DESIGN OF SAMPLING SET FOR BANDLIMITED GRAPH SIGNAL ESTIMATION

Xuan Xie, Hui Feng, Junlian Jia, Bo Hu
Research Center of Smart Networks and Systems, Fudan University, Shanghai 200433, China
Emails: {xxie15, hfeng, jljia10, bohu}@fudan.edu.cn

In this work, we connect the design of sampling set for noisy bandlimited graph signals with experimental design problem, which is closely related to the optimal experimental design problem. The original form of experimental design problem with given measurement size is a combinatorial problem and NP-hard. We use the convex relaxation which tries to obtain an optimal solution of optimal experimental design problem.

We consider the case when any node is allowed to be sampled multiple times. From the view of experimental design, the optimal solution usually achieves when we allow multiple-time sampling. This is very feasible for many real-world networks such as sensor network and social network. We not only decide the nodes to be sampled but also get the proportion of sample for each node of a given sample size by relaxing the combinational problem to convex optimization. In order to get the number of each node to be sampled, we use probabilistic quantization to unbiased quantify the solution of the relaxed problem to integers to get a suboptimal solution of the combinational problem. We also take the quantization error as the perturbation and analyze how sample size influence the performance of our algorithm by perturbation analysis. Moreover, we find a lower bound of the sample size to ensure the object function of our algorithm well-defined, which can provide a reference to practical problems. Finally, the performance of our method is analyzed and shown to have smaller reconstruction error and is more robust against additive noise.

The observation model:

$$\mathbf{y}_{\mathcal{S}} = \Psi \mathbf{V}_K \hat{\mathbf{f}}_K + \mathbf{w} = \mathbf{V}_{MK} \hat{\mathbf{f}}_K + \mathbf{w}$$

The best linear unbiased estimation of $\hat{\mathbf{f}}'_K$ from $\mathbf{y}_{\mathcal{S}}$:

$$\hat{\mathbf{f}}'_K = \mathbf{V}_{MK}^{\dagger} \mathbf{y}_{\mathcal{S}}$$

The covariance matrix of estimation error:

$$\mathbf{E} = \mathbb{E}[\mathbf{e}\mathbf{e}^{\mathsf{T}}] = \left(\mathbf{V}_{MK}^{\mathsf{T}} \mathbf{V}_{MK}\right)^{-1}$$

The scalarization of error covariance matrix:

$$f(\mathbf{E}) = \log\det\left(\left(\sum_{i=1}^{N} m_i \mathbf{u}_i \mathbf{u}_i^{\mathsf{T}}\right)^{-1}\right)$$

The basic experimental design problem:

$$\begin{aligned} \underset{m_i}{\text{minimize}} \quad & f(\mathbf{E}) \\ \text{subject to} \quad & m_i \geq 0, \quad m_1 + \cdots + m_N = M \\ & m_i \in \mathbf{Z} \end{aligned}$$

Related experimental design problem:

$$\begin{aligned} \underset{p_i}{\text{minimize}} \quad & f(\mathbf{E}) = \log\det\left(\left(\sum_{i=1}^{N} p_i \mathbf{u}_i \mathbf{u}_i^{\mathsf{T}}\right)^{-1}\right) \\ \text{subject to} \quad & \mathbf{p} \succeq 0, \quad \mathbf{1}^{\mathsf{T}} \mathbf{p} = 1 \end{aligned}$$

Quantization:

$$Q(p_i) = \begin{cases} \dfrac{k}{M} & \text{with probability } (\dfrac{k+1}{M} - p_i)M \\ \dfrac{k+1}{M} & \text{with probability } (p_i - \dfrac{k}{M})M \end{cases}$$

Theorem 1. Suppose that $\Delta p_1, \ldots, \Delta p_N$ are independent, and each $Q(p_i)$ is obtained by the quantization above. The probability of that $\hat{\mathbf{A}}$ is nonsingular is:

$$P(\|\delta \mathbf{A}\|_2 < \sigma_{\min}(\mathbf{A})) > \prod_{i=1}^{N}\left(1 - \frac{\mathrm{Var}[\Delta p_i]}{(\sigma_{\min}(\mathbf{A}))^2}\right)$$

Where $\delta\mathbf{A} = \hat{\mathbf{A}} - \mathbf{A} = \sum_{i=1}^{N}(Q(p_i) - p_i)\mathbf{u}_i\mathbf{u}_i^{\mathsf{T}} = \sum_{i=1}^{N} \Delta p_i \mathbf{u}_i \mathbf{u}_i^{\mathsf{T}}$ is the perturbation.

