Power Spectrum Blind Sampling Using Minimum Mean Square Error and Weighted Least Squares

Bamrung Tausiesakul and Nuria González-Prelcic
E.E. Telecomunicación
Universidade de Vigo
36313 Vigo, Spain
bamrung@gts.uvigo.es, nuria@gts.uvigo.es

***Abstract*—We present a new power spectrum recovery method in the context of power spectrum blind sampling. As sampling device we propose a multicoset sampler, which provides sub-Nyquist rate samples. A weighted least squares (WLS) criterion is adopted with the aim to define a power spectrum recovery algorithm that minimizes the mean square error (MSE) of the correlation estimate of the input signal. It is analytically shown that the optimal weighting matrix is equal to the inverse of the covariance matrix of the correlation estimate of the sub-Nyquist rate samples. The derived weight can also be shown to be optimal in MSE sense for power spectrum estimation. We also provide an optimization framework for the design of multicoset sampling patterns that minimize the MSE of the compressive WLS power spectrum estimator. The resulting integer nonlinear programming problem is solved by using exhaustive search.**

I. INTRODUCTION

Power spectrum blind sampling (PSBS), or equivalently, compressive power spectrum estimation techniques [1,2], define a sub-Nyquist rate sampling strategy and an efficient reconstruction method for the power spectrum of a wide sense stationary signal. These techniques are of particular interest in spectrum sensing for cognitive radio [3]–[7], since they combine the efficiency of compressed sensing when acquiring the sub-Nyquist rate samples needed to reconstruct the power spectrum, with the simplicity of some power spectrum estimators based on least squares (LS) solutions.

Although different sampling strategies are possible, previous works [1,2,8] concentrate on multicoset sampling [9,10] using minimal sparse rulers as sampling patterns, which provide a very low compression rate. Coprime sampling [11]–[13] has been also proposed as an interesting structure to obtain the sub-Nyquist rate samples, which can be combined with traditional power spectrum estimators which do not make use of any sparsity assumption on the input signal. Compressive sampling can be used without the sparsity assumption because the objective is not to obtain perfect signal reconstrution, but rather to perform power spectrum recovery.

This paper introduces a new criterion for the design of the sampling pattern, so that the mean square error (MSE) of the power spectrum estimate is minimized, while keeping minimal compression rate. In addittion, this work proposes a different power spectrum estimator from the sub-Nyquist rate samples, which makes use of the weighted least squares (WLS) criterion. Finally, we provide an optimization framework for the design of the minimum MSE (MMSE) multicoset sampling pattern. The resulting integer nonlinear programming problem is solved using exhaustive search. Numerical examples show that the proposed compressive power spectrum estimation strategy provides a lower MSE in the power spectrum reconstruction than the one obtained with the minimal-sparse-ruler-based pattern and the LS criterion.

Work supported by the European Regional Development Fund (ERDF) and the Spanish Government under projects TEC2010-21245-C02-02/TCM and CONSOLIDER-INGENIO 2010 CSD2008-00010, and the Galician Regional Government under projects CN 2012/260 "Consolidation of Research Units: AtlantTIC.

II. PROBLEM STATEMENT

Consider a complex-valued wide-sense stationary signal $x(t)$ with bandwith B. Our aim is to sample this signal at a rate lower than the Nyquist frequency $1/T$, such that the power spectrum of $x(t)$ can be accurately estimated.

For the acquisition stage, we consider a multicoset sampling strategy [9] implemented with M interleaved analog-to-digital converters working at a rate $1/NT$, being $1/T$ the Nyquist sampling rate and N the block length. This sampling device can be modeled as in [1,2]: a high rate integrate and dump process followed by a bank of M branches, consisting each one of a filtering operation followed by a downsampling operation, as illustrated in Figure 1. Taking into account that multicoset sampling consists of selecting M Nyquist-rate samples in each block of length N, the coefficients of the filter $c_i[n]$, $i = 1, \ldots, M$, can be written as

$$c_i[n] = \begin{cases} 1, & n = -n_i, \\ 0, & n \neq -n_i, \end{cases} \tag{1}$$

where there is no repetition in n_i, i.e.

