PARAMETRIC FRUGAL SENSING OF AUTOREGRESSIVE POWER SPECTRA

Aritra Konar and Nicholas D. Sidiropoulos

Dept. of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN

ABSTRACT

Estimating the power spectrum of a wide-sense stationary stochastic process is a core component of several signal processing tasks. Distributed spectrum sensing problems naturally emerge in cases where measurements of different realizations of a stochastic process are collected at multiple spatial locations. This paper describes a distributed power spectrum sensing scheme for stochastic processes which are well represented by an autoregressive (AR) process. The sensing model comprises a network of scattered low-end sensors which transmit randomly filtered, one bit quantized power measurements to a fusion center. The problem of AR power spectrum estimation from such binary power measurements is cast as a non-convex optimization problem, and an alternating minimization algorithm is proposed to obtain a stationary point. Simulations showcase the effectiveness of this scheme when the AR parametrization is valid.

1. INTRODUCTION

In several modern signal processing applications (e.g., cognitive radio sensing, radio astronomy), it is desirable to perform power spectrum (PS) estimation from compressed measurements drawn from the underlying wide-sense stationary (WSS) stochastic process. Non-parametric methods for PS estimation from compressed analog measurements were developed in [1, 2]. These results were extended to the distributed sensing setting in [3] where a non-parametric approach was adopted to reconstruct power spectra from one-bit compressed measurements. These methods are well suited for cases when there is little prior knowledge about the structure of the underlying WSS process. However, if it is known *apriori* that the process admits a parametric representation, then this information can be exploited for developing parametric PS estimation methods with improved estimation performance - see [4, 5] for a *moving average* (MA) parametrization in the context of [3]. Here we consider an autoregressive (AR) parametrization instead.

AR power spectrum (PS) estimation from analog measurements is a classic signal processing problem with many applications in geophysics, radar, sonar, radio astronomy, oceanography and speech processing, see [6] and references therein. Traditional AR PS estimation is a two-step process, where a non-parametric estimate of the autocorrelation is used to construct a system of linear equations that determine the AR model parameters. In contrast, we consider the problem of AR PS estimation in a distributed sensing scenario, where we employ a network of low-end sensors, each of which draws samples from the underlying WSS process, filters them using a random broadband filter, averages the power at the filter's output and then compresses the result to one bit. The single bit power measurements are then transmitted to a fusion center (FC), which aims to reconstruct the ambient PS. To the best of our knowledge, this is the first time that the problem of AR PS estimation from a small number of one-bit power measurements has been considered. Exploiting the underlying parametric structure and other pertinent properties of autocorrelation sequences, it is shown that problem of AR PS estimation can be formulated as a non-convex optimization problem, which appears hard to solve to global optimality in polynomial-time. Instead, the conditionally decomposable structure of the problem formulation is utilized to develop a simple alternating minimization algorithm for obtaining high quality sub-optimal solutions, and convergence to a stationary point of the original problem is established. Simulations indicate the effectiveness of the proposed scheme.

Contact: (konar006,nikos)@umn.edu. Supported in part by NSF AST-1247885, ECCS 1231504.

2. SYSTEM MODEL

We consider a network sensing scenario (first described in [3]), where M distributed, low-end sensors transmit randomly filtered, single bit quantized power measurements to a FC. Every sensor m acquires samples of $x(n)$ in the form of the data sample vector $\mathbf{x}_m^{(i)} = [x_m(i), x_m(i-1), \cdots, x_m(i-K+1)]^T \in \mathbb{C}^K$, (where i denotes the sampling instance). Furthermore, each sensor is equipped with an FIR filter with a broadband impulse response $\mathbf{g}_m \in \mathbb{C}^K$ which is generated independently and pseudo-randomly, with each element being drawn from the following uniform distribution defined on a set of 4 complex symbols

