Sampling Rate Distortion: Global Inference from Partial Measurements

V. Praneeth Boda and P. Narayan

University of Maryland 1856 — Institute for Systems Research, A. James Clark School of Engineering

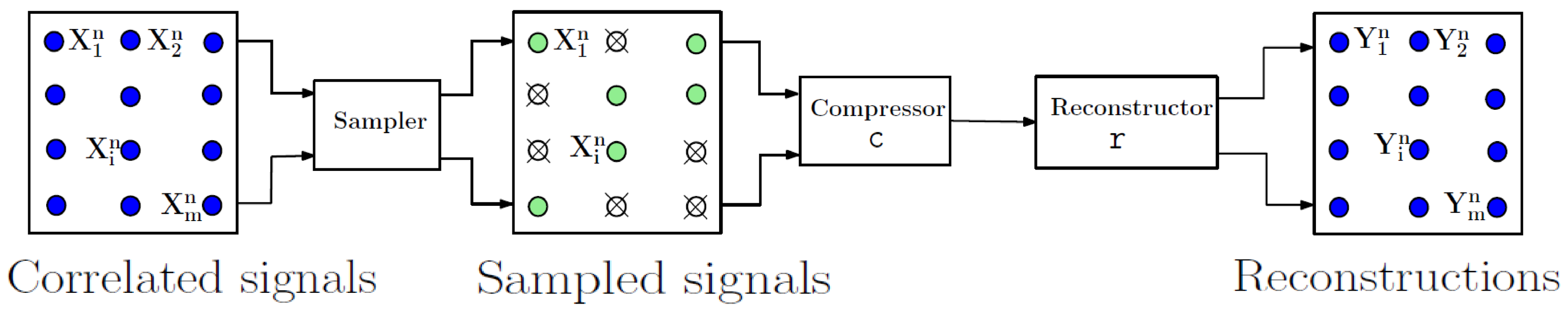
Problems?

- How do we place sensors for measurements? Dynamic thermal management in microprocessors.
- IoT: What can a central hub learn using compressed measurements from smart sensors?
- How do we find "central" nodes in large networks to capture the average behavior?

Reconstructing signals from spatially sampled signals

- Sampling rate distortion: Statistics known
  - Selection and compression of signals. Algorithms for reconstruction.
- Universal sampling rate distortion: Statistics unknown
  - Selection, estimation and compression.
- Multi-armed bandits: Statistics unknown, finite-time horizon
  - Selection, estimation and identification under budget.

- Correlated signals: $\{(X_{1t}, \ldots, X_{mt}) = X_{\mathcal{M}t}\}_{t=1}^{\infty}$
- k-random sampler: $P_{S^n|X_{\mathcal{M}}^n} = \prod_{t=1}^{n} P_{S_t|X_{\mathcal{M}}^t S^{t-1}}$



- Sampling Rate Distortion function:

$$R(\Delta) \triangleq \inf_{\substack{k\text{-RS},\ (\mathtt{c},\mathtt{r}) \\ \mathbb{E}[\mathsf{Distortion}(\mathsf{Signals},\, \mathtt{r}(\mathtt{c}(\mathsf{Sampled\ signals})))] \leq \Delta}} \frac{1}{n} \log ||\mathtt{c}||$$

Global inference when only subsets of signals can be accessed

Sampling Rate Distortion function

- For a non-adaptive random sampler,

$$P_{S_t|X_{\mathcal{M}}^t S^{t-1}} = P_{S_t},\ t = 1, \ldots$$

$$R_{IRS}(\Delta) = \min_{\substack{P_S,\, P_{Y_{\mathcal{M}}|S X_S} \\ \mathbb{E}[d(X_{\mathcal{M}}, Y_{\mathcal{M}})] \leq \Delta}} I(X_S \wedge Y_{\mathcal{M}}|S).$$

- For an adaptive random sampler,

$$P_{S_t|X_{\mathcal{M}}^t S^{t-1}} = P_{S_t|X_{\mathcal{M}t}},\ t = 1, \ldots$$

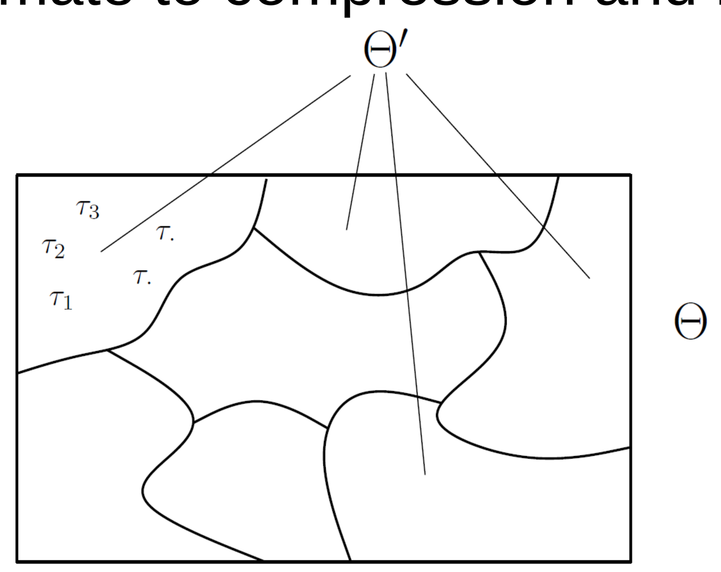
  - Conditional deterministic sampling is optimal!
  - Reduce computational complexity.
- Memory of previously sampled signals doesn't help.
- Does knowledge of sampling at reconstructor help?

