Citation for published version:
Golbabaee, M & Davies, M 2017, 'Cover tree compressed sensing', Paper presented at Signal Processing with Adaptive Sparse Structured Representations (SPARS), 1/06/17.

Publication date:
2017

Document Version
Peer reviewed version

COVER TREE COMPRESSED SENSING

Mohammad Golbabaee, Mike Davies
School of Engineering, University of Edinburgh

***Abstract*—We adopt data structure in the form of cover trees and iteratively apply approximate nearest neighbour searches for fast compressed sensing reconstruction of signals living on discrete manifolds. We study the convergence of our algorithm within a more general class of inexact gradient projection methods designated to solve constrained least-square problems. We show that the projection cost is mainly determined by the intrinsic dimension of the manifold and grows sub-linearly with its population. We apply our results to quantitative MRI compressed sensing and in particular within the Magnetic Resonance Fingerprinting framework. For a similar reconstruction accuracy, we report 2-3 orders of magnitude reduction in computations compared to the standard iterative method using brute-force searches.**

I. INTRODUCTION

Compressed sensing (CS) algorithms adopt efficient signal models to achieve accurate reconstruction given small number of measurements. First order methods in the form of iterative gradient projection are popular for CS recovery i.e. solving the following inverse problem:

$$x^* = \operatorname{argmin}_{x\in\mathcal{M}} \|y - Ax\|_2^2, \tag{1}$$

where $y \approx Ax_0$ denotes the CS measurements related to the true signal x_0 by a linear forward operator A, and the set $\mathcal{M}$ represents the signal model. A proper model should be chosen carefully to efficiently promote the low-dimensional structure of data meanwhile not bringing huge computational burden at the projection step. A suitable trade-off is often hard to achieve when dealing with practical datasets with complex models. Thus for saving in the computations one may alternatively consider an inexact iterative projection algorithm:

$$x^{t+1} = \mathcal{P}_{\mathcal{M}}^{\varepsilon_t}\left(x^t - \mu A^*(Ax^t - y)\right), \tag{2}$$

where ε_t is the level of approximation for calculating the projection step at iteration number t. In the next part we show that if ε_t decays rapidly enough, one can achieve a linear convergence for (2) with a recovery accuracy similar to that of using the exact projection.

II. ITERATIVE INEXACT PROJECTION

Assume A satisfies Bi-Lipschitz property with respect to $\mathcal{M}$ with constants $0 < \alpha \leqslant \beta$ such that $\forall x_1, x_2 \in \mathcal{M}$ it holds

$$\alpha\|x_1 - x_2\|^2 \leqslant \|A(x_1 - x_2)\|^2 \leqslant \beta\|x_1 - x_2\|^2 \tag{3}$$

Theorem 1. *Assume $\beta < 1.5\alpha$ and $2/3\alpha^{-1} < \mu \leqslant \beta^{-1}$, the sequence generated by algorithm* (2) *obeys the following bound:*

$$\|x^t - x_0\|^2 \leqslant \rho^t \left(\|x^{init} - x_0\|^2 + \frac{2}{\mu\alpha}\sum_{i=0}^{i=t}\rho^{-i}\varepsilon_i\right) + \frac{4(1-\rho^t)}{\alpha(1-\rho)}\|e\|^2.$$

Where $\rho = 2\left(\frac{1}{\mu\alpha} - 1\right)$, and $e = \|y - Ax_0\|$.

Remark 1. We achieve a linear convergence (up to a tolerance level determined by e) as long as ε_t geometrically decays to zero. For $\varepsilon_t = o(\rho^t)$ we get the same rate $r = \rho$ as the one bounding the convergence of iterative exact projection (see [1] Theorem.2). For $\varepsilon_t = \Theta(\rho^t)$, we have $r = \rho + \delta$ for a small $\delta > 0$. Finally for $\varepsilon_t = \Theta(\underline{\rho}^t)$ where $\underline{\rho} < \rho$ we have $r = \underline{\rho}$.

