DISTRIBUTED COMPRESSION AND MAXIMUM LIKELIHOOD RECONSTRUCTION OF FINITE AUTOCORRELATION SEQUENCES

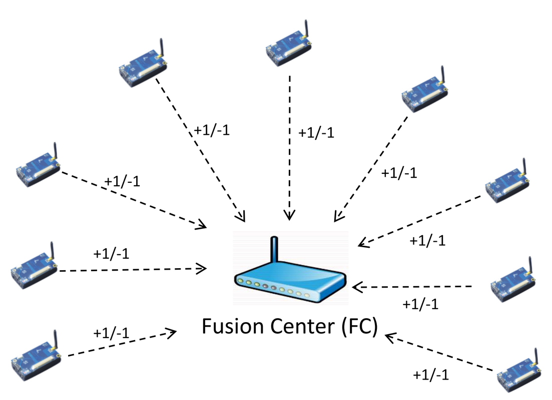
Aritra Konar and Nicholas D. Sidiropoulos
University of Minnesota, USA

University of Minnesota Driven to Discover

Introduction

- **Goal:** Devise a distributed sensing framework for compressing and reconstructing finite-length autocorrelation sequences of WSS processes.
    i. Employ network of sensors to transmit randomly filtered, single-bit quantized power measurements to fusion center.
    ii. Reconstruct autocorrelation sequence at fusion center from binary power measurements.
- **Contributions:**
    i. Maximum-Likelihood (ML) based reconstruction scheme proposed, vs. min. norm in previous work. ML is statistically more efficient, robust, but looks hard to solve.
    ii. Exploiting special problem structure, it is shown that the problem possesses **hidden convexity,** enabling optimal estimation. Extends hidden convexity result of previous work to ML case.

System Model [Mehanna *et al.*, 2013]



- Network of M scattered sensors with limited communication capabilities
- Sensor m equipped with FIR filter with random impulse response

$$g_m(n) = \begin{cases} \sim \mathcal{U}(\{1+j, 1-j, -1+j, -1-j\}) & : n \in [0, K-1] \\ 0 & : \text{otherwise} \end{cases}$$

- At i^{th} sampling instance, acquire data sample vector

$$\mathbf{x}_m^{(i)} = [x_m(i), x_m(i-1), \cdots, x_m(i-K+1)]^T \in \mathbb{C}^K$$

- Obtain random linear projections $z_m^{(i)} = \mathbf{g}_m^H \mathbf{x}_m^{(i)}$ with average power

$$\rho_m := \mathbb{E}[|z_m^{(i)}|^2] = \mathbb{E}[|\mathbf{g}_m^H \mathbf{x}_m^{(i)}|^2] = \mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m$$

- Soft power estimates: $\rho_m^{(N)} := \frac{1}{N}\sum_{n=1}^{N} |z_m^{(i)}|^2$
- Error due to insufficient sample averaging: $e_m := \rho_m - \rho_m^{(N)}$
- Errors approximately Gaussian by CLT, i.e., $e_m \sim \mathcal{N}(0, \sigma_m^2)$
- One bit quantization: $b_m = \text{sign}(\rho_m^{(N)} - t) = \text{sign}(\rho_m + e_m - t)$
- Errors may result in flipped bits

Signal Model

- Finite autocorrelation sequence of length L can be represented as

$$r(k) = \begin{cases} \mathbf{h}^H \mathbf{\Theta}_k^{L+1} \mathbf{h} & : |k| \le L \\ 0 & : |k| > L \end{cases}$$

where $\mathbf{h} \in \mathbb{C}^{L+1}$ and $\mathbf{\Theta}_k^{L+1} \in \mathbb{R}^{(L+1)\times(L+1)}$ has ones on the k^{th} diagonal and zeros elsewhere.

