MODEL-BASED POWER SPECTRUM SENSING FROM A FEW BITS

Omar Mehanna, Nicholas D. Sidiropoulos[*]

Dept. of ECE, Univ. of Minnesota
Minneapolis, MN 55455, USA

Efthymios Tsakonas[†]

ACCESS Linnaeus Centre, Royal Inst. Tech.
(KTH) 100 44 Stockholm, Sweden

ABSTRACT

Wideband power spectrum sensing is fundamental for numerous applications. When side information on the potentially active emitters is available, such as carriers and spectral masks, it should be exploited to improve sensing performance. Here the power spectrum is modeled as a weighted sum of candidate spectral density primitives. The objective is to estimate the unknown weights from a few randomly filtered broadband power measurement bits, taken using a network of low-end sensors. A linear programming formulation that exploits the sparsity in the unknown weights is proposed. A better approach follows, which exploits the approximately Gaussian distribution of the errors in the power measurements prior to quantization, in a maximum likelihood formulation that includes a sparsity-inducing penalty term. Simulations show that the model weights can be accurately estimated from few bits, even when the errors are significant.

1. INTRODUCTION

Wideband spectrum sensing is fundamental for many applications. Most work on spectrum sensing has focused on reconstructing the signal itself [1]. In cognitive radio and other applications, however, only the *power spectrum* (PS) is needed, i.e., the Fourier transform of the signal's autocorrelation. Based on this concept, the sampling rate requirements can be considerably decreased, without requiring spectrum sparsity [2]. Further reduction in the sampling rate is possible by exploiting available prior information on the spectrum, such as spectral masks, carrier frequencies, and bandwidths. In [3], a mixture of statistically independent signals is considered, where the PS of each signal is known up to a scaling factor representing the transmission power, which is estimated using sub-Nyquist measurements.

It is now widely accepted that effective spectrum sensing must be a collaborative endeavor, involving many sensors taking relatively sparse measurements across space, time, and frequency. Cooperative sensing and detection schemes are shown to increase reliability, reduce average detection time, help cope with fading, and improve throughput [4]. A parsimonious basis expansion model for the PS has been considered in [4], where multiple sensors cooperate in estimating the basis coefficients (i.e, powers), by exploiting sparsity as prior information to improve the estimation performance. In distributed spectrum sensing scenarios, however, sending analog or finely quantized signal sample streams to a fusion center (FC) is a heavy burden in terms of communication overhead and battery lifetime.

This assumption was subsequently relaxed in [5], where a network sensing scenario, comprising scattered low-end sensors and a FC, was considered. Each sensor reports a single randomly filtered power measurement bit to the FC, which then estimates the ambient PS from the collected bits. A linear programming (LP) formulation was introduced in [5], generalizing classical nonparametric PS estimation to the case where the data is in the form of inequalities, rather than equalities, and exploiting the autocorrelation parametrization and pertinent nonnegativity properties. This setting has been extended in [6] to handle line spectra, using a maximum likelihood formulation.

In this paper, we consider the same network sensing scenario as [5, 6], and a PS model similar to [3, 4]. The overall PS is modeled as a weighted sum of candidate spectral density primitives, where the weights are unknown. Assuming accurate power measurements at the sensors, an LP formulation is proposed to estimate the unknown model weights by exploiting their sparsity, thus obtaining a full PS estimate. The problem formulation in this case may be reminiscent of *one-bit compressed sensing* [7]. Unlike [7], we operate on the PS, not the signal *per se*, which enables us to exploit positivity constraints that are not present in the one-bit compressed sensing framework. In addition, the choice of (positive) thresholds mitigates the scaling problem, so we do not use the unit sphere constraint imposed in [7]. The performance (i.e., estimation error vs. measurement bits) is significantly improved by exploiting the additional constraints in our formulation.

Next, the errors in the power measurements prior to quantization are considered. The distribution of these errors can be shown to be approximately Gaussian, due to sample averaging and the effects of frequency-selective fading. This is then exploited in a maximum likelihood (ML) formulation that in-

[*]E-mail: (meha0006,nikos)@umn.edu, Tel: +16126251242, Fax: +16126254583. Supported in part by NSF ECCS 1231504, and NSF AST-1247885.