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

where $\mathcal{U}(S)$ denotes the uniform probability mass function defined over the finite set S. Using $\mathbf{g}_m$, sensor m obtains random linear projections of $\mathbf{x}_m^{(i)}$ of the form $z_m^{(i)} = \mathbf{g}_m^H \mathbf{x}_m^{(i)}$. The average power of the random linear projections at each sensor is defined as,

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

where $\mathbf{R}_x = \mathbb{E}[\mathbf{x}_m^{(i)} \mathbf{x}_m^{(i)H}] \in \mathbb{C}^{K \times K}$ is the Toeplitz-Hermitian autocorrelation matrix of $\mathbf{x}_m^{(i)}$ and is given by

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(K-1) \\ r_x(-1) & r_x(0) & \cdots & r_x(K-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(-K+1) & r_x(-K+2) & \cdots & r_x(0) \end{bmatrix} \tag{3}$$

Each sensor obtains soft estimates of ρ_m by a simple sample averaging operation defined as

$$\rho_m^{(N)} := \frac{1}{N} \sum_{n=1}^{N} |z_m(n)|^2 \tag{4}$$

Finally, each power estimate $\rho_m^{(N)}$ is compared to a single, predetermined threshold t. If $\rho_m^{(N)} \geq t$, the sensor transmits a bit $b_m = 1$ to the FC, otherwise, it sends $b_m = -1$. Thus, each transmitted power measurement bit can be expressed as

$$b_m = \text{sign}(\rho_m^{(N)} - t) \tag{5}$$

where $\text{sign}(u) = 1$ if $u \geq 0$ and -1 otherwise $\forall u \in \mathbb{R}$. Assuming certain ergodic mixing conditions hold [7, p. 171], we have that $\lim_{N\to\infty} \rho_m^{(N)} = \rho_m$, i.e., sample averages converge to ensemble averages, which in turn corresponds to

$$b_m = \text{sign}(\mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m - t) \tag{6}$$

Thus, on receipt of a bit $b_m = 1$ (or $b_m = -1$) from sensor m, the FC infers that the inequality $\mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m \geq t$ (or $\mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m < t$) is satisfied. The task of the FC is to reconstruct the ambient PS from the bit measurements $\{b_m\}_{m=1}^M$.

3. PROBLEM FORMULATION

We seek to formulate the problem of PS reconstruction from binary power measurements as an optimization problem which takes into account all pertinent prior available information so as to reduce the under-determinacy of our estimation setup (cf. the *binary* measurements).

First, note that since $\mathbf{R}_x$ is Toeplitz-Hermitian, it can be expressed as $\mathbf{R}_x = \sum_{k=-(K-1)}^{K-1} r_x(k)\mathbf{\Theta}_k^K$, where $\mathbf{\Theta}_k^K \in \mathbb{R}^{K\times K}$ is a elementary Toeplitz matrix with ones on the k^{th} diagonal and zeros elsewhere (by our notation, $k = 0$ corresponds to the main diagonal, $k > 0$ correspond to the super-diagonals and $k < 0$ are the sub-diagonals). Exploiting this structure, the average power $\rho_m = \mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m$ can be expressed as

$$\mathbf{g}_m^H \mathbf{R}_x \mathbf{g}_m = \mathbf{g}_m^H \left(\sum_{k=-(K-1)}^{K-1} r_x(k)\mathbf{\Theta}_k^K \right) \mathbf{g}_m \tag{7a}$$

$$= \sum_{k=-(K-1)}^{K-1} \underbrace{\mathbf{g}_m^H \mathbf{\Theta}_k^K \mathbf{g}_m}_{c_m(k)} r_x(k) \tag{7b}$$

$$= \mathbf{q}_m^T \mathbf{r}_x \tag{7c}$$

where $c_m(k)$ represents the k^{th} lag of the deterministic autocorrelation sequence of the m^{th} broadband filter with impulse response $\mathbf{g}_m$, and in the last step we have defined the vectors

