Frugal Sensing

Spectral Analysis from Power Inequalities

Nikos Sidiropoulos

Joint work with Omar Mehanna

IEEE SPAWC 2013 Plenary, June 17, 2013, Darmstadt, Germany

Wideband Spectrum Sensing (for CR/DSM)

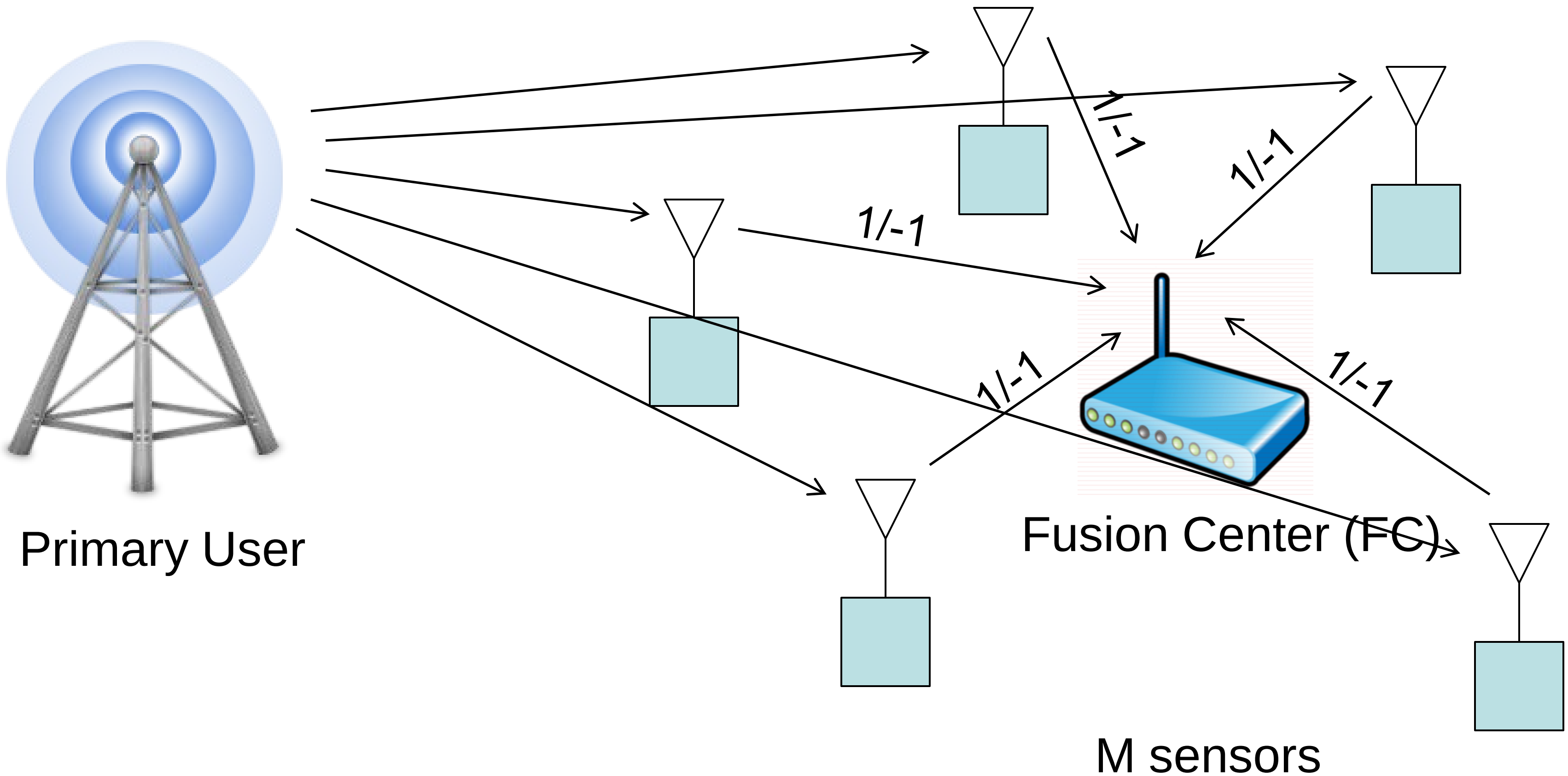
- Split in narrowband channels + channel-by-channel sensing
  ↔ Filterbank / frequency sweeping (hardware / delay), correlation ignored
- Wideband scanning with high-rate ADC
  - Hard to implement, expensive, high power consumption
- Compressive sampling
  - Requires frequency-domain sparsity for sub-Nyquist sampling
- *[Leus et al, '11]:* no need to reconstruct received signal spectrum; *power spectrum* sufficient / more relevant for CR, certain other applications (e.g., radio astronomy)
- Can estimate from FT of truncated autocorrelation → finite parameterization

State-of-the-art

- Power spectrum sensing *[Leus et al, '11]*
  - Neither Nyquist-rate sampling nor full-band scanning is necessary
  - Signal passed through bank of filters → Cross-correlations of outputs are used to build an over-determined system of linear equations in the signal autocorrelation for a finite number of lags
  - Analog amplitude samples not suitable in network sensing setting using low-end sensors with limited communication capabilities

- One-bit compressed sensing *[Boufounos et al, '08]*
  - Signal recovered (within scaling factor) from sign info of compressed measurements
  - Does not exploit additional autocorrelation-specific constraints
  - Requires signal sparsity

Frugal Sensing



Estimate of the power spectrum using few bits
Spectral estimation from inequalities instead of equalities

Outline

1. Problem Formulation
2. Nonparametric Estimation
   - Error-free case
   - Gaussian errors
3. Parametric Estimation (Line Spectra)
4. Adaptive Thresholding (Active FC)
5. Summary