[†]E-mail: tsakonas@kth.se

cludes a sparsity-inducing penalty term. Interestingly, the ML formulation is similar to the formulation considered in [8] for a seemingly very different problem: so-called *conjoint analysis*. In order to reduce the number of bits transmitted from the sensors so as to prolong battery lifetimes and to minimize the communication overhead [9], a censoring scheme is also proposed, where only sensors that provide the most useful information bits are permitted to send, while other *less-informative* sensors remain silent. Simulations show that the underlying model weights can be accurately estimated using relatively few bits, even when the errors are significant.

2. PROBLEM FORMULATION

2.1. Setup

Consider M scattered sensors measuring the ambient signal power and reporting to a FC. The received signal at sensor $m \in \{1, \ldots, M\}$ is sampled using an analog-to-digital converter operating at Nyquist rate, yielding the discrete-time signal $y_m(n)$. Constant scaling differences in $\{y_m(n)\}_{m=1}^M$ across sensors due to path loss, shadowing, and frequency-flat fading can be factored out via automatic gain control (AGC), and the Nyquist sampling requirement can be lifted by using an equivalent analog processing and integration chain, as shown in [5]. In presence of frequency-selective fading, the received signal $y_m(n)$ is the convolution of the primary wide-sense stationary (WSS) signal $x(n)$ with the linear finite impulse response (FIR) fading channel $\{h_m(\ell)\}_{\ell=0}^{T-1}$, expressed as $y_m(n) = \sum_{\ell=0}^{T-1} h_m(\ell)x(n-\ell)$, where T is the number of channel taps. The channel is assumed time-invariant during a sensing epoch.

Sensor $m \in \{1, \ldots, M\}$ then passes $y_m(n)$ through a wideband FIR filter with random complex pseudo-noise (PN) impulse response $g_m(n)$ of length $K > T$, where:

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

The filter sequence $g_m(n)$ can be generated using a PN linear shift register, whose initial seed is unique for each sensor (e.g., its serial number), and is assumed to be known to the FC. Using random PN filters can also be motivated from a random projections viewpoint, as for the compression matrix applied to sparse signals [7]. The filter's output sequence is expressed as $z_m(n) = \sum_{k=0}^{K-1} g_m(k)y_m(n-k)$.

Let α_m denote the average power of the WSS signal $z_m(n)$, i.e., $\alpha_m := \mathbb{E}[|z_m(n)|^2]$. Each sensor estimates α_m using a sample average:

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

Finally, each sensor compares the estimated $\hat{\alpha}_m$ to a sensor-specific threshold t_m. If $\hat{\alpha}_m \ge t_m$, then sensor m sends a bit $b_m = 1$ to the FC, otherwise it sends $b_m = -1$.

2.2. Signal Model

We assume that the PS of the signal $x(n)$ can be expressed (approximated) by the model:

$$S_x(\omega) = \sum_{\ell=1}^{L} \rho_\ell \Psi_\ell(\omega - \omega_\ell) \tag{3}$$

where $\Psi_\ell(\omega)$ can be any function. For example, $\Psi(\omega) = \delta(\omega)$ corresponds to a spectral line [6]. Another example is the (overlapping) raised cosine bases which can model transmit-spectra of multicarrier systems [4]. In general, the PS model (3) can describe the spectrum of L primary transmitters, where each transmitter $\ell \in \{1, \ldots, L\}$ is transmitting at center (angular) frequency ω_ℓ, with transmit-power ρ_ℓ, and a spectral mask that is characterized by the function $\Psi_\ell(\omega)$. In this setting we assume that the large-scale fading channel between transmitter ℓ and sensor m is known a-priori (e.g., using training sequences), and is removed from the received signal. Note that the bandwidth can be different across $\{\Psi_\ell(\omega)\}_{\ell=1}^L$.

If L and $\{\Psi_\ell(\omega)\}_{\ell=1}^L$ are known and L is relatively small (e.g. $L < 5$), the parametric ML estimation technique that has been considered in [6] for estimating line spectra can be used to estimate $\{\omega_\ell\}_{\ell=1}^L$ and $\{\rho_\ell\}_{\ell=1}^L$ for the PS model (3). On the other hand, the estimation problem can be linearized by considering all possible center frequencies $\{\omega_\ell\}_{\ell=1}^L$ if L is large or unknown (i.e., an overcomplete dictionary), such that the (sparse) weights $\{\rho_\ell\}_{\ell=1}^L$ are the only unknowns in the PS model (3). This is the scenario considered herein.