$$n_i \neq n_j, \quad \forall i \neq j. \tag{2}$$

The output of the i-th branch of this PSBS scheme is given by

$$y_i[k] = z_i[kN], \tag{3}$$

where $z_i[\cdot]$ is given by

$$z_i[n] = c_i[n] * x[n] = \sum_{m=1-N}^{0} c_i[m]x[n-m]. \tag{4}$$

In compressive power spectrum estimation, the objective is to estimate the power spectrum of $x(t)$ from the sub-Nyquist samples $\{y_i[k]\}_{i,k}$, that is, to estimate the power spectrum

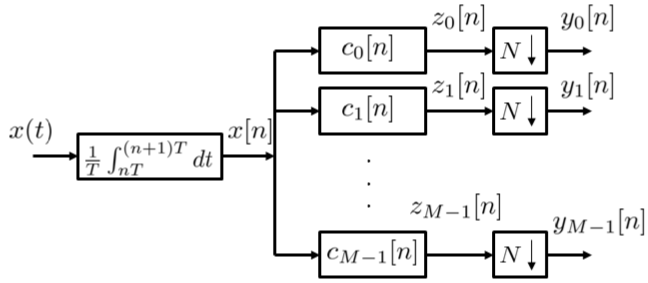

Fig. 1. Digital model of the sampling device.

spectrum of $x[n]$, which is equivalent to obtain the autocorrelation function of $x[n]$ given by $r_x[n] = \mathcal{E}\{x[m]x^*[m-n]\}$. In [2], Ariananda and Leus propose a method to recover the autocorrelation function $r_x[n]$ given the cross correlations $r_{y_i,y_j}[k]$, for $i, j = 0, \ldots, M-1$, given by

$$r_{y_i,y_j}[k] = \mathcal{E}\{y_i[l]y_j^*[l-k]\}. \tag{5}$$

They show that

$$r_{y_i,y_j}[k] = \sum_{l=0}^{1} \mathbf{r}_{c_i,c_j}^{\mathsf{T}}[l]\mathbf{r}_x[k-l], \tag{6}$$

where

$$\begin{aligned} \boldsymbol{r}_{c_i,c_j}[k] = \big[r_{c_i,c_j}[kN] \quad & r_{c_i,c_j}[kN-1] \\ \cdots \quad & r_{c_i,c_j}[(k-1)N+1]\big]^{\mathsf{T}}, \end{aligned} \tag{7a}$$

$$\begin{aligned} \boldsymbol{r}_x[k] = [r_x[kN] \quad & r_x[kN+1] \\ \cdots \quad & r_x[(k+1)N-1]]^{\mathsf{T}}. \end{aligned} \tag{7b}$$

By cascading all these cross correlation functions they compose vector $\boldsymbol{r}_y[k] = [\ldots, r_{y_i,y_j}[k], \ldots]^{\mathsf{T}}$, which can be written as:

$$\boldsymbol{r}_y[k] = \sum_{l=0}^{1} \boldsymbol{R}_c[l]\boldsymbol{r}_x[k-l], \tag{8}$$

where

$$\begin{aligned} \boldsymbol{R}_c[k] = [\boldsymbol{r}_{c_0,c_0}[k] \quad & \cdots \quad \boldsymbol{r}_{c_0,c_{M-1}}[k] \\ & \boldsymbol{r}_{c_1,c_1}[k] \quad \cdots \quad \boldsymbol{r}_{c_{M-1},c_{M-1}}[k]]^{\mathsf{T}}, \end{aligned} \tag{9}$$