Sensor Measurement Chain

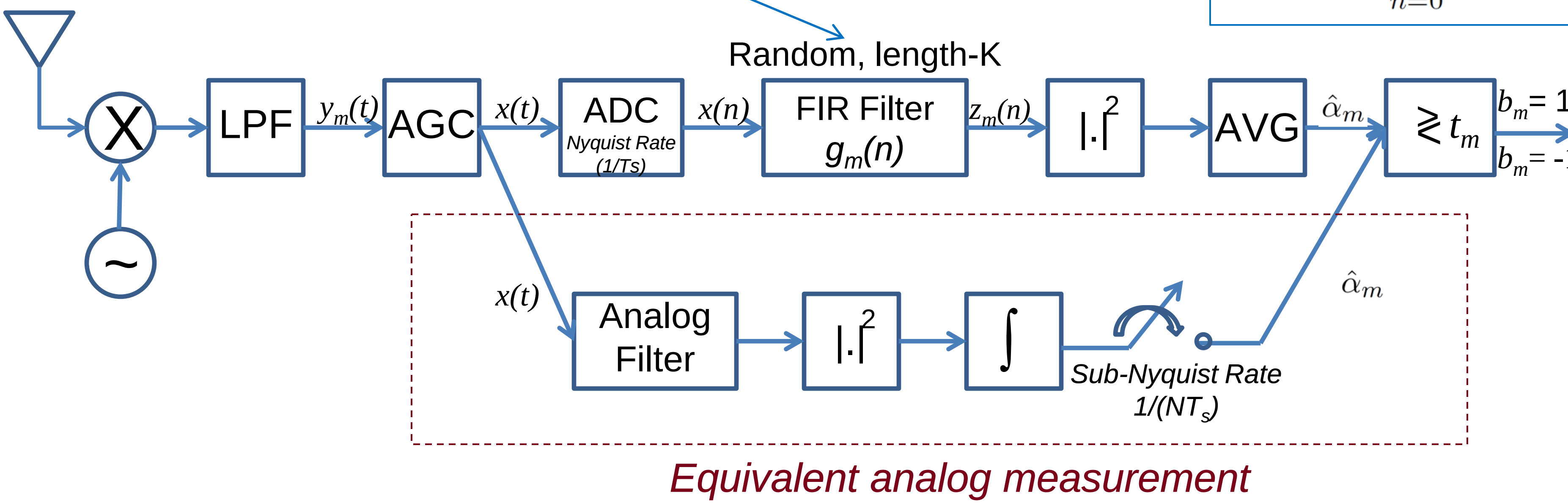
Known at the FC

$$g_m(n) = \begin{cases} (1/\sqrt{2K})(\pm 1 \pm j) & \text{if} \quad 0 \le n \le K-1 \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_m := \mathbb{E}[|z_m(n)|^2]$$

$$\hat{\alpha}_m = \frac{1}{N}\sum_{n=0}^{N-1} |z_m(n)|^2$$

Random, length-K



Fading (1)

- Received (discrete-time) signal

$$\tilde{y}_m(n) = \gamma_m \sum_{\ell=0}^{L-1} h_m(\ell) x(n-\ell) \xrightarrow{\text{AGC}} y_m(n) = \tilde{y}_m(n)/\gamma_m$$

Sensor-specific loss

- *Assumption:* L-tap channel is random, time-invariant, correlation between taps is only function of ordinal distance

$$\mathbb{E}[h_m(\ell_1) h_m^*(\ell_2)] = r_{h_m}(\ell_1 - \ell_2)$$

- Frequency response $H_m(\omega) = \sum_{\ell=0}^{L-1} h_m(\ell) e^{-j\omega\ell}$

- Power

$$\mathbb{E}[|H_m(\omega)|^2] = \sum_{\ell_1=0}^{L-1} \sum_{\ell_2=0}^{L-1} \mathbb{E}[h_m(\ell_1) h_m^*(\ell_2)] e^{-j\omega(\ell_1-\ell_2)}$$

$$= \sum_{\ell=-L+1}^{L-1} (L - |\ell|) r_{h_m}(\ell) e^{-j\omega\ell}$$

Fading (2)

- Received signal autocorrelation

$$\begin{aligned}
\mathbb{E}[y_m(n)y_m^*(n-k)] &= \mathbb{E}\left[\sum_{\ell_1=0}^{L-1} h_m(\ell_1)x(n-\ell_1)\sum_{\ell_2=0}^{L-1} h_m^*(\ell_2)x^*(n-k-\ell_2)\right] \\
&= \sum_{\ell_1=0}^{L-1}\sum_{\ell_2=0}^{L-1} r_{h_m}(\ell_1-\ell_2)r_x(k+\ell_2-\ell_1) \\
&= \sum_{\ell=-L+1}^{L-1}(L-|\ell|)r_{h_m}(\ell)r_x(k-\ell)
\end{aligned}$$

$$r_x(k) := \mathbb{E}[x(n)x^*(n-k)]$$

- PS

$$\begin{aligned}
S_{y_m}(\omega) &= \sum_{k=-\infty}^{\infty}\mathbb{E}[y_m(n)y_m^*(n-k)]e^{-j\omega k} \\
&= \sum_{\ell=-L+1}^{L-1}(L-|\ell|)r_{h_m}(\ell)\sum_{k=-\infty}^{\infty} r_x(k-\ell)e^{-j\omega k} \\
&= \underbrace{\sum_{\ell=-L+1}^{L-1}(L-|\ell|)r_{h_m}(\ell)e^{-j\omega\ell}}_{\mathbb{E}[|H_m(\omega)|^2]}\underbrace{\sum_{\bar{k}=-\infty}^{\infty} r_x(\bar{k})e^{-j\omega\bar{k}}}_{S_x(\omega)}
\end{aligned}$$

Fading (3)

- Consistent power spectrum measurements if
  - $\mathbb{E}[|H_m(\omega)|^2]$ same across all sensors
  - Sensors acquire sufficient samples with different channel realizations

- In practice
  - Sensor periodically senses spectrum encountering new channel realization each time (drift and carrier/phase lock)
  - Reported measurements reflect averaging over many epochs

Power Measurement

- Signal autocorrelation $r_x(\ell) := \mathbb{E}[x(n)x^*(n-\ell)]$
- *Deterministic* filter autocorrelation $q_m(\ell) = \sum_{n=0}^{K-1} g_m(n) g_m^*(n+\ell)$
- Power measurement $\alpha_m = r_{z_m}(0) = \sum_{\ell=1-K}^{K-1} r_x(\ell) q_m^*(\ell) = \mathbf{q}_m^T \tilde{\mathbf{r}}_x$

$$\tilde{\mathbf{r}}_x := [r_x(0), \mathrm{Re}\{r_x(1)\}, \ldots, \mathrm{Re}\{r_x(K-1)\}, \mathrm{Im}\{r_x(1)\}, \ldots, \mathrm{Im}\{r_x(K-1)\}]^T$$
$$\mathbf{q}_m := [q_m(0), 2\,\mathrm{Re}\{q_m(1)\}, \ldots, 2\,\mathrm{Re}\{q_m(K-1)\}, 2\,\mathrm{Im}\{q_m(1)\}, \ldots, 2\,\mathrm{Im}\{q_m(K-1)\}]^T$$

- Power spectrum estimate $\hat{\mathbf{s}}_x = \overbrace{\tilde{\mathbf{F}}}^{\text{Permuted DFT}} \tilde{\mathbf{r}}_x$

FC Goal: Estimate the real vector $\tilde{\mathbf{r}}_x$ from $\{b_m\}_{m=1}^M$

Nonparametric Estimation (Passive FC)

Autocorrelation Reconstruction

- ***Constraints:***
  1. The bounds $0 \leq r_x(0) \leq P_{\max}$, $|r_x(\ell)| \leq r_x(0)$, $\ell = 1, \ldots, K-1$ define a bounded $\mathcal{P}_0$ as the initial feasible region for $\tilde{\mathbf{r}}_x$
  2. Receiving $\{b_m\}_{m=1}^{M} \Rightarrow b_m(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m) \geq 0$, $m = 1, \ldots, M$ *(ignoring α estimation errors)*
  3. $\mathbf{R}_x = \text{Toeplitz}(\mathbf{r}_x) \succeq 0$ and $\hat{\mathbf{s}}_x = \tilde{\mathbf{F}} \hat{\mathbf{r}}_x \geq 0$

