary is usually fixed and we cannot assume that is satisfies a restricted isometry property.
Another way to characterise the performance of a sparse approximation algorithm is to consider its average performance. That is, if we assume a random model on the sparse signals we want to know how likely an algorithm is to recover the best S-sparse approximation. The only two algorithms/schemes for which such results are known are thresholding, [2], and the Basis Pursuit Principle, [4]. However, thresholding only works for signals whose coefficients have small dynamic range, $\max_{i \in I} |x_i| \approx \min_{i \in I} |x_i|$, while the Basis Pursuit Principle is computationally costly and it is unknown how stable the results are in a noisy setting. The goal of this project is to investigate the average sparse approximation properties of Hard Thresholding Pursuit (HTP), an algorithm, which has theoretical recovery guarantees in the context of compressed sensing, works also for larger dynamic ranges and has computational complexity between thresholding and BP.

Tasks:

- Read [1], implement the HTP algorithm and reproduce the simulation results in the paper.
- Deduce a worst case result for the performance of HTP for sparse approximation.
- Test the performance and stability of the HTP algorithm in the context of average sparse approximation on synthetic data.

* Compare the performance of HTP to thresholding, [2], and OMP, [3], on synthetic data.
* Compare the performance of HTP to thresholding and OMP on image data.
* Derive a theoretical statement about the average case performance of HTP.

References

[1] S. Foucart. Hard thresholding pursuit: An algorithm for compressive sensing. *SIAM Journal on Numerical Analysis*, 49(6):2543–2563, 2011.

[2] K. Schnass and P. Vandergheynst. Average performance analysis for thresholding. *IEEE Signal Processing Letters*, 14(11):828–831, 2007.

[3] J.A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, October 2004.

[4] J.A. Tropp. On the conditioning of random subdictionaries. *Applied Computational Harmonic Analysis*, 25(1-24), 2008.