- If $K > L$, $\mathbf{R}_x$ is positive semidefinite,Toeplitz-Hermitian and banded with bandwidth L
- Exploiting the structure of $\mathbf{R}_x$, we have $\rho_m = \mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m = \mathbf{h}^H \mathbf{C}_m \mathbf{h}$

where $\mathbf{C}_m = \sum_{k=-L}^{L} c_m(k) \mathbf{\Theta}_k^{L+1}$ and $c_m(k) = \mathbf{g}_m^H \mathbf{\Theta}_k^K \mathbf{g}_m$

- $\mathbf{C}_m$ is positive semidefinite and Toeplitz-Hermitian

Problem Formulation

- Find $\mathbf{h}$ from $\{b_m\}_{m=1}^M$ utilizing Gaussian distribution of $\{e_m\}_{m=1}^M$ in a Maximum-Likelihood (ML) formulation
- Define $\mathcal{M}_a = \{m | b_m = +1\}$, $\mathcal{M}_b = \{m | b_m = -1\}$
- Likelihood function:

$$f(b_1, \cdots, b_m; \mathbf{h}) = \prod_{m \in \mathcal{M}_a} \Pr(\mathbf{h}^H \mathbf{C}_m \mathbf{h} + e_m \ge t) \prod_{m \in \mathcal{M}_b} \Pr(\mathbf{h}^H \mathbf{C}_m \mathbf{h} + e_m < t)$$

$$= \prod_{m \in \mathcal{M}} \Phi\left(\frac{b_m(\mathbf{h}^H \mathbf{C}_m \mathbf{h} - t)}{\sigma_m}\right)$$

- ML formulation:

$$\boxed{\max_{\mathbf{h}} \ \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\mathbf{h}^H \mathbf{C}_m \mathbf{h} - t)}{\sigma_m}\right)}$$

Possesses hidden convexity

Problem Reformulation

- Note $\mathbf{h}^H \mathbf{C}_m \mathbf{h} = \sum_{k=-L}^{L} c_m(k) \mathbf{h}^H \mathbf{\Theta}_k^{L+1} \mathbf{h} = \sum_{k=-L}^{L} c_m(k) r(k) = \text{Re}(\mathbf{c}_m^T \mathbf{r})$

where $\mathbf{c}_m = [c_m(0), 2c_m(1), \cdots, 2c_m(L)]^T \in \mathbb{C}^{L+1}, \mathbf{r} = [r(0), \cdots, r(L)]^T \in \mathbb{C}^{L+1}$

- Define $\mathbf{Q} = \mathbf{h}\mathbf{h}^H$. Then,

$$\begin{aligned} \max_{\mathbf{Q}, \mathbf{r}} \quad & \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\text{Re}(\mathbf{c}_m^T \mathbf{r}) - t)}{\sigma_m}\right) \\ \text{s.t.} \quad & r(k) = \text{trace}(\mathbf{\Theta}_k^{L+1} \mathbf{Q}), \ k = 0, \cdots, L \\ & \mathbf{Q} \succeq \mathbf{0}, \\ & \text{Rank}(\mathbf{Q}) = 1 \end{aligned}$$

$$\xrightarrow{\text{Relax}}$$

$$\begin{aligned} \max_{\mathbf{Q}, \mathbf{r}} \quad & \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\text{Re}(\mathbf{c}_m^T \mathbf{r}) - t)}{\sigma_m}\right) \\ \text{s.t.} \quad & r(k) = \text{trace}(\mathbf{\Theta}_k^{L+1} \mathbf{Q}), \ k = 0, \cdots, L \\ & \mathbf{Q} \succeq \mathbf{0} \end{aligned}$$

- Rank-1 constraint is redundant [Alkire and Vandenberghe, 2002]
- High rank solution can always be converted to rank-1 solution with same cost via spectral factorization.
- Solution non-unique (i.e., $\mathbf{h}$ is non-identifiable) but all equivalence classes define the same autocorrelation sequence.
- Hence, original ML formulation can be solved to global optimality.

Final Formulation

$$\begin{aligned} \max_{\mathbf{Q}} \quad & \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\text{trace}(\mathbf{C}_m \mathbf{Q}) - t)}{\sigma_m}\right) \\ \text{s.t.} \quad & \mathbf{Q} \succeq \mathbf{0} \end{aligned}$$

- Equivalent to the previous formulation.
- Fewer variables, hence more computationally efficient.
- Construct sequence $r(k) = \text{trace}(\mathbf{\Theta}_k^{L+1} \mathbf{Q})$ followed by spectral factorization to obtain optimal rank-1 solution.