     Proposition: $\boxed{\tilde{\mathbf{F}} \hat{\mathbf{r}}_x \geq 0 \Rightarrow \mathbf{R}_x \succeq 0}$

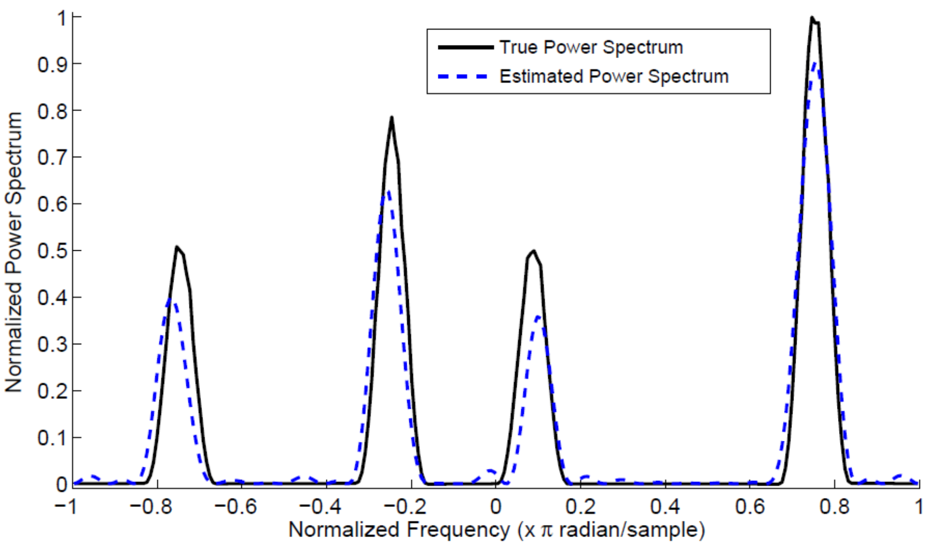
- ***Cost Function:***

  Minimize total signal power $E[|x(n)|^2] = r_x(0) = \frac{1}{N_F} \sum_{f=0}^{N_F-1} \hat{s}_x(f) = \frac{1}{N_F} ||\hat{\mathbf{s}}_x||_1$

- ***Linear Program:***

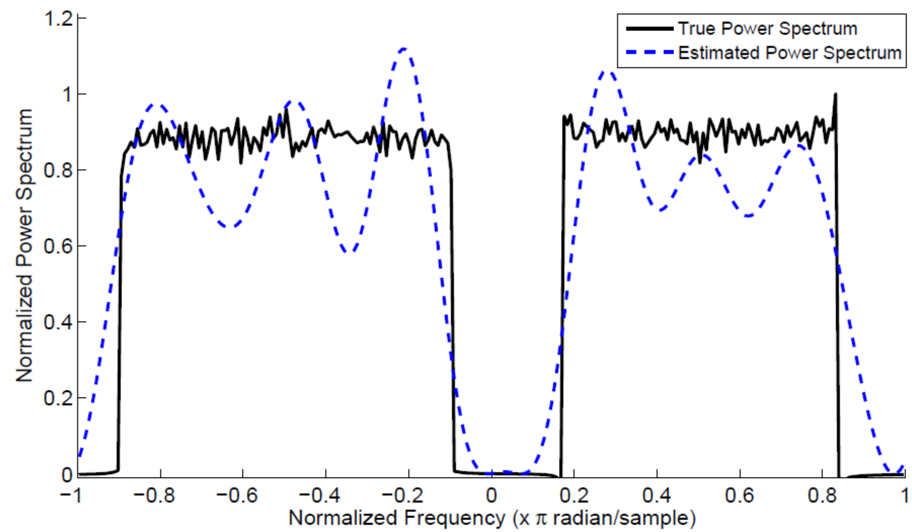
$$\boxed{\begin{aligned} \min_{\tilde{\mathbf{r}}_x \in \mathcal{P}_0} \quad & r_x(0) \\ \text{s.t.} : \quad & b_m(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m) \geq 0, \\ & \tilde{\mathbf{F}} \tilde{\mathbf{r}}_x \geq 0, \quad m = 1, \ldots, M. \end{aligned}}$$

Simulations



(a) Frugal sensing with sparse spectrum.

M=100, K=24, $t_m=t$*, 30 sensors send* $b_m=1$

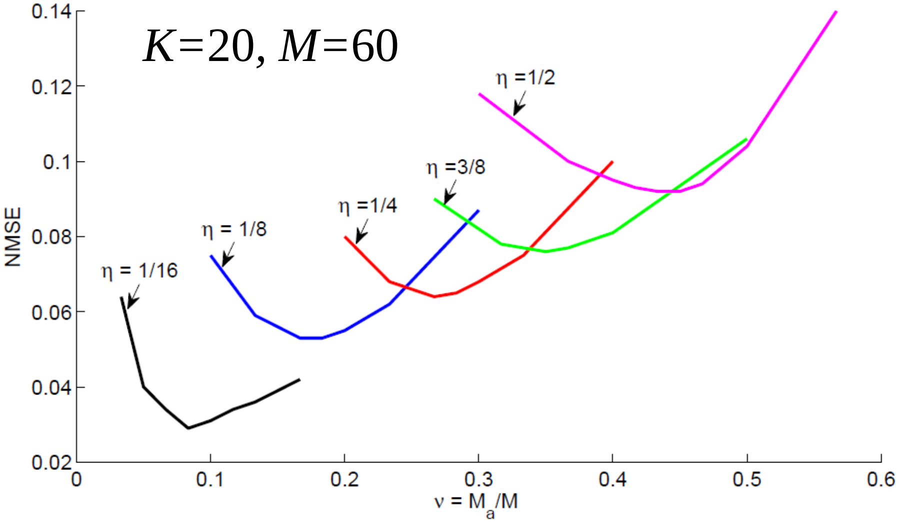


(b) Frugal sensing with dense spectrum.

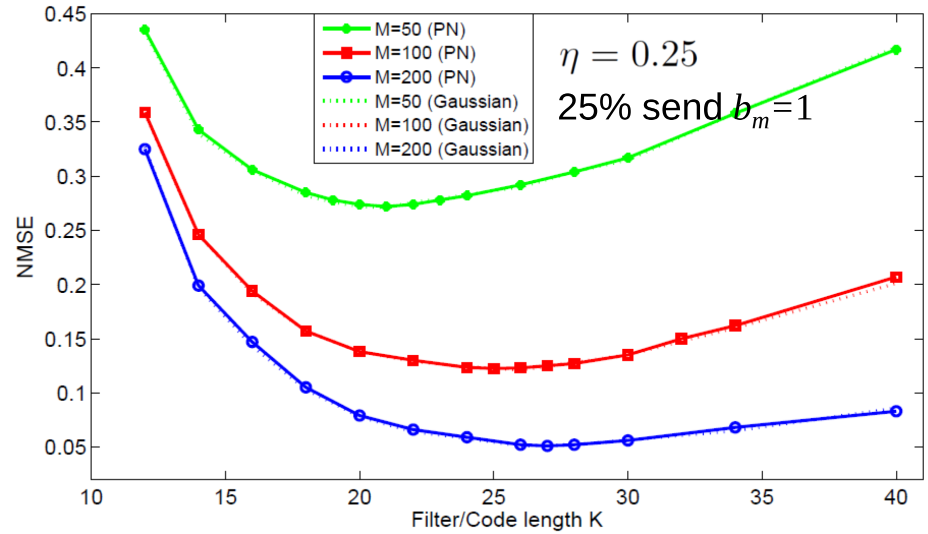
M=100, K=10, $t_m=t$, 50 *sensors send* $b_m=1$

100 bits equivalent to 3 single precision IEEE floats (*r(0)* and *r(1)*)

Threshold Selection & Filter Length



(a) $\mathrm{NMSE_s}$ vs. ν for different η Sparsity ratio



(b) $\mathrm{NMSE_s}$ vs. K for different M.

