Adaptive Sampling on Markov Trees

Nikhil Rao*, Gonguo Tang*, Robert Nowak*,
*Department of Electrical and Computer Engineering , University of Wisconsin-Madison

***Abstract*—In this work, we develop a method to adaptively obtain samples from signals that can be modeled as lying on a Hidden Markov Tree. This model finds applications in image acquisition [1], fault detection in sensor networks [2], and disease propagation in social networks, etc. We employ a method based on the Upward Downward algorithm to infer posterior probabilities on Markov Trees. Simulations on toy data show that our method outperforms non adaptive sampling schemes, as well as current methods to sample on trees.**

A. *Method*

Our algorithm uses update equations governed by the Upward-Downward (UD) Algorithm for Hidden Markov Trees [3]. A simplistic version of our algorithm is provided below:

- **Initialize**: Sample the root node and infer its state
- **Iterate until no more samples remain**:
  - Compute the posterior the probabilities of the children of the sampled nodes being active
  - Sample the child that has the highest probability of being active
  - Compute the posterior probability of the sampled node being active based on the update equations in the UD algorithm
- **Return**: Sampled set of nodes, and locations that are estimated to be active.

Our method differs from the UD algorithm, in that once we compute the probability of a node being active, we do not change it. This allows the algorithm to be analyzed easily, and empirically we have observed that the degradation in performance is negligible. Our method also does not assume the sparsity pattern to be a rooted tree, hence allowing us to deal with a richer and more realistic signal model.

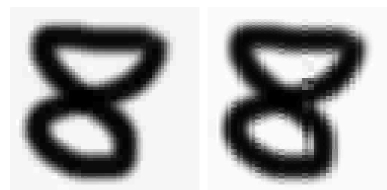
B. *Preliminary Results*

We generated Markov dyadic tree-sparse signals of length $n = 4095$ (12 levels in total) such that the probability of a node being active given its parent is active is $\gamma = 0.8$, and given its parent is inactive is $1 - \delta$(varied as in Table I). We proceed on the tree until we get $k = 256$ non-zeros, or reach the end of the tree. We sample 1000 nodes, keeping the magnitude of the active locations to be $\mu = 1$, and corrupting each measurement with AWGN $\sigma = 0.1$. We compared our method to non adaptive sensing, using the ℓ_1 penalty and i.i.d. Gaussian sensing vectors of unit norm, and to the method by [1], which assumes the sparsity pattern to lie on a rooted tree. We see that, when the pattern is not a rooted tree, our method yields superior performance (Table I). Also, since the ℓ_1 penalized method does not depend on the structure, the performance does not change as we vary δ

We also test our method on the adaptive imaging case, the result of which is shown in Fig 1.

TABLE I
EFFECT OF VARYING δ. THE NUMBERS REPRESENT AVERAGE HAMMING DISTANCE, AND THE NUMBERS IN PARENTHESES ARE THE STD. DEVIATIONS OVER 100 TESTS.

δ	ℓ_1	**Rooted Tree**	**Our Method**
1	0.1479(0.0676)	$\mathbf{0(0)}$	$\mathbf{0(0)}$
0.8	0.1451(0.0632)	0.0367(0.0133)	$\mathbf{0.0197(0.0082)}$
0.6	0.1478(0.0633)	0.0402(0.0119)	$\mathbf{0.0209(0.0083)}$
0.4	0.1478(0.0620)	0.0430(0.0124)	$\mathbf{0.0214(0.0070)}$



(a) original image (b) Recovery using our method

Fig. 1. Image Recovery from 6000 measurements, for a 128×128 dimensional image

C. *Analysis*

For a k sparse signal on a rooted tree, we assume that the variable at each node follows Gaussian distributions $\mathcal{N}(\mu, \sigma^2)$ or $\mathcal{N}(0, \sigma^2)$, depending respectively on whether the node is active or inactive. We establish that we need, as both a necessary and sufficient condition,

$$\mu \approx C\sigma\sqrt{\log(k)}$$

in order to get support recovery with vanishing probability of error, the same as [1]

In the extended version of this work, we include detailed descriptions of the algorithm, perform extensive numerical experiments on real world data, and lay out the full proofs of our theoretical results. We will also analyze the performance of our algorithm for non-rooted trees, and for the case where the node variables follow arbitrary distributions.

REFERENCES

[1] A. Soni and J. Haupt, "Efficient adaptive compressed sensing using sparse hierarchical learned dictionaries," *Asilomar Conference on Signals, Systems and Computers*, 2011.
[2] A. Singh, R. Nowak, and R. Calderbank, "Detecting weak but hierarchically-structured patterns in networks," *arXiv preprint arXiv:1003.0205*, 2010.
[3] R. Nowak, "Multiscale hidden markov models for bayesian image analysis," *Bayesian Inference in Wavelet Based Models, Springer*, 1999.
[4] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian tree structured image modeling using wavelet domain hidden markov models," *Transactions on Image Processing*, March 2000.