III. COVER TREES FOR FAST PROJECTION

In this part we consider compressed sensing of signals living on a compact discrete manifold $\mathcal{M}$ with low intrinsic dimension. The set $\mathcal{M}$ may be a dense collection of samples from a continuous manifold with a non trivial projection. With such formalism the projection step reduces to searching the nearest atom, however (potentially) in a very large dictionary. We address the computational shortcoming of this step by preprocessing $\mathcal{M}$ and form a *cover tree* structure suitable for fast approximate nearest neighbour (NN) searches [2]. Cover tree is a levelled tree whose nodes at different scales form covering nets for $\mathcal{M}$ at multiple resolutions (i.e., coarse-to-fine dyadic coverage levels). This structure hierarchically partitions the metric space and enables using branch-and-bound methods for fast NN search. Meanwhile the tree construction is blind to the explicit structure of data, several key growth properties such as the tree's explicit depth and the number of children per node are characterized by the intrinsic dimension $\dim(\mathcal{M})$ of the manifold. In particular, we have [2, 3]:

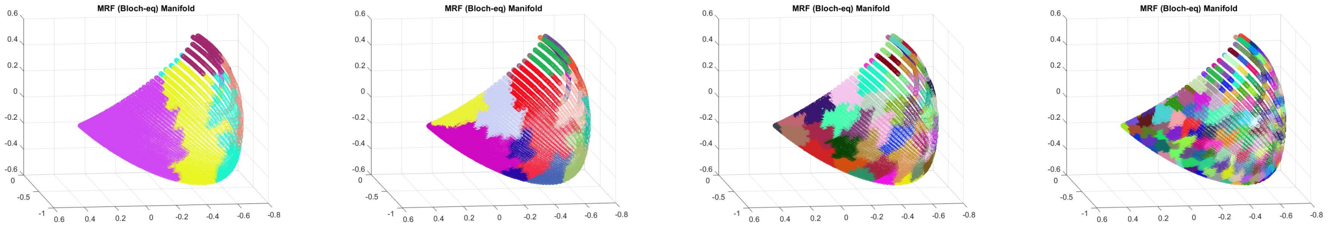
Theorem 2. *The approximate (1+e)-NN search on a cover tree takes at most $2^{O(\dim(\mathcal{M}))}\log\Delta + (1/e)^{O(\dim(\mathcal{M}))}$ computations in time, where Δ is the aspect ratio of $\mathcal{M}$.*

In practical datasets $\Delta \approx n := \operatorname{Card}(\mathcal{M})$ and thus the approximate search complexity grows sub-linearly with manifold population (as opposed to the brute-force search). We use cover tree's (1+e)-NN approximate search to efficiently implement the projection step of algorithm (2). We link between the results of Theorems 1 and 2 and discuss that for a suitable constant e one can achieve enough fast decay in the approximation level ε_t and thus, accurately solve problem (1) with small computational effort. Note that forming the cover tree takes $O(n\log(n))$ computations and it may be computed only once for a given application.

IV. APPLICATION IN MR FINGERPRINTING

Recently MR Fingerprinting [4] proposed a fast CS acquisition scheme for quantifying the NMR properties (namely the $T1, T2$ relaxation times) of tissues. Small number of excitations in form of rotating the magnetic field applies to the tissue, and between each two excitations, the *partial* k-space information is measured. An iterative exact projection algorithm (BLIP) is proposed for the MRF problem which achieves a great parameter estimation accuracy [5]. The projection consists of NN searches on a densely sampled manifold $\mathcal{M}$ of fingerprints driven by the Bloch dynamic equations.

We investigate application of our proposed scheme for enhancing the computational cost of the MRF reconstruction (as we know that the fingerprints are driven from a low-dimensional manifold i.e., here parametrized by two values $T1$, $T2$). We construct a cover tree on a dictionary composed of $n = 48682$ fingerprints and apply the (1+e)-NN tree search for different values of e (we refer to this algorithm as BLIP-CT). We compare the performances of BLIP and BLIP-CT on a synthetic brain phantom with six $(T1, T2)$ segments. Figures 1-3 provides more details on this experiment. Results in Figure 3 indicates that BLIP-CT achieves a similar level of accuracy by saving 2-3 orders of magnitude in computations.