- Threshold should be tuned such that number of sensors reporting b_m=1 (above threshold) decreases as the power spectrum becomes more sparse

- Small K → smeared power spectrum estimate
- Large K → more unknowns vs. inequality constraints (more under-determined) → high uncertainty
- More M → optimal K^* increases
- Binary PN simpler than Gaussian

Gaussian Errors – ML

$$\boxed{\hat{\alpha}_m = \mathbf{q}_m^T \tilde{\mathbf{r}}_x + e_m}$$

Frequency-selective fading
+ insufficient sample averaging

$$\mathcal{M}_+ := \{m | b_m = 1\}$$
$$\mathcal{M}_- := \{m | b_m = -1\}$$

$$\begin{aligned} f(b_1, \ldots, b_M | \tilde{\mathbf{r}}_x) &= \prod_{m \in \mathcal{M}_+} \Pr(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m + e_m \geq 0) \prod_{m \in \mathcal{M}_-} \Pr(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m + e_m \leq 0) \\ &= \prod_{m \in \mathcal{M}_+} \Phi\left(\frac{\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m}{\sigma_m}\right) \prod_{m \in \mathcal{M}_-} \Phi\left(-\frac{\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m}{\sigma_m}\right) \end{aligned}$$

Gaussian CDF

- Log-likelihood function:

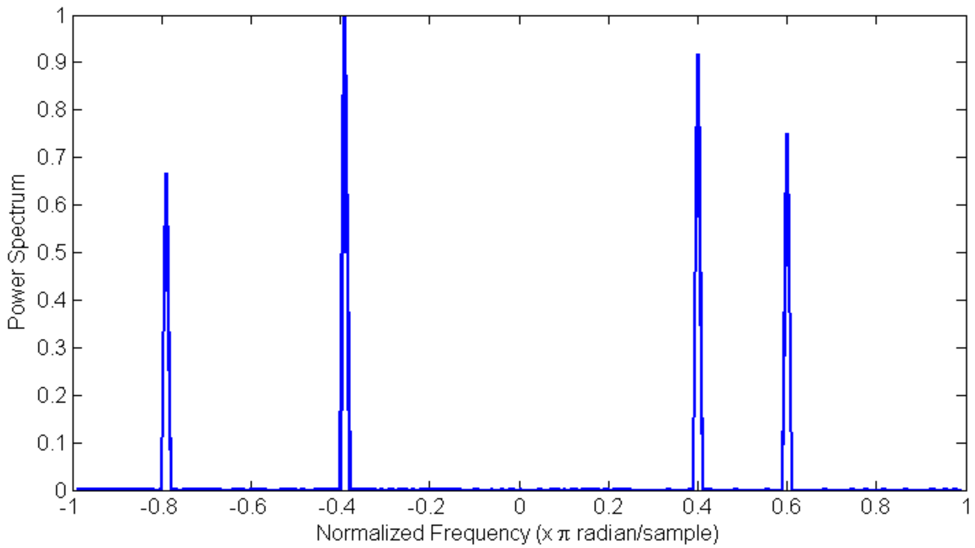
$$L(\tilde{\mathbf{r}}_x) := \log f(b_1, \ldots, b_M | \tilde{\mathbf{r}}_x) = \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m)}{\sigma_m}\right)$$

- Constrained ML - Convex optimization problem:

$$\boxed{\max_{\tilde{\mathbf{r}}_x \in \mathcal{P}_0} L(\tilde{\mathbf{r}}_x) \quad \text{s.t. :} \quad \tilde{\mathbf{F}} \tilde{\mathbf{r}}_x \geq \mathbf{0}}$$

Parametric Estimation (Line Spectra)

Line spectrum



- L tones (spectral components) $$r(k) = \sum_{\ell=1}^{L} p_\ell e^{j\omega_\ell k}$$
- Estimate frequencies and powers $\{\omega_\ell, p_\ell\}_{\ell=1}^{L}$

Line spectrum estimation from few bits

→ Estimation from inequalities (instead of equalities)

1) Nonparametric LP + MUSIC

1. Nonparametric estimation of autocorrelation

$$\boxed{\begin{aligned} \min_{\tilde{\mathbf{r}}_x \in \mathcal{P}_0} \quad & r_x(0) \\ \text{s.t.}: \quad & b_m(\mathbf{q}_m^T \tilde{\mathbf{r}}_x - t_m) \geq 0, \\ & \tilde{\mathbf{F}} \tilde{\mathbf{r}}_x \geq 0, \quad m = 1, \ldots, M. \end{aligned}}$$

2. Parametric estimation of frequencies using MUSIC (MUltiple SIgnal Classification)

   L strongest peaks of: $\hat{S}_{\mathrm{MUSIC}}(\omega) = \dfrac{1}{\sum_{i=L+1}^{K} |\mathbf{e}(\omega)^H \mathbf{u}_i|^2}$

   $\mathbf{u}_i$ eigenvector corresponding to i^th strongest eigenvalue of autocorrelation matrix and $\mathbf{e}(\omega) := [1, e^{j\omega}, \ldots, e^{j(K-1)\omega}]^T$

3. LS for powers: $\{\hat{p}_\ell\}_{\ell=1}^{L} = \arg \min_{\{p_\ell\}_{\ell=1}^{L}} \sum_{k=0}^{K-1} \left| \hat{r}(k) - \sum_{\ell=1}^{L} p_\ell e^{jk\hat{\omega}_\ell} \right|^2$

2) Nonparametric ML + MUSIC

1. Exploit Gaussian distribution of errors $\{e_m\}_{m=1}^{M}$

$$\boxed{\max_{\tilde{\mathbf{r}}\in\mathcal{P}_0} \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\mathbf{q}_m^T\tilde{\mathbf{r}} - t_m)}{\sigma_m}\right) \qquad \text{s.t. :} \quad \tilde{\mathbf{F}}\tilde{\mathbf{r}} \geq \mathbf{0}}$$