Simulation Results

- MA(7) model, M = 200 sensors, K = 30, 80 sensors transmit b_m=1, random errors resulted in 52 bit flips (26% of measurements), true model order known at FC, 500 Monte-Carlo trials.

Fig 1: Power Spectrum Reconstruction

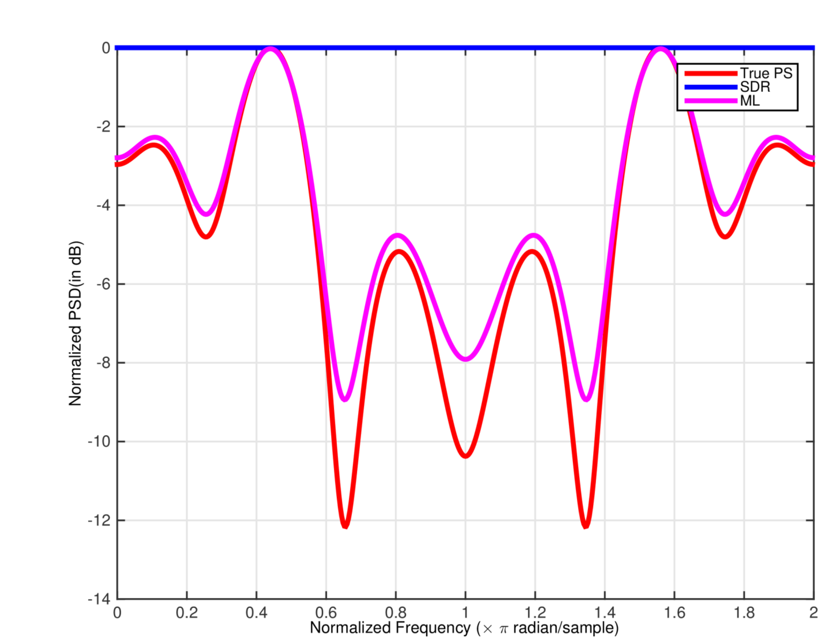
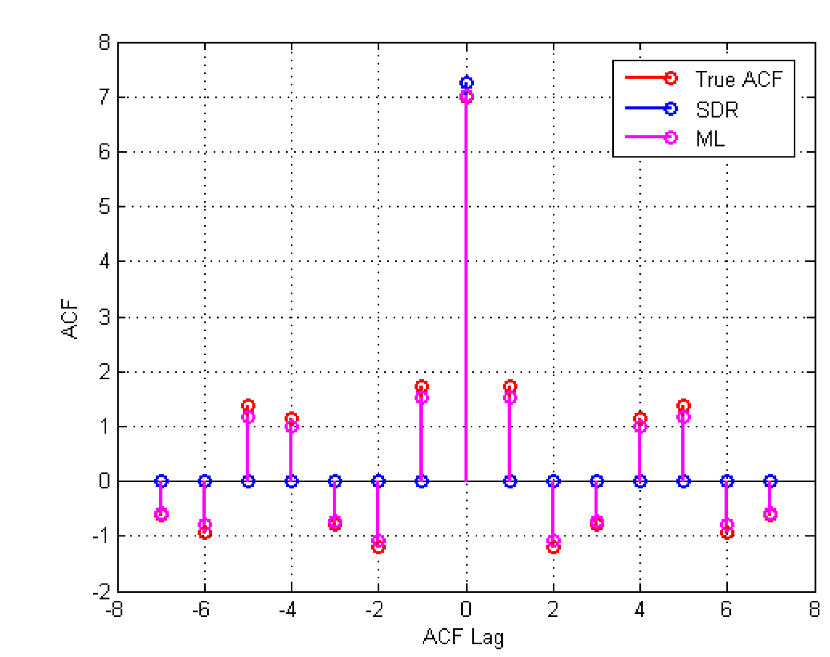


Fig 2: Autocorrelation Function Reconstruction



- Very satisfactory performance, ML formulation robust to bit flips

Future Work:

- Using CRLB for threshold design
- Active sensing version