Let $r_x(k) := \mathbb{E}[x(n)x^*(n-k)]$ denote the autocorrelation sequence of $x(n)$. From (3), the autocorrelation function can be expressed as:

$$\begin{aligned} r_x(k) &= \frac{1}{2\pi} \int_0^{2\pi} S_x(\omega) e^{j\omega k} d\omega \\ &= \sum_{\ell=1}^{L} \psi_\ell(k) \rho_\ell e^{j\omega_\ell k} \end{aligned} \tag{4}$$

where $\psi_\ell(k) := \frac{1}{2\pi} \int_0^{2\pi} \Psi_\ell(\omega) e^{j\omega k} d\omega$ is the inverse discrete-time Fourier transform (DTFT) of $\Psi_\ell(\omega)$.

Define $\tilde{z}_m(n) := \sum_{k=0}^{K-1} g_m(k)x(n-k)$ as the convolution of the primary signal $x(n)$ and the random filter $g_m(n)$ (i.e., ignoring fading), and let $\tilde{\alpha}_m := \mathbb{E}[|\tilde{z}_m(n)|^2]$. Also, define the deterministic autocorrelation of $g_m(n)$ as:

$$q_m(k) := \sum_{\ell=0}^{K-1} g_m(\ell) g_m^*(\ell + k) \tag{5}$$

Hence, it can be shown that:

$$\begin{aligned} \tilde{\alpha}_m &= \sum_{k=1-K}^{K-1} r_x(k) q_m^*(k) \\ &= \sum_{\ell=1}^{L} \rho_\ell \sum_{k=1-K}^{K-1} \psi_\ell(k) e^{jk\omega_\ell} q_m^*(k) \\ &= \mathbf{v}_m^T \boldsymbol{\rho} \end{aligned} \tag{6}$$

where $\boldsymbol{\rho} := [\rho_1, \ldots, \rho_L]^T$ and $v_m(\ell) := \sum_{k=1-K}^{K-1} \psi_\ell(k) e^{jk\omega_\ell} q_m^*(k)$ is the ℓ-th entry of the vector $\mathbf{v}_m$, $\ell = 1, \ldots, L$. The filter length K should be sufficiently large to ensure that the random filtering (1) yields non-redundant vectors $\{\mathbf{v}_m\}_{m=1}^M$ with high probability. Note that $\boldsymbol{\rho}$ can be upper bounded by $\mathbf{p}_{\max}$ (due to the use of AGC at the front-end of the sensor processing chain). This yields the box constraint $\boldsymbol{\rho} \in \mathcal{B}$, where $\mathcal{B} := \{\boldsymbol{\rho} \in \mathbb{R}^L \mid \mathbf{0} \leq \boldsymbol{\rho} \leq \mathbf{p}_{\max}\}$.

2.3. Error Model

Assuming that $\{h_m(\ell)\}_{\ell=0}^{T-1}$ are i.i.d. complex Gaussian random variables with zero-mean and $1/T$-variance, it can be shown that the errors due to fading, defined as $\tilde{e}_m := \alpha_m - \tilde{\alpha}_m$, can be approximated as i.i.d. Gaussian random variables with zero-mean and variances $\{\tilde{\sigma}_m^2\}_{m=1}^M$ using the Lyapunov central limit theorem when T is large. Furthermore, the estimation errors due to insufficient sample averaging, $\bar{e}_m := \hat{\alpha}_m - \alpha_m$, can also be modeled as i.i.d. Gaussian random variables with zero-mean and variances $\{\bar{\sigma}_m^2\}_{m=1}^M$, by the central limit theorem. This means that $\hat{\alpha}_m$ can be modeled as: $\hat{\alpha}_m = \tilde{\alpha}_m + e_m$, where $e_m := \tilde{e}_m + \bar{e}_m$ is a Gaussian random variable with zero mean and variance $\sigma_m^2 = \tilde{\sigma}_m^2 + \bar{\sigma}_m^2$. Thus, the power measurement bit of each sensor can be expressed as: $b_m = \text{sign}(\mathbf{v}_m^T \boldsymbol{\rho} + e_m - t_m)$. Note that the performance with errors is identical to the error-free case as long as no measurement bits are flipped.

Objective. Given the measurement bits $\{b_m\}_{m=1}^M$, the goal is to estimate the sparse vector $\boldsymbol{\rho}$. In the following two sections, different formulations are proposed for estimating $\boldsymbol{\rho}$.