From this equation, and after some algebraic manipulations, they arrive to the following matrix equation:

$$\boldsymbol{r}_{\mathrm{y}} = \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}, \tag{10}$$

where $\boldsymbol{r}_{\mathrm{y}} \in \mathbb{C}^{\frac{1}{2}M(2L+1)(M+1)\times 1}$ and $\boldsymbol{r}_{\mathrm{x}} \in \mathbb{C}^{N(2L+1)\times 1}$ are given by

$$\boldsymbol{r}_{\mathrm{y}} = \big[\boldsymbol{r}_y^{\mathsf{T}}[0] \quad \cdots \quad \boldsymbol{r}_y^{\mathsf{T}}[L] \quad \boldsymbol{r}_y^{\mathsf{T}}[-L] \quad \cdots \quad \boldsymbol{r}_y^{\mathsf{T}}[-1]\big]^{\mathsf{T}}, \tag{11a}$$

$$\boldsymbol{r}_{\mathrm{x}} = \big[\boldsymbol{r}_x^{\mathsf{T}}[0] \quad \cdots \quad \boldsymbol{r}_x^{\mathsf{T}}[L] \quad \boldsymbol{r}_x^{\mathsf{T}}[-L] \quad \cdots \quad \boldsymbol{r}_x^{\mathsf{T}}[-1]\big]^{\mathsf{T}}, \tag{11b}$$

with L being a design parameter related to the support of $\boldsymbol{r}_x[k]$ and $\boldsymbol{R}_{\mathrm{c}} \in \mathbb{C}^{\frac{1}{2}M(2L+1)(M+1)\times N(2L+1)}$ is given by

$$\boldsymbol{R}_{\mathrm{c}} = \begin{bmatrix} \boldsymbol{R}_c[0] & \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{R}_c[1] \\ \boldsymbol{R}_c[1] & \boldsymbol{R}_c[0] & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R}_c[1] & \boldsymbol{R}_c[0] & \boldsymbol{O} & \vdots \\ \vdots & \ddots & \ddots & \ddots & \boldsymbol{O} \\ \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{R}_c[1] & \boldsymbol{R}_c[0] \end{bmatrix}, \tag{12}$$

with $\boldsymbol{O}$ being a dimension-corresponding zero matrix.

The power spectrum of $x[n]$ can be writtten as $\boldsymbol{s}_{\mathrm{x}} = \mathbf{F}_{2L+1}\boldsymbol{r}_x$, being $\boldsymbol{F}_n \in \mathbb{C}^{n\times n}$ the DFT matrix of size n. From (10), a time domain (TD) estimator for the power spectrum is proposed as [2]

$$\hat{\boldsymbol{s}}_{\mathrm{x}} = \boldsymbol{F}_{(2L+1)N}\left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\boldsymbol{R}_{\mathrm{c}}\right)^{-1}\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\hat{\boldsymbol{r}}_{\mathrm{y}}, \tag{13}$$

where $(\cdot)^{-1}$ is the inverse, $(\cdot)^{\mathsf{H}}$ is the conjugate transpose, and $\hat{\boldsymbol{r}}_{\mathrm{y}} \in \mathbb{C}^{M^2(2L+1)\times 1}$ is an estimate of $\boldsymbol{r}_{\mathrm{y}}$. An unbiased estimator of $\boldsymbol{r}_{\mathrm{y}}$ can be obtained using that

$$\hat{r}_{y_i,y_j}[k] = \frac{1}{K-|k|}\sum_{l=\max(0,k)}^{K-1+\min(0,k)} y_i[l]y_j^*[l-k], \tag{14}$$

where K is the number of measurements.

From (13), it is clear that the selected sampling pattern has to lead to a full column rank matrix $\boldsymbol{R}_{\mathrm{c}}$. In [2], a suboptimal solution for the sampling patterns is proposed based on minimal sparse rulers (SR). In this paper we propose a WLS criterion for the design of the power spectrum estimator, and a MMSE criterion for the design of the sampling pattern, so that the full column rank constraint can be verified and at the same time the power spectrum is recovered with minimum mean square error.