2. MUSIC for $\{\omega_\ell\}_{\ell=1}^{L}$ then LS for $\{p_\ell\}_{\ell=1}^{L}$

3) Parametric ML

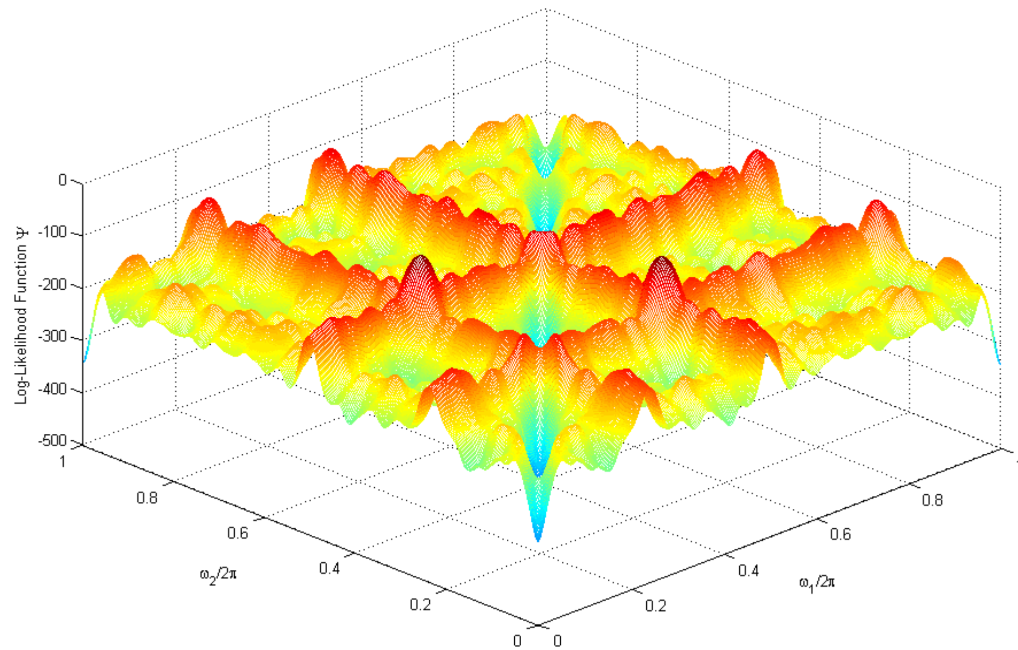
$$\alpha_m = \mathbf{q}_m^T \mathbf{r} = \sum_{\ell=1}^{L} \sum_{k=1-K}^{K-1} p_\ell e^{jk\omega_\ell} q_m(k)$$

- Estimate $\{\omega_\ell, p_\ell\}_{\ell=1}^{L}$ directly by maximizing the log-likelihood:

$$\sum_{m=1}^{M} \log \Phi \left(b_m \left(\sum_{\ell=1}^{L} \sum_{k=1-K}^{K-1} p_\ell e^{jk\omega_\ell} q_m(k) - t_m \right) \Big/ \sigma_m \right)$$

- Nonconvex in $\{\omega_\ell\}_{\ell=1}^{L}$

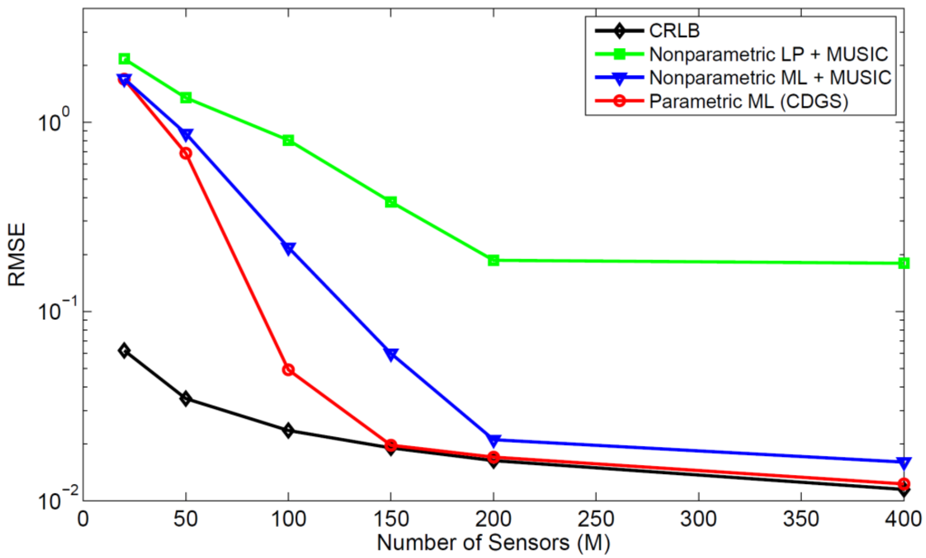
  Solve with Coordinate Descent Grid Search (CDGS)



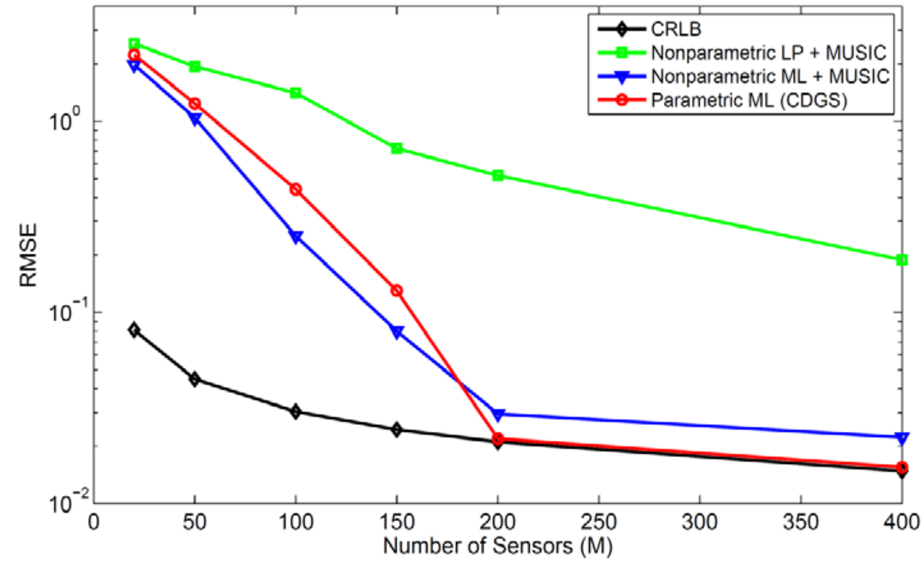
Numerical Results

$K=25$, $t_m=t$, □$^2=1$

2 far-apart tones



2 close tones $(|\omega_1 - \omega_2| = 0.06\pi < 0.08\pi = 2\pi/K)$



- Parametric ML (solved with CDGS) outperforms other techniques and meets the CRLB for large M
- Nonparametric ML + MUSIC can do better for small M when tones are very close

Adaptive Thresholding (Active FC)

Adaptive Thresholding Algorithm

- The volume of $\mathcal{P}_M = \mathcal{P}_0 \cap \{\mathbf{y} \mid b_m(\mathbf{q}_m^T \mathbf{y} - t_m) \geq 0,\ m = 1, \ldots, M\}$ gives a measure of ignorance / uncertainty about $\tilde{\mathbf{r}}_x$