(a) Cover tree segments at scale 2 (b) Cover tree segments at scale 3 (c) Cover tree segments at scale 4 (d) Cover tree segments at scale 5

Fig. 1: The Bloch response manifold sampled from 48682 pairs of $T1, T2$ relaxation times ranging between 100-5000 (ms) and 20-1800 (ms), respectively. Samples $\mathcal{M} \in \mathbb{R}^{1024}$ are magnetization responses of the corresponding parameters to a sequence of 1024 excitation pulses, and are shown along the first three principal components. A cover tree is formed on $\mathcal{M}$ composed of 14 scales: (a-d) data partitions are highlighted in different colours and demonstrated for scales 2-5. Low-scale partitions divide into finer segments by traversing down the cover tree i.e. increasing the scale.

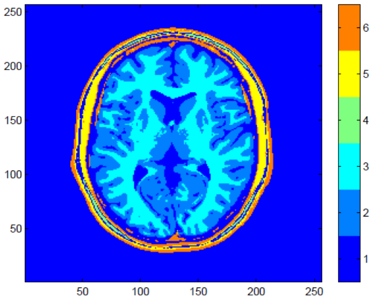
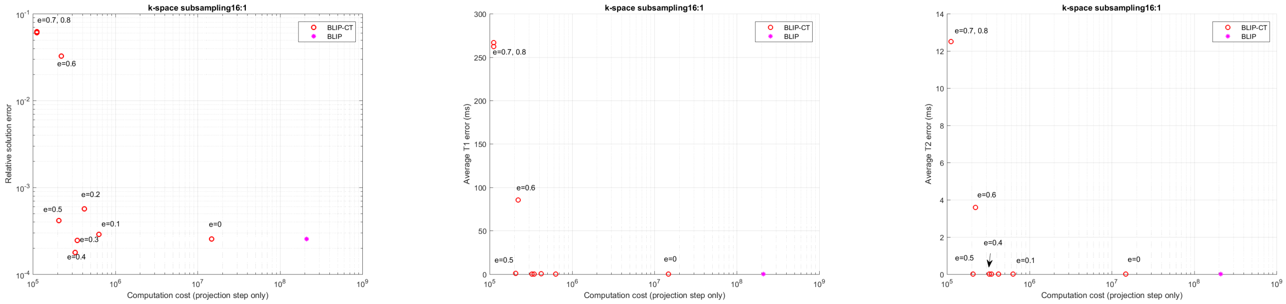


Fig. 2: Anatomical brain phantom. Segments correspond to Background, CSF, Grey Matter, White Matter, Muscle, Skin.



(a) Image reconstruction error vs. computation (b) T1 map accuracy vs. computation (c) T2 map accuracy vs. computation

Fig. 3: Computational cost vs. accuracy comparisons between BLIP and BLIP-CT, tested for brain phantom MRF reconstruction given 16 times subsampled k-space measurements. We only consider cost of the projection steps which is measured by the total number of tree nodes to visit (complexity of the NN searches) per pixel until convergence. As we can observe, BLIP-CT with the choice of parameter $e = 0.5$ achieves an accurate parameter map with about 3 orders of magnitude less computations compared to BLIP.

REFERENCES

[1] T. Blumensath, “Sampling and reconstructing signals from a union of linear subspaces,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4660–4671, July 2011.

[2] A. Beygelzimer, S. Kakade, and J. Langford, “Cover trees for nearest neighbor,” in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 97–104.

[3] R. Krauthgamer and J. R. Lee, “Navigating nets: Simple algorithms for proximity search,” in *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA ’04, 2004.

[4] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. Sunshine, J. Durek, and M. Griswold, “Magnetic resonance fingerprinting,” *Nature*, vol. 495, no. 7440, pp. 187–192, 2013.

[5] M. Davies, G. Puy, P. Vandergheynst, and Y. Wiaux, “A compressed sensing framework for magnetic resonance fingerprinting,” *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2623–2656, 2014.