III. Power Spectrum Recovery based on WLS

In a least squares sense, the estimate of $\boldsymbol{r}_{\mathrm{x}}$ in (10) can be written as

$$\begin{aligned} \hat{\boldsymbol{r}}_{\mathrm{x}} &= \arg\min_{\boldsymbol{r}_{\mathrm{x}}} \|\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\|_{\mathsf{E}}^2 \\ &= \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\boldsymbol{R}_{\mathrm{c}}\right)^{-1}\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\hat{\boldsymbol{r}}_{\mathrm{y}}, \end{aligned} \tag{15}$$

where $\|\cdot\|_{\mathsf{E}}$ is the Euclidean norm. We consider here an alternative approach based on WLS from

$$\begin{aligned} \hat{\boldsymbol{r}}_{\mathrm{x}} &= \arg\min_{\boldsymbol{r}_{\mathrm{x}}} \|\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\|_{\boldsymbol{W}}^2 \\ &= \arg\min_{\boldsymbol{r}_{\mathrm{x}}} \left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\right)^{\mathsf{H}}\boldsymbol{W}\left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\right), \end{aligned} \tag{16}$$

where $\boldsymbol{W} \in \mathbb{C}^{M^2(2L+1)\times M^2(2L+1)}$ is a positive-definite weighting matrix that does not depend on $\boldsymbol{r}_{\mathrm{x}}$. Let the cost function in (16) be

$$f_{\mathrm{WLS}}(\boldsymbol{r}_{\mathrm{x}}) = \left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\right)^{\mathsf{H}}\boldsymbol{W}\left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{R}_{\mathrm{c}}\boldsymbol{r}_{\mathrm{x}}\right). \tag{17}$$

Since $f_{\mathrm{WLS}}(\boldsymbol{r}_{\mathrm{x}})$ is a convex function of $\boldsymbol{r}_{\mathrm{x}}$, the critical point $\frac{\partial}{\partial \boldsymbol{r}_{\mathrm{x}}} f_{\mathrm{WLS}}(\boldsymbol{r}_{\mathrm{x}}) = \mathbf{0}$ leads to the minimum according to

$$\hat{\boldsymbol{r}}_{\mathrm{x}} = \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{R}_{\mathrm{c}}\right)^{-1}\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}\boldsymbol{W}\hat{\boldsymbol{r}}_{\mathrm{y}}, \tag{18}$$

which results in the new compressive power spectrum estimator that we propose:

$$\hat{\boldsymbol{s}}_{\mathrm{x}} = \boldsymbol{F}_{(2L+1)N} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \hat{\boldsymbol{r}}_{\mathrm{y}}. \tag{19}$$

Note that if $\boldsymbol{W}$ in (18) and (19) is the identity matrix, the estimates in (18) and (19) become (15) and (13) respectively. Next subsections are devoted to the computation of the optimal weighting matrix in a MMSE sense.

A. Error Covariance of Weighted Least Squares

The covariance of $\hat{\boldsymbol{r}}_{\mathrm{x}}$ in (18) is given by

$$\begin{aligned} \boldsymbol{C}_{\hat{r}_{\mathrm{x}}} &= \mathcal{E}_{\mathrm{x}} \left\{ \left(\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\right) \left(\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\right)^{\mathsf{H}} \right\} \\ &= \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \boldsymbol{W}^{\mathsf{H}} \boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W}^{\mathsf{H}} \boldsymbol{R}_{\mathrm{c}}\right)^{-1}, \end{aligned} \tag{20}$$

where $\boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \in \mathbb{C}^{M^2(2L+1)\times M^2(2L+1)}$ is given by

$$\boldsymbol{C}_{\hat{r}_{\mathrm{y}}} = \mathcal{E}_{\mathrm{y}} \left\{ \left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{r}_{\mathrm{y}}\right) \left(\hat{\boldsymbol{r}}_{\mathrm{y}} - \boldsymbol{r}_{\mathrm{y}}\right)^{\mathsf{H}} \right\}. \tag{21}$$

If the weighting matrix $\boldsymbol{W}$ is Hermitian, we have

$$\boldsymbol{C}_{\hat{r}_{\mathrm{x}}} = \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1}. \tag{22}$$

An optimal method to determine $\boldsymbol{W}$ can be derived from the minimization of the MSE in the estimate of the autocorrelation, that is, $\mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\}$. This can be written as

$$\begin{aligned} \hat{\boldsymbol{W}}_{\mathrm{MMSE}} &= \arg\min_{\boldsymbol{W}} \operatorname{tr}\left(\boldsymbol{C}_{\hat{r}_{\mathrm{x}}}\right) \\ &= \arg\min_{\boldsymbol{W}} \operatorname{tr}\Big(\left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \\ &\qquad \boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \Big). \end{aligned} \tag{23}$$