$$\begin{aligned} \mathbf{q}_m :=& [c_m(0), 2\,\text{Re}\{c_m(1)\}, \cdots, 2\,\text{Re}\{c_m(K-1)\}, \\ & 2\,\text{Im}\{c_m(1)\}, \cdots, 2\,\text{Im}\{c_m(K-1)\}]^T \in \mathbb{R}^{2K-1} \\ \mathbf{r}_x :=& [r_x(0), 2\,\text{Re}\{r_x(1)\}, \cdots, 2\,\text{Re}\{r_x(K-1)\}, \\ & 2\,\text{Im}\{r_x(1)\}, \cdots, 2\,\text{Im}\{r_x(K-1)\}]^T \in \mathbb{R}^{2K-1} \end{aligned} \tag{8}$$

From (6) and (7c), we have that each measurement bit b_m corresponds to the following linear inequality in $\mathbf{r}_x$.

$$b_m(\mathbf{q}_m^T \mathbf{r}_x - t) \geq 0, \forall\, m \in \mathcal{M} \tag{9}$$

Additionally, the fact that the autocorrelation matrix $\mathbf{R}_x$ associated with any autocorrelation vector $\mathbf{r}_x$ of any order must be positive semi-definite, can be exploited. This property ensures the non-negativity of the PS $\forall\, \omega \in [0, 2\pi]$. However, since we employ a finite parametrization of the autocorrelation sequence, the windowed PS estimate obtained from the discrete-time Fourier Transform (DTFT) of $\mathbf{r}_x$ is not necessarily non-negative at all frequencies. In spite of this, it was demonstrated in [3] that incorporating the non-negativity of the windowed PS estimate as a constraint improves estimation quality by reducing the under-determinacy of the problem setup. This constraint can be represented as $\mathbf{F}\mathbf{r}_x \geq \mathbf{0}$, where $\mathbf{F}\mathbf{r}_x$ is the discrete N_F-point PS estimate, $\mathbf{F} := \tilde{\mathbf{F}}\mathbf{W}$, $\tilde{\mathbf{F}} \in \mathbb{C}^{N_F\times(2K-1)}$ is the phase shifted discrete Fourier transform (DFT) matrix, and

$$\mathbf{W} := \begin{bmatrix} \mathbf{0}_{K-1} & \mathbf{J}_{K-1} & -j\mathbf{J}_{K-1} \\ 1 & \mathbf{0}_{K-1}^T & \mathbf{0}_{K-1}^T \\ \mathbf{0}_{K-1} & \mathbf{I}_{K-1} & j\mathbf{I}_{K-1} \end{bmatrix} \tag{10}$$

where $\mathbf{0}_{K-1}$ is a vector of $K-1$ zeros, $\mathbf{I}_{K-1}$ is the $K-1$ identity matrix, and $\mathbf{J}_{K-1}$ is the $K-1$ anti-identity matrix. It has already been established [3, Appendix C] that $\mathbf{F}\mathbf{r}_x \geq \mathbf{0} \implies \mathbf{R}_x \succeq \mathbf{0}$, thus the latter constraint is redundant.

Finally, we exploit the fact that $x(n)$ admits an AR representation. The second-order statistics of an AR process of order p are parametrized by the Yule-Walker equations (YWEs) [8], which are given by

$$r_x(l) + \sum_{k=1}^{p} \alpha(k) r_x(l-k) = \delta(l),\ \forall\, l \in \mathbb{Z}^+ \tag{11}$$

where $\boldsymbol{\alpha} = [\alpha(1), \cdots, \alpha(p)]^T \in \mathbb{C}^p$ are the AR parameters and $\mathbb{Z}^+$ is the set of all non-negative integers. The AR parameters correspond to the coefficients of the minimum-phase polynomial $A_p(z) = 1 + \sum_{k=1}^{p} \alpha(k) z^{-k}$ and the parametric form of the PS associated with the AR process is given by

$$S_x(e^{j\omega}) = \frac{1}{|A_p(e^{j\omega})|^2} = \frac{1}{|1 + \sum_{k=1}^{p} \alpha(k) e^{-j\omega k}|^2} \tag{12}$$