Corollary 1. $\hat{\mathbf{A}} = \sum_{i=1}^{N} Q(p_i)\mathbf{u}_i\mathbf{u}_i^{\mathsf{T}}$ is invertible with probability η if the same size M satisfies:

$$M \geq \left\lceil \left(\frac{5}{192(1-\sqrt[N]{\eta})(\sigma_{\min}(\mathbf{A}))^2}\right)^{\frac{1}{3}} \right\rceil$$

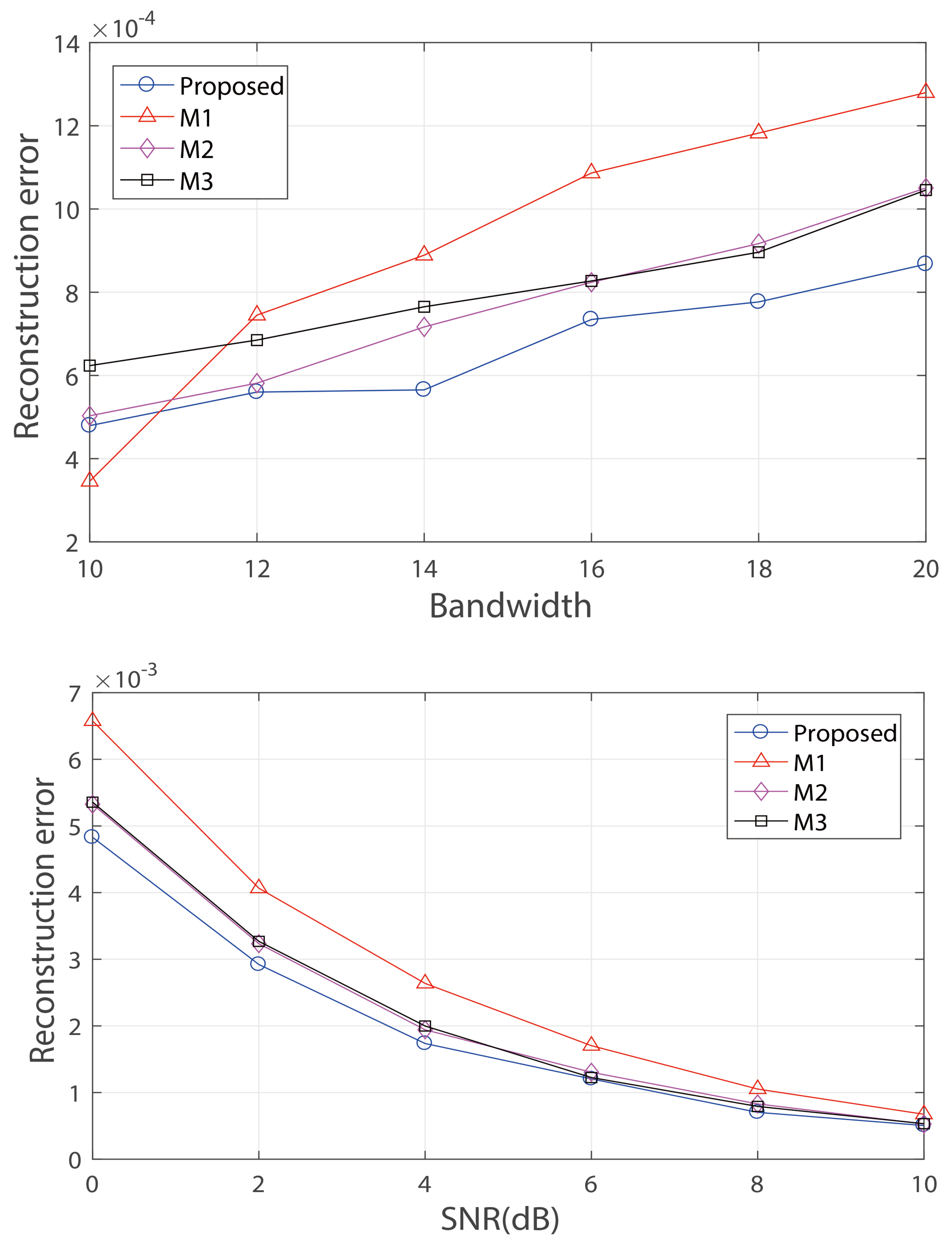

Fig: Reconstruction results of random geometric graph for different signals.