In fact, the minimal MSE (MMSE) weight in (23) is also optimal in MSE sense for the power spectrum recovery in (19) by considering that

$$\begin{aligned} \hat{\boldsymbol{W}}_{\mathrm{MMSE}} &= \arg\min_{\boldsymbol{W}} \mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\} \\ &= \arg\min_{\boldsymbol{W}} \mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{s}}_{\mathrm{x}} - \boldsymbol{s}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\}, \end{aligned} \tag{24}$$

where the second equality holds from the properties of $\boldsymbol{F}_{(2L+1)N}$ in (13) and (19), i.e. $\boldsymbol{F}_{(2L+1)N}^{\mathsf{T}} = \boldsymbol{F}_{(2L+1)N}$ and $\boldsymbol{F}_{(2L+1)N}^{-1} = \frac{1}{(2L+1)N} \boldsymbol{F}_{(2L+1)N}^{*}$.

B. Minimum Mean Square Error Weight

After some algebraic manipulation, one can show that

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{W}^*} \mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\} &= 2\boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-2} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \\ &- 2\boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-1} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{C}_{\hat{r}_{\mathrm{y}}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}} \left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \boldsymbol{W} \boldsymbol{R}_{\mathrm{c}}\right)^{-2} \boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}}. \end{aligned} \tag{25}$$

Proof of this derivation can be found in [14].

If we choose $\boldsymbol{W} = \boldsymbol{C}_{\hat{r}_{\mathrm{y}}}^{-1}$, we have

$$\left. \frac{\partial}{\partial \boldsymbol{W}^*} \mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{r}}_{\mathrm{x}} - \boldsymbol{r}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\} \right|_{\boldsymbol{W} = \boldsymbol{C}_{\hat{r}_{\mathrm{y}}}^{-1}} = \boldsymbol{O}. \tag{26}$$

Therefore, the MMSE weighting matrix in (23) becomes

$$\hat{\boldsymbol{W}}_{\mathrm{MMSE}} = \boldsymbol{C}_{\hat{r}_{\mathrm{y}}}^{-1}. \tag{27}$$

Let the second-order statistics of $x[n]$ be

$$\mathcal{E}_{\mathrm{x}} \left\{ x[n] x^*[m] \right\} = \sigma_{\mathrm{x}}^2 \delta[n-m], \qquad \forall m, n, \tag{28a}$$

$$\mathcal{E}_{\mathrm{x}} \left\{ x[n] x[m] \right\} = 0, \qquad \forall m, n, \tag{28b}$$

where σ_{x}^2 is the variance of $x[n]$. Under the signal assumption in (28), one can show that [14]

$$\boldsymbol{C}_{\hat{r}_{\mathrm{y}}} = \sigma_{\mathrm{x}}^4 \left(\boldsymbol{\Lambda}(\boldsymbol{\beta}) \otimes \boldsymbol{I}_{M^2} \right), \tag{29}$$

where $\boldsymbol{\Lambda}(\cdot)$ is a diagonal matrix whose diagonal is $\boldsymbol{\beta} \in \mathbb{R}^{(2L+1)\times 1}$, given by

$$\boldsymbol{\beta} = \begin{bmatrix} \frac{1}{K} & \frac{1}{K-1} & \cdots & \frac{1}{K-L} & \frac{1}{K-L} & \cdots & \frac{1}{K-1} \end{bmatrix}^{\mathsf{T}}, \tag{30}$$

$\otimes$ is the Kronecker product, and $\boldsymbol{I}_m$ is the identity of size m.

Although this signal assumption leads to a power spectrum which is flat, and obviously does not need to be estimated, it provides a way to obtain an analytical expression of the optimal weighting matrix for the worst case scenario, when the power spectrum has to be estimated under a low SNR. In addition, this assumption paves the way for the design of an alternative multicoset sampling pattern based on the minimization of the mean square error in the power spectrum estimate, as proposed in the next section of this paper.