⇨ *adaptively select* $\{t_m\}_{m=1}^M$ *to ensure* $\mathcal{P}_M$ *is as small as possible*

CCAT Algorithm:

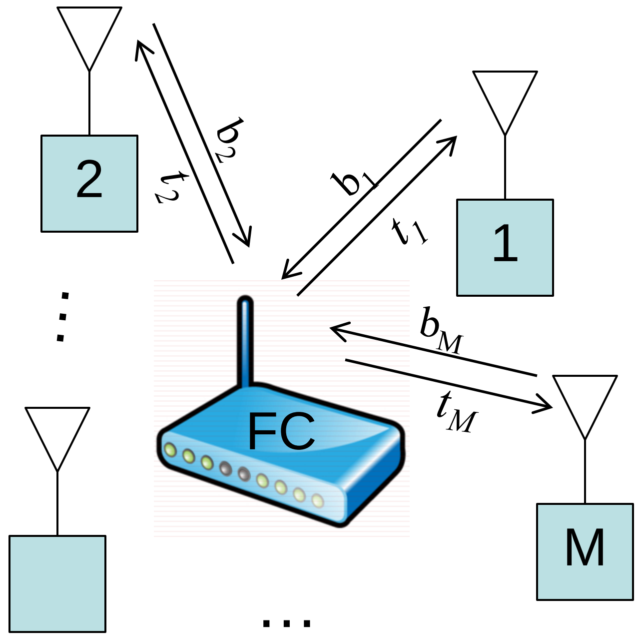
Given $\mathcal{P}_0$*, its Chebyshev center (CC),* $\mathbf{y}_{cc}^{(0)}$*, and* $\{\mathbf{q}_m\}_{m=1}^M$

For $m=1,\ldots M$*, do*

1. Set $t_m = \mathbf{q}_m^T \mathbf{y}_{cc}^{(m-1)}$, send it to senor m
2. Upon receiving b_m update:

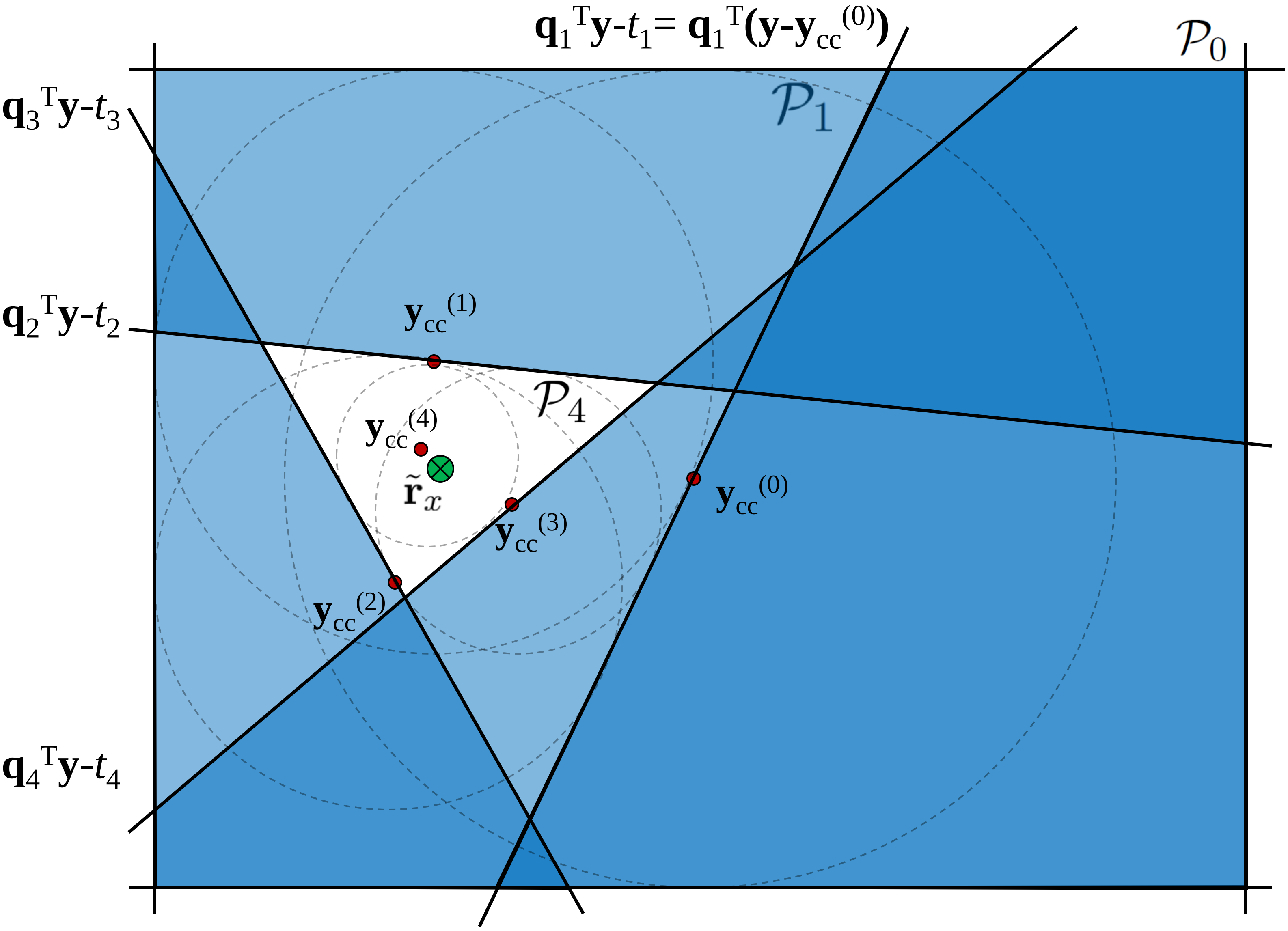
$$\mathcal{P}_m := \begin{cases} \mathcal{P}_{m-1} \cap \{\mathbf{y} \mid \mathbf{q}_m^T \mathbf{y} \geq t_m\} & \text{if } b_m = 1 \\ \mathcal{P}_{m-1} \cap \{\mathbf{y} \mid \mathbf{q}_m^T \mathbf{y} < t_m\} & \text{if } b_m = -1 \end{cases}$$

3. Compute the CC, $\mathbf{y}_{cc}^{(m)}$, of $\mathcal{P}_m$



$$\hat{\mathbf{r}}_x = \mathbf{y}_{cc}^{(M)}$$

2-D Example



Significant portion of the feasible region is cut-off after each iteration

CC Computation and Convergence

- For $\mathcal{P} := \{\mathbf{y} \mid \mathbf{a}_i^T \mathbf{y} \leq d_i,\ \ i = 1, \ldots, L\}$ the CC is computed by solving the LP:

$$\boxed{\begin{array}{ll} \displaystyle\max_{R \geq 0, \mathbf{y}} & R \\ \text{s.t.} : & \mathbf{a}_i^T \mathbf{y} + R||\mathbf{a}_i||_2 \leq d_i,\ \ i = 1, \ldots, L \end{array}}$$