If the true autocorrelation of the AR process were known, then one could have formed a square system of $p+1$ linear equations in $\boldsymbol{\alpha}$ obtained from (11) by taking $l = 0 : p$. Since the autocorrelation matrix is positive definite if and only if the AR parameters are minimum-phase [9, p. 228], it follows that this system of linear equations can be uniquely solved for $\boldsymbol{\alpha}$, which also corresponds to the true minimum-phase solution. In practice when the true autocorrelations are unknown, the traditional two-step approach uses sample autocorrelation estimates obtained from the WSS process to solve the YWEs for $\boldsymbol{\alpha}$. The sample autocorrelation matrix in that case can also be shown to be positive definite under mild conditions [8, p. 93], and hence the YWEs again admit a unique solution. Using sample autocorrelation estimates to solve the YWEs can also be interpreted as solving an approximate Maximum-Likelihood estimation problem for $\boldsymbol{\alpha}$ [10, p.196], which yields the true AR parameters when the sample size is large. Hence, the sample autocorrelations lags in the range $[-p, p]$ constitute a *sufficient statistic* for estimating $\boldsymbol{\alpha}$.

Note that sample autocorrelation estimates of $\mathbf{r}_x$ are not available in our setup since our problem involves estimation from a finite number of bits, and not samples of the WSS process. An estimate of the window of $2p-1$ autocorrelation lags obtained from a few bits is not guaranteed to be a sufficient statistic for estimating $\boldsymbol{\alpha}$. Thus, we propose to estimate $\mathbf{r}_x$ and $\boldsymbol{\alpha}$ jointly. In order to make our problem less under-determined, we use the information contained in the higher autocorrelation lags. Assuming $K \geq p$, we form the following overdetermined system of K bilinear equations in $\mathbf{r}_x$ and $\boldsymbol{\alpha}$.

$$\begin{bmatrix} r_x(0) & r_x(-1) & \cdots & r_x(-p) \\ r_x(1) & r_x(0) & \cdots & r_x(-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(p) & r_x(p-1) & \cdots & r_x(0) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(K-1) & r_x(K-2) & \cdots & r_x(K-p-1) \end{bmatrix} \begin{bmatrix} 1 \\ \alpha(1) \\ \vdots \\ \alpha(p) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{13}$$

Denoting $\tilde{\mathbf{R}}_x \in \mathbb{C}^{K\times(p+1)}$ as the matrix defined in (13), $\tilde{\boldsymbol{\alpha}} := [1; \boldsymbol{\alpha}]^T \in \mathbb{C}^{p+1}$ and $\mathbf{e}_1^K$ as the first canonical basis vector in $\mathbb{R}^K$, we can compactly express (13) as

$$\tilde{\mathbf{R}}_x \tilde{\boldsymbol{\alpha}} = \mathbf{e}_1^K \tag{14}$$

which we refer to as the *extended Yule Walker equations* for an AR process. In order to impose AR structure on $\mathbf{r}_x$, we propose to use the following formulation

$$\min_{\mathbf{r}_x,\boldsymbol{\alpha}} \quad \|\tilde{\mathbf{R}}_x\tilde{\boldsymbol{\alpha}} - \mathbf{e}_1^K\|_2^2 \tag{15a}$$

$$\text{s.t.} \quad b_m(\mathbf{q}_m^T\mathbf{r}_x - t) \geq 0,\ \forall\, m \in \mathcal{M} \tag{15b}$$

$$\mathbf{F}\mathbf{r}_x \geq \mathbf{0} \tag{15c}$$

Note that the problem is non-convex since the cost function is the composition of a convex function with a bilinear function in $\mathbf{r}_x$ and $\boldsymbol{\alpha}$. Hence, it is *unlikely* that all instances of (15) can be solved to global optimality in polynomial time. In the next section, we describe an alternating minimization (AM) algorithm for obtaining approximate solutions of (15).