IV. Minimum Mean Square Error Pattern Design

Let $\boldsymbol{n} \in \mathbb{N}^{M\times 1}$ be the vector of indices given by

$$\boldsymbol{n} = \begin{bmatrix} n_0 & n_1 & \cdots & n_{M-1} \end{bmatrix}^{\mathsf{T}}. \tag{31}$$

These indices correspond to the positions of the Nyquist rate samples to be obtained when acquiring the signal, that is, the indices of $c_i[n]$ which contain a nonzero element. Our goal is to find vector $\boldsymbol{n}$ that leads to a minimum MSE estimation of the power spectrum of the acquired signal. To this aim, we consider the WLS power spectrum estimator in (19), which obviously depends on the cross correlations between vectors $c_i[n]$ to be designed, and define the cost function to be minimized as $\mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{s}}_{\mathrm{x}} - \boldsymbol{s}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\}$. After some algebraic manipulations one can show that

$$\begin{aligned} \mathcal{E}_{\mathrm{x}} \left\{ \|\hat{\boldsymbol{s}}_{\mathrm{x}} - \boldsymbol{s}_{\mathrm{x}}\|_{\mathsf{E}}^2 \right\} &= \frac{1}{K(2L+1) - L(L+1)} \sigma_{\mathrm{x}}^4 (2L+1) N \\ &\qquad \sum_{n_r=1}^{N} \frac{1}{\alpha_{n_r}(\boldsymbol{n})}, \end{aligned} \tag{32}$$

where $\alpha_{n_r}(\boldsymbol{n}) \in \mathbb{Z}^{1\times 1}$ is given by

$$\begin{aligned} \alpha_{n_r}(\boldsymbol{n}) = \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} & \delta[-n_r + 1 - (n_{m_2} - n_{m_1})] \\ & + \delta[N - n_r + 1 - (n_{m_2} - n_{m_1})]. \end{aligned} \tag{33}$$

In order to recover the power spectrum in (19) the existence condition of the inverse $\left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \hat{\boldsymbol{W}}_{\mathrm{MMSE}} \boldsymbol{R}_{\mathrm{c}}\right)^{-1}$ must be verified.

Let $\mathcal{I}(\boldsymbol{X})$ be a set of the invertibility of a matrix $\boldsymbol{X}$, defined as

$$\mathcal{I}(\boldsymbol{X}) = \left\{\boldsymbol{Y} | \boldsymbol{Y} = \boldsymbol{X}^{-1}, |\boldsymbol{X}| \neq 0\right\}, \tag{34}$$

where $|\cdot|$ is the determinant of matrix $\cdot$.

Proposition 1 (Invertibility Condition): When K is larger than L, there exists a loosely sufficient condition that enables perfect power spectrum reconstruction, such as

$$\begin{aligned} &\alpha_{n_r}(\boldsymbol{n}) \geq 1, \quad \forall n_r \in \{2, 3, \ldots, N\} \\ &\qquad \Longleftrightarrow \mathcal{I}\left(\boldsymbol{R}_{\rm c}^{\sf H} \hat{\boldsymbol{W}}_{\rm MMSE} \boldsymbol{R}_{\rm c}\right) \neq \emptyset. \end{aligned} \tag{35}$$

Proof: The derivation of (35) is given in Appendix A. ■

From (32) and (35), the optimization problem for the MMSE pattern can be written as

$$\begin{aligned} \hat{\boldsymbol{n}}_{\rm WMMSE} &= \arg\min_{\boldsymbol{n}} \sum_{n_r=1}^{N} \frac{1}{\alpha_{n_r}(\boldsymbol{n})} \\ &\alpha_{n_r}(\boldsymbol{n}) \geq 1, \quad \forall n_r \in \{2, 3, \ldots, N\}, \\ &n_0 = 0, \\ \text{s.t.} \quad &n_{M-1} = \lfloor \tfrac{1}{2} N \rfloor, \\ &n_m \in \{n_{m-1}+1, \ldots, N_{\max} - M + m + 1\}, \\ &\forall m \in \{1, 2, \ldots, M-2\}. \end{aligned} \tag{36}$$

This optimization problem can be solved by exhaustive search. We call WLS-MMSE pattern a vector of indices which is a solution of this optimization problem. Next section show some examples of different WLS-MMSE patterns for different values of (M, N). There are few patterns after the search that correspond to SR patterns, that is, for some values of (M, N) the above optimization problem leads to a SR pattern.