3. LINEAR PROGRAM FORMULATION

In this first approach, we assume sufficiently small errors such that $\text{sign}(\hat{\alpha}_m - t_m) = \text{sign}(\tilde{\alpha}_m - t_m), \forall m$. This means that the FC learns that: $b_m(\tilde{\alpha}_m - t_m) \geq 0$, from each received bit b_m. The linear inequality constraints: $b_m(\mathbf{v}_m^T \boldsymbol{\rho} - t_m) \geq 0$, for $m = 1, \ldots, M$, in addition to the box constraint $\mathcal{B}$, define the feasible region (polyhedron) for $\boldsymbol{\rho}$. In order to find a *good* estimate for $\boldsymbol{\rho}$ inside the feasible region, the sparsity of $\boldsymbol{\rho}$ can be exploited by minimizing the sparsity-inducing ℓ_1-norm: $||\boldsymbol{\rho}||_1 = \sum_{\ell=1}^L \rho_\ell$, for $\boldsymbol{\rho} \geq \mathbf{0}$. This yields the following LP:

$$\boxed{\begin{aligned} &\min_{\boldsymbol{\rho} \in \mathcal{B}} \quad \sum_{\ell=1}^{L} \rho_\ell \\ &\text{s.t.}: \quad b_m(\mathbf{v}_m^T \boldsymbol{\rho} - t_m) \geq 0, \quad m = 1, \ldots, M. \end{aligned}} \tag{7}$$

The LP (7) is efficiently solved using specialized algorithms. The solution of (7) is the sparsest solution attainable via the ℓ_1-norm that satisfies the constraints, whereas sparser solutions can be achieved using the iterative weighted ℓ_1-norm algorithm of [10].

It is worth mentioning that problem (7) is similar to the nonparamteric estimation problem considered in [5], where the autocorrelation function is first estimated for a finite number of lags, then Fourier transformed to obtain a PS estimate. It is also worth mentioning that, unlike the 1-bit compressed sensing framework of [7], the sphere constraint, i.e., $||\boldsymbol{\rho}||_2 = 1$, is not required in (7) due to the positivity of $\boldsymbol{\rho}$ and $\{t_m\}_{m=1}^M$.

The assumption that $e_m \approx 0, \forall m$, enables using $\hat{\alpha}_m = \mathbf{v}_m^T \boldsymbol{\rho}$ in the constraints of (7). As the error variances, $\{\sigma_m^2\}_{m=1}^M$, increase, the estimation problem (7) becomes inaccurate, and the constraints may become inconsistent as the number of flipped bits due to errors increase. This motivates exploiting the Gaussian distribution of $\{e_m\}_{m=1}^M$ in a maximum-likelihood formulation, which is presented in the next section.

4. MAXIMUM LIKELIHOOD FORMULATION

In this approach, the Gaussian distribution of $\{e_m\}_{m=1}^M$ is used in defining a convex optimization problem that yields a (sparse) ML estimate of the power vector $\boldsymbol{\rho}$.

Define $\mathcal{M}_+ := \{m | b_m = 1\}$ and $\mathcal{M}_- := \{m | b_m = -1\}$. Since the error samples $\{e_m\}_{m=1}^M$ are i.i.d Gaussian random variables with zero-mean and variance $\{\sigma_m^2\}_{m=1}^M$, the probability of receiving the bits $b_1, \ldots, b_M$, given $\boldsymbol{\rho}$, is:

$$\begin{aligned} &f(b_1, \ldots, b_M | \boldsymbol{\rho}) \\ &\quad = \prod_{m \in \mathcal{M}_+} \Pr(\tilde{\alpha}_m + e_m \geq t_m) \prod_{m \in \mathcal{M}_-} \Pr(\tilde{\alpha}_m + e_m < t_m) \\ &\quad = \prod_{m \in \mathcal{M}_+} \Phi\left(\frac{\mathbf{v}_m^T \boldsymbol{\rho} - t_m}{\sigma_m}\right) \prod_{m \in \mathcal{M}_-} \Phi\left(\frac{t_m - \mathbf{v}_m^T \boldsymbol{\rho}}{\sigma_m}\right) \end{aligned} \tag{8}$$

where $\Phi(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt$ is the cumulative distribution function (CDF) of the Gaussian distribution. The log-likelihood function can thus be written as:

$$\log f(b_1, \ldots, b_M | \boldsymbol{\rho}) = \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\mathbf{v}_m^T \boldsymbol{\rho} - t_m)}{\sigma_m}\right) \tag{9}$$