4. ALTERNATING MINIMIZATION

Although problem (15) is non-convex, if either of the variables $\mathbf{r}_x$ or $\boldsymbol{\alpha}$ is fixed, the resulting subproblem is convex in the other variable. This suggests that $\mathbf{r}_x$ and $\boldsymbol{\alpha}$ can be updated in an alternating fashion, which results in a simple AM algorithm consisting of the following steps.

1. *Initialization:* First, we compute a solution of the problem

$$\text{find} \quad \mathbf{r}_x \tag{16a}$$

$$\text{s.t.} \quad b_m(\mathbf{q}_m^T\mathbf{r}_x - t) \geq 0,\ \forall\, m \in \mathcal{M} \tag{16b}$$

$$\mathbf{F}\mathbf{r}_x \geq \mathbf{0} \tag{16c}$$

which is a Linear Programming (LP) feasibility problem. The solution of (16) corresponds to a truncated K-lag autocorrelation sequence that is consistent with the inequalities corresponding to the bit measurements $\{b_m\}_{m=1}^M$, which is then used to initialize the algorithm.

2. $\boldsymbol{\alpha}$ *Update:* When $\mathbf{r}_x$ is fixed, the update for $\boldsymbol{\alpha}$ is a Least Squares problem which can be represented as

$$\|\tilde{\mathbf{R}}_x\tilde{\boldsymbol{\alpha}} - \mathbf{e}_1^K\|_2^2 = \|\bar{\mathbf{R}}_x\boldsymbol{\alpha} + \boldsymbol{\rho} - \mathbf{e}_1^K\|_2^2 = \|\bar{\mathbf{R}}_x\boldsymbol{\alpha} + \tilde{\boldsymbol{\rho}}\|_2^2 \tag{17}$$

where $\boldsymbol{\rho} \in \mathbb{C}^K$ is the first column of $\tilde{\mathbf{R}}_x$, $\bar{\mathbf{R}}_x \in \mathbb{C}^{K\times p}$ is obtained by deleting $\boldsymbol{\rho}$ from $\tilde{\mathbf{R}}_x$ and $\tilde{\boldsymbol{\rho}} = \boldsymbol{\rho} - \mathbf{e}_1^K$. The closed form solution of (17) is given by

$$\boldsymbol{\alpha} = -(\bar{\mathbf{R}}_x^H\bar{\mathbf{R}}_x + \epsilon\mathbf{I}_K)^{-1}(\bar{\mathbf{R}}_x^H\tilde{\boldsymbol{\rho}}) \tag{18}$$

where $\epsilon > 0$ is an extra regularization parameter which guards against ill-conditioning of $\bar{\mathbf{R}}_x$.

3. $\mathbf{r}_x$ *Update:* When $\boldsymbol{\alpha}$ is fixed, the update for $\mathbf{r}_x$ can be cast as a Quadratic Programming (QP) problem. This is shown by expressing the cost function (15a) in terms of $\mathbf{r}_x$. Define the matrix $\mathbf{E}_{p+1}^K := [\mathbf{e}_1^K, \cdots, \mathbf{e}_{p+1}^K]$, whose columns are the first $(p+1)$ canonical basis vectors in $\mathbb{R}^K$. Then, we have

$$\tilde{\mathbf{R}}_x\tilde{\boldsymbol{\alpha}} = \mathbf{R}_x^T\mathbf{E}_{p+1}^K\tilde{\boldsymbol{\alpha}} \tag{19a}$$

$$= \left(\sum_{k=-(K-1)}^{K-1} r_x(k)\boldsymbol{\Theta}_{-k}^K\right)\mathbf{E}_{p+1}^K\tilde{\boldsymbol{\alpha}} \tag{19b}$$