V. Numerical Examples

This section presents some numerical results that show the theoretical performance of the designed sampling patterns and the proposed WLS-TD approach for power spectrum reconstruction. Thus, we compare here the theoretical MSE in terms of

$$\epsilon(\boldsymbol{n}) = \frac{1}{\sigma_{\rm x}^4} \mathcal{E}_{\rm x}\left\{\|\hat{\boldsymbol{s}}_{\rm x}(\boldsymbol{n}) - \boldsymbol{s}_{\rm x}\|_{\sf E}^2\right\}, \tag{37}$$

which has been computed for the following approaches:

1) A minimal sparse ruler as sampling pattern, denoted as $\boldsymbol{n}_{\rm SR}$, and the TD power spectrum estimator based on LS in (13) defined in [2].
2) An MMSE sampling pattern, denoted as $\hat{\boldsymbol{n}}_{\rm MMSE}$, obtained in [14] and the TD power spectrum estimator in (13).
3) An MMSE sampling pattern, denoted as $\hat{\boldsymbol{n}}_{\rm WMMSE}$, and the WLS-TD power spectrum estimator proposed in this paper.

Thus, Tables I, II and III show examples of the sampling patterns designed minimizing the cost function in (36), for different parameters, in addition to the sparse rulers of the same length and the MMSE patterns designed in [14].

Sampling Patterns for $N = 39$	
$\boldsymbol{n}_{\rm SR}$	$[0\ \ 1\ \ 10\ \ 11\ \ 13\ \ 15\ \ 17\ \ 19]^{\sf T}$
$\hat{\boldsymbol{n}}_{\rm MMSE}^{\rm LS}$	$[0\ \ 1\ \ 2\ \ 5\ \ 10\ \ 13\ \ 17\ \ 19]^{\sf T}$
$\hat{\boldsymbol{n}}_{\rm MMSE}^{\rm WLS}$	$[0\ \ 1\ \ 3\ \ 7\ \ 9\ \ 14\ \ 18\ \ 19]^{\sf T}$.

TABLE I. Designed sampling patterns for $N = 39$, $M = 8$.

Sampling Patterns for $N = 78$	
$\boldsymbol{n}_{\rm SR}$	$[0\ \ 1\ \ 12\ \ 15\ \ 26\ \ 29\ \ 31\ \ 33\ \ 35\ \ 37\ \ 39]^{\sf T}$
$\hat{\boldsymbol{n}}_{\rm MMSE}^{LS}$	$[0\ \ 1\ \ 2\ \ 10\ \ 15\ \ 16\ \ 28\ \ 32\ \ 35\ \ 37\ \ 39]^{\sf T}$
$\hat{\boldsymbol{n}}_{\rm MMSE}^{\rm WLS}$	$[0\ \ 1\ \ 3\ \ 5\ \ 7\ \ 17\ \ 20\ \ 30\ \ 31\ \ 38\ \ 39]^{\sf T}$

TABLE II. Designed sampling patterns for $N = 78$, $M = 11$.

Figure 2 show the MSE obtained by the different approaches as a function of the compression ratio. It can be clearly observed that the strategy proposed in this paper outperforms the approaches presented in previous works.

Finally, Tables IV show the WLS-MMSE sampling patterns obtained by exhaustive search for different values of (M, N).

VI. Conclusions

We have presented a new power spectrum recovery or power spectral estimation method that is based on WLS and multicoset sampling patterns that minimize the MSE in the reconstruction. The weighting matrix is shown in a closed form and can reduce the MSE of the power spectrum recovery. We also have presented the sampling pattern that is derived from the optimal weight. Exhaustive search is used to obtain the new sampling pattern designs. Numerical results shows that the WLS provides a significant improvement of the MSE over the former approach using minimal sparse ruler criterion and the MMSE criterion with conventional least squares.