- *Convergence:* $\mathbf{y}_{cc}^{(M)} \rightarrow \tilde{\mathbf{r}}_x$ as $M \rightarrow \infty$
  – Radius of largest inscribed ball at each iteration goes to zero
  – Convergence with independence conditions across $\{\mathbf{q}_m\}_{m=1}^{M}$

- *Dropping Constraints*
  – Linear inequalities increase with each iteration → complexity increases
  – Drop redundant constraints, or keep fixed number of most relevant ones
  – Sensor 1 is redundant in example

Positivity Constraints

- Spectrum positivity constraints
  - For truncated K-lag autocorrelation $\hat{\mathbf{s}}_x = \tilde{\mathbf{F}}\tilde{\mathbf{r}}_x \ngeq 0$
  - Can prevent convergence to true autocorrelation vector
  - Beneficial with small M

- Relaxed positivity constraints

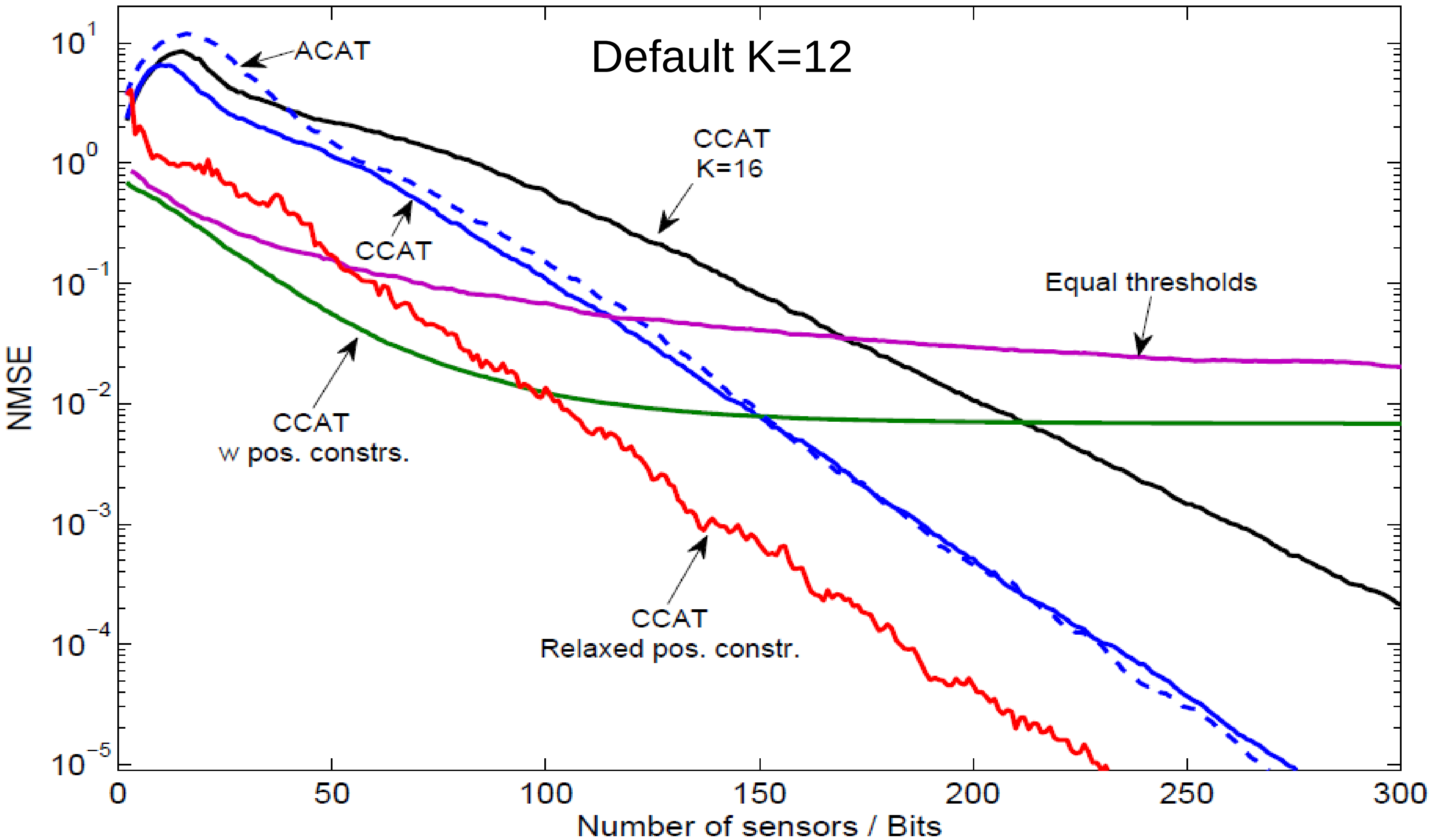
  Define $\mathbf{v}(\omega) := [1, e^{j\omega}, \ldots, e^{j\omega(K-1)}]^T$

$$\mathbf{R}_x \succeq 0 \Rightarrow \mathbf{v}(\omega)^H \mathbf{R}_x \mathbf{v}(\omega) \geq 0,\ \forall \omega \in [0, 2\pi]$$

$$\Rightarrow \sum_{\ell=-K+1}^{K-1} (K - |\ell|) r(\ell) e^{-j\omega\ell} \geq 0,\ \forall \omega \in [0, 2\pi]$$

$$\Rightarrow \tilde{\mathbf{F}}\mathbf{D}\tilde{\mathbf{r}}_x \geq 0$$

Numerical Results



Sensor Polling Algorithm

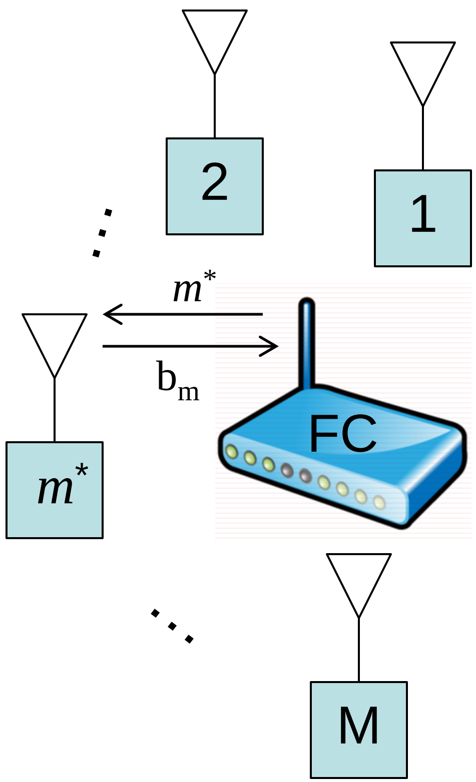
- Avoid downlink threshold communication overhead
- Each sensor pseudo-randomly chooses its threshold

CCSP Algorithm:

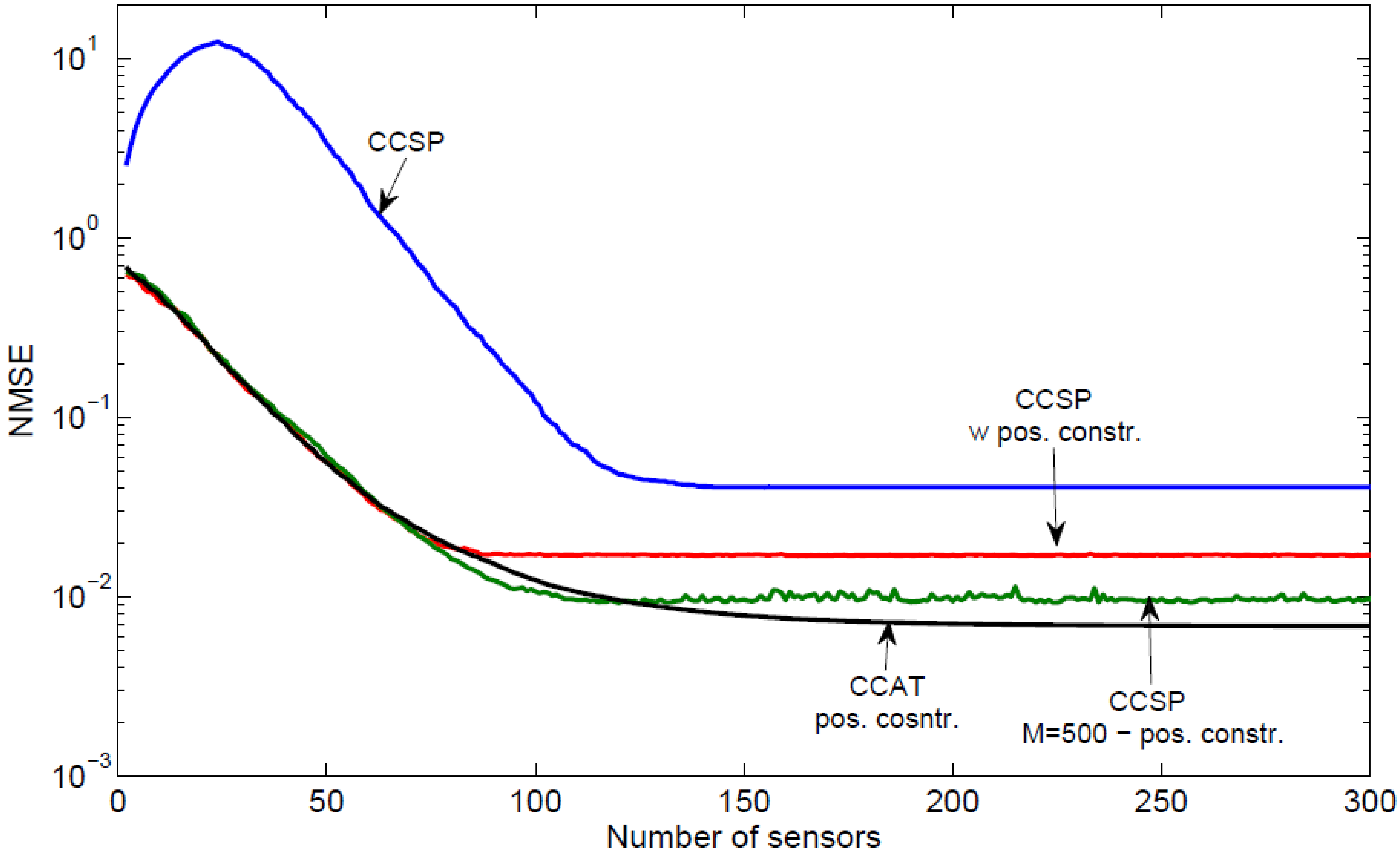
Given $\mathcal{P}_0$, $\mathbf{y}_{cc}^{(0)}$, $\{\mathbf{q}_m\}_{m=1}^{M}$, $\{t_m\}_{m=1}^{M}$, $k=1$

While $k \leq M$, *do*

1. For each $m \in \bar{\mathcal{J}}$, find the distance between $\{\mathbf{y} \mid \mathbf{q}_m^T \mathbf{y} - t_m = 0\}$ and $\mathbf{y}_{cc}^{(k-1)}$: $d_m = \dfrac{|\mathbf{q}_m^T \mathbf{y}_{cc}^{(k-1)} - t_m|}{||\mathbf{q}_m||}$
2. Poll sensor $m^* = \arg\min d_m$
3. Upon receiving b_m, delete m^* from $\bar{\mathcal{J}}$, and update
$$\mathcal{P}_k := \begin{cases} \mathcal{P}_{k-1} \cap \{\mathbf{y} \mid \mathbf{q}_{m^*}^T \mathbf{y} \geq t_{m^*}\} & \text{if } b_{m^*} = 1 \\ \mathcal{P}_{k-1} \cap \{\mathbf{y} \mid \mathbf{q}_{m^*}^T \mathbf{y} < t_{m^*}\} & \text{if } b_{m^*} = 0 \end{cases}$$
4. Compute the CC, $\mathbf{y}_{cc}^{(k)}$
5. Increment k and repeat, or terminate



Numerical Results



Summary

Adequate power spectrum sensing is possible from few bits

- Nonparametric estimation

  K-lag autocorrelation reconstruction
  - LP formulation with perfect sensor power measurement s
  - Constrained ML formulation exploiting Gaussian errors

- Parametric line spectrum estimation
  - Parametric ML solved with CDGS meets the CRLB for large M

- Adaptive thresholding (active FC)
  - FC adaptively picks the threshold so as to cut off a half-space from the feasible region along its Chebyshev center
  - FC judiciously polls sensors with pseudo-random thresholds

Thank You !

Proposition $\mathbf{F}\mathbf{r}_x \geq 0 \Rightarrow \mathbf{R}_x \succeq 0$

$$\bar{\mathbf{r}}_x := [\,\underbrace{0\ldots 0}_{\frac{N_F-2K+1}{2}}\quad \mathbf{r}_x^T \quad \underbrace{0\ldots 0}_{\frac{N_F-2K+1}{2}}\,]^T \Rightarrow \bar{\mathbf{F}}\bar{\mathbf{r}}_x = \mathbf{F}\mathbf{r}_x = \mathbf{s}_x$$

Square DFT matrix

$$\mathbf{R}_c = \begin{bmatrix} r(0) & r(-1) & 0 & 0 & r(1) \\ r(1) & r(0) & r(-1) & 0 & 0 \\ 0 & r(1) & r(0) & r(-1) & 0 \\ 0 & 0 & r(1) & r(0) & r(-1) \\ r(-1) & 0 & 0 & r(1) & r(0) \end{bmatrix}$$

$K = 2$ and $N_F = 5$

Circulant matrices are diagonalized by a DFT: $\mathbf{R}_c = \frac{1}{N_F}\mathbf{W}^H\mathbf{\Lambda}\mathbf{W}$

$\bar{\mathbf{F}}\bar{\mathbf{r}}_x = \mathbf{s}_x \geq 0 \Rightarrow$ $\mathbf{R}_c$ is positive semidefinite

$\mathbf{R}_x$ is the K-th order leading principal submatrix of $\mathbf{R}_c \Rightarrow \mathbf{R}_x \succeq 0$