$$= \Bigg(r_x(0)\boldsymbol{\Theta}_0 + \sum_{k=1}^{K-1}\mathrm{Re}\{r_x(k)\}(\underbrace{\boldsymbol{\Theta}_k^K + \boldsymbol{\Theta}_{-k}^K}_{\boldsymbol{\Phi}_k^K}) + \sum_{k=1}^{K-1}\mathrm{Im}\{r_x(k)\}(\underbrace{j\boldsymbol{\Theta}_{-k}^K - j\boldsymbol{\Theta}_k^K}_{\boldsymbol{\Psi}_k^K})\Bigg)\underbrace{\mathbf{E}_{p+1}^K\tilde{\boldsymbol{\alpha}}}_{\boldsymbol{\gamma}} \tag{19c}$$

$$= r_x(0)\boldsymbol{\gamma} + \sum_{k=1}^{K-1}\mathrm{Re}\{r_x(k)\}\boldsymbol{\Phi}_k^K\boldsymbol{\gamma} + \sum_{k=1}^{K-1}\mathrm{Im}\{r_x(k)\}\boldsymbol{\Psi}_k^K\boldsymbol{\gamma} \tag{19d}$$

$$= \underbrace{[\boldsymbol{\gamma}, \boldsymbol{\Phi}_1^K\boldsymbol{\gamma}, \cdots, \boldsymbol{\Phi}_{K-1}^K\boldsymbol{\gamma}, \boldsymbol{\Psi}_1^K\boldsymbol{\gamma}, \cdots, \boldsymbol{\Psi}_{K-1}^K\boldsymbol{\gamma}]}_{\boldsymbol{\Pi}}\mathbf{r}_x \tag{19e}$$

$$= \boldsymbol{\Pi}\mathbf{r}_x \tag{19f}$$

where $\boldsymbol{\Pi} \in \mathbb{C}^{K\times 2K-1}$. Overall, we have a problem of the form

$$\min_{\mathbf{r}_x} \quad \|\boldsymbol{\Pi}\mathbf{r}_x - \mathbf{e}_1^K\|_2^2 \tag{20a}$$

$$\text{s.t.} \quad b_m(\mathbf{q}_m^T\mathbf{r}_x - t) \geq 0,\ \forall\, m \in \mathcal{M} \tag{20b}$$

$$\mathbf{F}\mathbf{r}_x \geq \mathbf{0} \tag{20c}$$

which is a QP problem in $\mathbf{r}_x$, and can be efficiently solved via convex programming.

The overall algorithm can be summarized as follows.

Algorithm 1 : AM for AR PS Estimation

Initialization: Solve (16) to obtain $\mathbf{r}_x^{(0)}$. Set $k := 0$.
Repeat

- Fix $\mathbf{r}_x^{(k)}$. Update $\boldsymbol{\alpha}^{(k+1)}$ according to (18).
- Fix $\boldsymbol{\alpha}^{(k+1)}$. Update $\mathbf{r}_x^{(k+1)}$ by solving the QP problem (20).
- Compute cost value $v^{(k+1)} = \|\tilde{\mathbf{R}}_x^{(k+1)}\tilde{\boldsymbol{\alpha}}^{(k+1)} - \mathbf{e}_1^K\|_2^2$
- Set $k := k + 1$.

Until Improvement in cost function $< tolerance\ factor$ in the last 10 iterations OR specified no. of iterations exceeded.

Since the update step for each variable is conditionally optimal given the other variable, the algorithm produces a monotonically non-increasing cost sequence. Moreover, we have the following proposition.

Proposition 1 *Every limit point of Algorithm 1 is a stationary point of* (15)*.*

Proof 1 *In [11], it is proven that every limit point of AM is a stationary point of the original problem, provided that the cost function of the original problem is continuously differentiable, each subproblem has a solution, and the constraint set corresponding to each variable block is closed and convex . Obviously, the first two conditions hold in our case. Regarding the final condition, the constraint set of* $\mathbf{r}_x$