Similar to (7), the sparsity of $\boldsymbol{\rho}$ can be exploited by adding the sparsity-inducing penalty $||\boldsymbol{\rho}||_1$ to the cost function. Therefore, a *sparse ML* estimate of $\boldsymbol{\rho}$ is obtained by solving the optimization problem:

$$\boxed{\max_{\boldsymbol{\rho} \in \mathcal{B}} \sum_{m=1}^{M} \log \Phi\left(\frac{b_m(\mathbf{v}_m^T \boldsymbol{\rho} - t_m)}{\sigma_m}\right) - \lambda \sum_{\ell=1}^{L} \rho_\ell} \tag{10}$$

where $\lambda \geq 0$ is a tuning parameter that controls the sparsity of the solution. Since the Gaussian CDF $\Phi(\cdot)$ is log-concave [11, pp. 104], problem (10) is convex and can be solved efficiently using interior-point algorithms. Note that the maximizer of problem (10) always exists since $\boldsymbol{\rho}$ is bounded in $\mathcal{B}$. It is worth mentioning that (10) is similar to the formulation considered in [8].

Constrained Cramer-Rao Bound. In order to find a lower bound on the estimation error variance of (10), we compute the Cramer-Rao bound (CRB) assuming perfect knowledge of the support set of $\boldsymbol{\rho}$ (i.e, oracle performance). Define κ as the number of nonzero entries (cardinality) of $\boldsymbol{\rho}$, i.e., $||\boldsymbol{\rho}||_0 = \kappa$. The CRB with cardinality constraint was studied in [12]. First, the Fisher information matrix (FIM) is computed as [8]:

$$\mathbf{J} = \sum_{m=1}^{M} \frac{e^{\frac{-(\mathbf{v}_m^T\boldsymbol{\rho}-t_m)^2}{\sigma_m^2}}}{2\pi\sigma_m^2} \left[\frac{1}{\Phi\left(\frac{\mathbf{v}_m^T\boldsymbol{\rho}-t_m}{\sigma_m}\right)} + \frac{1}{\Phi\left(\frac{t_m-\mathbf{v}_m^T\boldsymbol{\rho}}{\sigma_m}\right)} \right] \mathbf{v}_m\mathbf{v}_m^T \tag{11}$$

Let $\mathrm{supp}(\boldsymbol{\rho})$ denote the support set of $\boldsymbol{\rho}$, and define the $L \times \kappa$ matrix $\mathbf{U}$ as the matrix of feasible directions consisting of the subset of columns of the identity matrix corresponding to the set $\mathrm{supp}(\boldsymbol{\rho})$. The (constrained) CRB is thus computed as [12]:

$$\mathrm{Trace}(\mathbf{U}(\mathbf{U}^T\mathbf{J}\mathbf{U})^{-1}\mathbf{U}^T) \tag{12}$$

This gives the best achievable mean square error obtained by estimators that have perfect knowledge of the support set of the vector to be estimated.

Censoring. It is important to reduce the number of bits transmitted from the sensors so as to prolong battery lifetimes and to minimize the communication overhead [9]. With censoring, only sensors that provide the most useful information bits are permitted to send, while other *less-informative* sensors remain silent. It is easy to see that the log-likelihood function in (9) and the FIM in (11) are almost unaffected if the value of $|\tilde{\alpha}_m - t_m|/\sigma_m$ is too large, because of the properties of the Gaussian CDF, i.e., $\Phi(u) \approx 1$ (≈ 0) if $u > 4$ (resp. $u < -4$). This means that the measurement bit b_m is almost useless in the estimation problem if $|\tilde{\alpha}_m - t_m|/\sigma_m$ is too large. Thus, censoring can be employed such that sensor m sends b_m only if $|\hat{\alpha}_m - t_m| \leq \zeta_m$, where $\zeta_m > 0$ is a censoring threshold that is known at the sensor. More insight on the choice of ζ_m is discussed in the following section.

5. NUMERICAL RESULTS

To test the performance of the proposed estimation techniques, we assume that the primary signal PS model is the combination of $L = 20$ identical raised cosine functions, each with roll-off factor $= 0.5$ and bandwidth $= 3\pi/L$. The corresponding center frequencies are $\{\omega_\ell = 2\pi(\ell-1)/L\}_{\ell=1}^{L}$, implying equispaced and overlapping functions. We use the mean-square error (MSE), defined as $\mathrm{MSE}_\rho := E[||\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}||^2]$, to measure the performance of the proposed estimation techniques, where the expectation is taken with respect to the random impulse responses of the FIR filters, the random error samples, and the random power vector $\boldsymbol{\rho}$, obtained via 1000 Monte-Carlo simulation runs. Typically, the MSE should be computed with respect to the estimated PS; however, the symmetry in the considered PS model allows using MSE_ρ instead.