What if statistics are unknown?

- Distribution of $\{(X_{1t}, \ldots, X_{mt}) = X_{\mathcal{M}t}\}_{t=1}^{\infty}$ known only to

$$P_{X_{\mathcal{M}}} \in \mathcal{P} = \{P_\tau,\ \tau \in \Theta\}$$

- Compressor:
  - Forms an "estimate" of statistics from partial observations
  - Cannot learn all correlations
  - Uses estimate to compression and reconstruction.

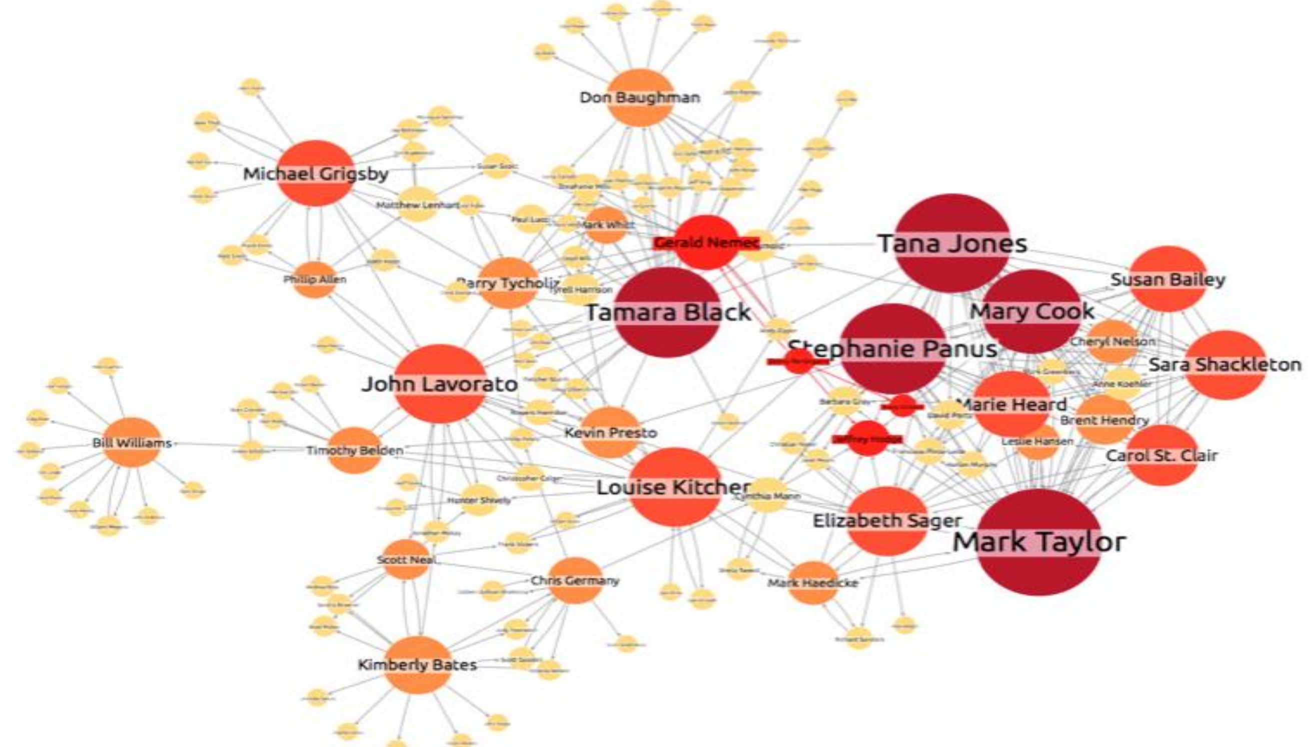


Ambiguity atoms over class of distributions

Identify "central" nodes in networks

(Joint work with Prashanth L. A.)

- Capture *average* behavior in communication/social networks.
- Statistics unknown; data acquisition expensive.
- Observations only from subsets of nodes/people.
- Sequential decision algorithms require less data.



Thanks to NSF - CCF

Acknowledgements: Ankur Srivastava, Himanshu Tyagi