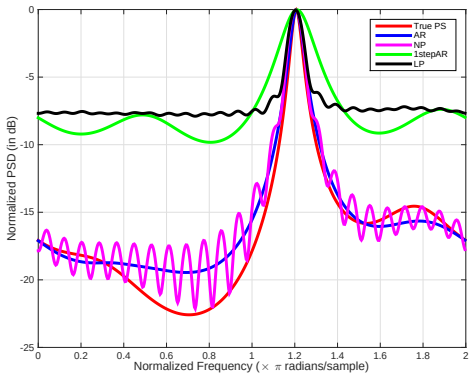

Fig. 1: Mean Normalized spectra for a complex AR(3) model

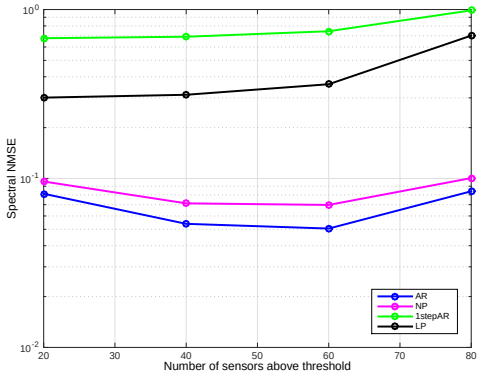

Fig. 2: NMSE vs No. of sensors above threshold for AR(5) models

is a convex polyhedron, which is closed, while $\boldsymbol{\alpha} \in \mathbb{C}^{p+1}$*, which is also closed and convex.*

Once an estimate of $\boldsymbol{\alpha}$ [1] is obtained from the algorithm, it can be plugged into (12) to generate an estimate of the AR PS.

5. NUMERICAL RESULTS

For the purpose of benchmarking our proposed AM algorithm, we consider the method of solving (16) followed by a single of step of fitting AR model parameters according to (18). This two-step approach corresponds to the traditional AR estimation method in our setup. The LP feasibility problem (16) and the QP problem (20) were modeled using YALMIP [12] and solved using the solver Sedumi [13]. In all experiments, the maximum iteration counter of the AM algorithm was set to 50 iterations (each iteration consisting of 2 alternating updates) and the exit tolerance factor was set to 10^{-5}. In order to qualitatively evaluate the performance of the methods with regard to the quality of the PS estimates generated, we selected the Normalized Mean Square Error (NMSE) as a performance criterion, which is defined as

$$\text{NMSE} = \mathbb{E}\left[\frac{\|S_x - \hat{S}_x\|_2^2}{\|S_x\|_2^2}\right] \tag{21}$$

where S_x is the true PS and $\hat{S}_x$ is the estimated PS, with both spectra normalized by their peak values. The expectation is taken with respect to the randomness of the signal and the broadband filters.

We now present an illustrative example showcasing the effectiveness of our approach. A complex AR(3) model was used to generate a WSS stochastic process, and a sensing scenario was considered with 100 sensors, $K = 25$ and threshold t empirically tuned to select 40 sensors to transmit $b_m = 1$ to the FC. Knowledge of the true model order is assumed *apriori*. Both methods were initialized from the same instance of the LP feasibility problem. The averaged results over 400 Monte-Carlo trials are depicted in Figure 1, with each PS estimate normalized by its peak value. The non-parametric LP initial estimate (black), obtained by taking the N_F point Discrete Fourier Transform (DFT) correctly estimates the peak of true AR PS, but the follow up AR fitting procedure (green) does not improve the estimation performance (in fact the spectral lobe widens). The parametric AR estimate (blue), obtained from the AM algorithm, exhibits the best estimation performance. Note that the quality of the spectral estimate is very satisfactory considering that only 100 bits (roughly equivalent to 6 floats in IEEE 16-bit precision standard) were used. Hence, even though the AM algorithm is incapable of solving (15) exactly, it generates high quality approximate solutions, as evidenced from the figure. We also plotted the non-parametric estimate obtained by taking the DFT of the autocorrelation vector returned by the AM algorithm (in magenta). Ripples appear in the spectral estimate since we only estimate a finite window of the autocorrelation sequence, which degrades its overall quality.