A sparse vector $\boldsymbol{\rho}$ with 5 uniformly distributed nonzero entries out of 20 is randomly generated and normalized by $\sum_{\ell=1}^{L} \rho_\ell$ in each simulation run. For brevity, we name the estimate of $\boldsymbol{\rho}$ obtained using (7) as the *LP estimate*, whereas the estimate obtained using (10) is named the *ML estimate*. A single threshold $t_m = t = 0.052$ and a filter length $K = 25$ were used at all sensors. This t is numerically computed as the minimizer of the expected CRB across different $\boldsymbol{\rho}$ and filter realizations. The sparsity tuning parameter in (10) was fixed to $\lambda = 100$ in the simulations. For the same error variance, $\sigma_m = \sigma$, $\forall m$, we define the signal-to-error ratio (SER) as: $\gamma := \sum_{\ell=1}^{L} \rho_\ell \psi_\ell(0)/\sigma^2$. The MSE of the LP estimate, the MSE of the ML estimate, and the (oracle) CRB computed using (12), are plotted versus the number of sensors M in Figs. 1 and 2, for $\gamma = 5 \times 10^4$ and $\gamma = 500$, respectively.

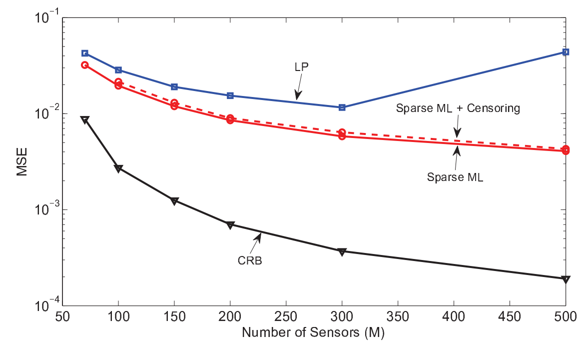

Fig. 1. The MSE of the LP estimate, the MSE of the ML estimate, and the (oracle) CRB, versus M, at large SER.

In Fig. 1, where a relatively large SER is considered, the random errors in this case cause the flipping of 2% of the sensor measurement bits, on average. In other words, $\mathrm{sign}(\tilde{\alpha}_m - t_m) \neq \mathrm{sign}(\hat{\alpha}_m - t_m)$ for $0.02M$ sensors, on average. The figure shows the decrease of the CRB and the MSE of the ML estimate as M increases, as expected. The figure also shows that the MSE of the LP estimate is decreasing and is close to MSE of the ML estimate for $M \leq 300$, then it increases when M increases to 500. The reason is that as M increases, the constraints in (7) become more stringent, and the flipped bits due to errors drive the solution far away from the true estimate. In this setting, a feasible solution is guaranteed at $\hat{\rho}_\ell = t/\nu$, $\forall \ell$, where $\nu := \sum_{\ell=1}^{L} v_m(\ell) = L\psi(0)$, which is constant for all m. In other more general settings, the constraints in (7) can become inconsistent with large M when many bits are flipped.

The performance of the censoring scheme proposed at the end of Section 4 is also considered in Fig. 1. The dashed line in the figure shows the MSE of the ML estimate when the censoring threshold ζ was selected such that only the *best* 80 out of M sensors are active in each simulation run (i.e., $|\hat{\alpha}_m - t| \leq \zeta$ for 80 sensors), on average. As shown in the figure, the performance with censoring is very close to the case when all sensors are reporting. Censoring in this case is very efficient for large M (e.g., performance with 80/500

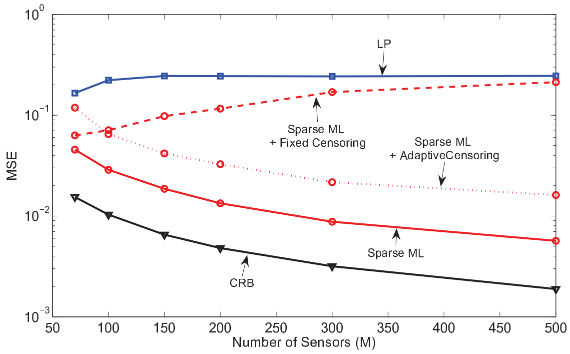

Fig. 2. The MSE of the LP estimate, the MSE of the ML estimate, and the (oracle) CRB, versus M, at small SER.

reporting sensors almost same as all 500 reporting sensors).

In Fig. 2, where a relatively small SER is considered, the random errors in this case cause the flipping of 16% of the sensor measurement bits, on average. The figure shows the decrease of the CRB and the MSE of the ML estimate as M increases, albeit at larger values than their counterparts in Fig. 1, due to the increased flipped bits due to errors. It is interesting to see that good ML estimates can be obtained, despite the flipping of so many measurement bits. We also note that the gap between the MSE of the ML estimate and the CRB decreased as the SER decreased from Fig. 1 to Fig. 2. The performance of the LP estimate in this figure is very limited by the increased flipped bits, which is not accounted for in (7).

The MSE of the ML estimate when employing the same *fixed censoring* setting as in Fig. 1, represented by the dashed line in Fig. 2, is increasing with M. This is because the number of flipped bits among the 80 measurement bits that are reported each simulation run, on average, increases as M increases, which deteriorates the performance. Instead, we considered an adaptive censoring scheme, where the censoring threshold ζ was selected such that 67% of the sensors (i.e., $0.67M$) are reporting in each simulation run, on average. The MSE of the ML estimate when this censoring scheme was employed is represented by the dotted line in Fig. 2. Although the adaptive censoring performance is improving with M, we can see that the performance is significantly worse than the case when all sensors are reporting. Therefore, we conclude that when the SER is small, it is better that all sensors report to combat the increasing number of flipped measurement bits due to errors, whereas censoring is more efficient when the SER is large.

6. CONCLUSIONS

Assuming available prior information on the spectrum, such as spectral masks, center frequencies, and bandwidths, an over-complete dictionary model can be considered for the PS. We proposed an LP formulation that exploits sparsity to estimate the unknown PS model weights when accurate measurements are assumed at the sensors, and an ML formulation with a sparsity-inducing penalty that exploits the Gaussian distribution of the errors in the sensor measurements prior to quantization. Simulations have shown that the underlying model weights can be accurately estimated using relatively few bits, even when the errors are significant.

7. REFERENCES

[1] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Processing Mag.,* vol. 29, no. 3, pp. 101–116, May 2012.

[2] D.D. Ariananda, and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Processing*, vol. 60, no. 9, pp. 4775–4789 , Sep. 2012.

[3] D. Romero, G. Leus, and R. Lopez-Valcarce, "Compressive wideband spectrum sensing with spectral prior information" *in proc. of the 38th International Conference on Acoustics, Speech, and Signal Processing,* Vancouver, Canada, May 26-31, 2013.

[4] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Group-Lasso on splines for spectrum cartography," *IEEE Transactions on Signal Processing,* vol. 59, no. 10, pp. 4648-4663, October 2011.

[5] O. Mehanna and N. D. Sidiropoulos, "Frugal sensing: wideband power spectrum sensing from few bits," *IEEE Transactions on Signal Processing,* June 2013 (To appear).

[6] O. Mehanna, N. D. Sidiropoulos, and E. Tsakonas, "Line spectrum estimation from broadband power detection bits," *submitted to the 14th Workshop on Signal Processing Advances in Wireless Communications,* Darmstadt, Germany, June 16-19, 2013.

[7] P. Boufounos and R. Baraniuk, "1-bit compressive sensing," *in Proc. Conf. Inform. Science and Systems (CISS),* Princeton, NJ, Mar. 2008.

[8] E. Tsakonas, J. Jalden, N. D. Sidiropoulos, and B. Ottersten, "Sparse conjoint analysis through maximum likelihood estimation," *submitted to IEEE Transactions on Signal Processing,* Oct. 2012.

[9] E. J. Msechu and G. B. Giannakis, "Sensor-centric data reduction for estimation with WSNs via censoring and quantization," *IEEE Transactions on Signal Processing,* vol. 60, no. 1, pp. 400-414, January 2012.

[10] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, Dec. 2008.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[12] Z. Ben-Haim and Y. C. Eldar, "The Cramer-Rao bound for estimating a sparse parameter vector," *IEEE Transactions on Signal Processing,* vol. 58, no. 6, June 2010.