A more comprehensive simulation is presented in Figure 2, where we considered a sensing scenario with $M = 100$ sensors, set $K = 50$ and tuned the threshold t to vary the number of sensors reporting above threshold. The spectral NMSE was computed for 30 randomly drawn AR(5) models, with the NMSE for each value of t being averaged out over 100 Monte-Carlo trials for each AR model. Prior knowledge of the true model order was again assumed. Both methods were again initialized from the solution of the same instance of the LP feasibility problem. Again, the superior performance of the AM algorithm is noted, with the parametric AR estimate (blue) exhibiting lowest spectral NMSE. The one step AR fitting method is much worse-off in comparison. Extensive simulations across a range of model orders and sensing scenarios revealed that the AM algorithm always delivers the best performance, thus providing supporting evidence of its approximation quality.

6. CONCLUSIONS

A network sensing scenario was considered, consisting of scattered low-end sensors transmitting randomly filtered, one-bit quantized power measurements to a FC. Under the assumption that the underlying stochastic process can be parametrized by an autoregressive time series, the problem of estimating the ambient AR power spectrum from binary measurements was formulated as a non-convex optimization problem, which is hard to solve exactly in polynomial-time. Instead, the conditionally decomposable structure of the problem formulation was exploited to develop an alternating minimization algorithm, with guaranteed convergence to a stationary point of the original problem. The performance of the proposed algorithm was compared against the traditional AR estimation method based on obtaining a feasible autocorrelation estimate followed by a single step of fitting AR model parameters. Simulations revealed the superior performance of the AM algorithm with regard to spectral NMSE, even though it only yields approximate solutions.

[1] Although we established convergence to a stationary point of (15), there is no guarantee that the AR parameter estimates obtained are minimum-phase, since the solution of the extended YWEs is not guaranteed to be minimum-phase in general.

7. REFERENCES

[1] D. D. Ariananda, and G. Leus, "Compressive Wideband Power Spectrum Estimation," *IEEE Trans. Signal Processing*, vol. 60, no. 9, pp. 4775–4789 , Sep. 2012.

[2] D. Romero, and G. Leus, "Wideband spectrum sensing from compressed measurements using spectral prior information," *IEEE Trans. Signal Processing*, vol. 61, no. 24, pp. 6232–6246, Dec. 2013.

[3] O. Mehanna, and N. D. Sidiropoulos, "Frugal Sensing: Wideband power spectrum sensing from few bits," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2693–2703, May 2013.

[4] A. Konar, N. D. Sidiropoulos, and O. Mehanna, "Parametric Frugal Sensing of Power Spectra for Moving Average Models," *IEEE Trans. Signal Processing*, vol. 63, no. 5, pp. 1073–1085, March 2015.

[5] A. Konar and N. D. Sidiropoulos," Hidden convexity in QCQP with Toeplitz-Hermitian quadratics," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1623–1627, Oct. 2015.

[6] S. M. Kay and S. L. Marple Jr, "Spectrum Analysis - A Modern Perspective," *Proc. of the IEEE*, vol. 61, no. 11, pp 1380-1419, Nov. 1981.

[7] R. Gray and L. Davidson, *Random Processes: A Mathematical approach for Engineers*, Englewood Cliffs, NJ, USA: Prentice Hall, 1986.

[8] P. Stoica, and R. L. Moses, *Spectral Analysis of Signals*, Upper Saddle River, NJ, USA: Prentice Hall, 2005.

[9] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, USA, 1996.

[10] S. M. Kay, *Fundamentals of statistical signal processing: Estimation Theory*, Upper Saddle River, NJ, USA: Prentice Hall, 1993.

[11] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Oper. Res. Lett.*, vol. 26, no. 3, pp. 127-136, April 2000.

[12] J. Lofberg, "Yalmip: A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD*, Taipei, Sep. 4, 2004.

[13] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1, pp. 625–653, 1999.