Appendix A
Proof of Proposition 1

Taking into account (**??**), the inverse of $\boldsymbol{R}_{\rm c}^{\sf H} \hat{\boldsymbol{W}}_{\rm MMSE} \boldsymbol{R}_{\rm c}$ can be written as

$$\left(\boldsymbol{R}_{\rm c}^{\sf H} \hat{\boldsymbol{W}}_{\rm MMSE} \boldsymbol{R}_{\rm c}\right)^{-1} = \sigma_{\rm x}^4 \left(\boldsymbol{R}_{\rm c}^{\sf H}\left(\boldsymbol{\Lambda}^{-1}(\boldsymbol{\beta}) \otimes \boldsymbol{I}_{M^2}\right) \boldsymbol{R}_{\rm c}\right)^{-1}. \tag{38}$$

Based on (38), the inverse of $\boldsymbol{R}_{\rm c}^{\sf H} \hat{\boldsymbol{W}}_{\rm MMSE} \boldsymbol{R}_{\rm c}$ exists, if and only if each diagonal block-matrix in (**??**) is invertible, i.e.

$$\begin{aligned} &\mathcal{I}\left(\boldsymbol{R}_{\rm c}^{\sf H} \hat{\boldsymbol{W}}_{\rm MMSE} \boldsymbol{R}_{\rm c}\right) \neq \emptyset \\ &\Longleftrightarrow \bigcap_{l=1}^{2L} \mathcal{I}\left(\frac{1}{\beta_l} \boldsymbol{\Lambda}\left(\bar{\boldsymbol{\alpha}}(\boldsymbol{n})\right) + \frac{1}{\beta_{l+1}} \boldsymbol{\Lambda}\left(\tilde{\boldsymbol{\alpha}}(\boldsymbol{n})\right)\right) \\ &\qquad \cap \mathcal{I}\left(\frac{1}{\beta_{2L+1}} \boldsymbol{\Lambda}\left(\bar{\boldsymbol{\alpha}}(\boldsymbol{n})\right) + \frac{1}{\beta_1} \boldsymbol{\Lambda}\left(\tilde{\boldsymbol{\alpha}}(\boldsymbol{n})\right)\right) \neq \emptyset. \end{aligned} \tag{39}$$

Usually, K is larger than L. Thus, β_l is positive for all $l \in \{1, 2, \ldots, 2L+1\}$. Since $\bar{\alpha}_n$ and $\tilde{\alpha}_n$ in (**??**) are non-negative for all $n \in \{1, 2, \ldots, N\}$, we can derive

$$\begin{aligned} &\mathcal{I}\left(\boldsymbol{\Lambda}\left(\bar{\boldsymbol{\alpha}}(\boldsymbol{n})\right) + \boldsymbol{\Lambda}\left(\tilde{\boldsymbol{\alpha}}(\boldsymbol{n})\right)\right) \neq \emptyset \\ &\Longleftrightarrow \mathcal{I}\left(\frac{1}{\beta_{l_1}} \boldsymbol{\Lambda}\left(\bar{\boldsymbol{\alpha}}(\boldsymbol{n})\right) + \frac{1}{\beta_{l_2}} \boldsymbol{\Lambda}\left(\tilde{\boldsymbol{\alpha}}(\boldsymbol{n})\right)\right) \neq \emptyset, \end{aligned} \tag{40}$$

Sampling Patterns for $N = 128$	
$\boldsymbol{n}_{\mathrm{SR}}$	[0 1 3 6 13 20 27 34 41 48 55 59 63 64]$^{\mathsf{T}}$
$\hat{\boldsymbol{n}}_{\mathrm{MMSE}}^{\mathrm{LS}}$	[0 1 2 6 8 20 29 38 47 50 53 60 63 64]$^{\mathsf{T}}$
$\hat{\boldsymbol{n}}_{\mathrm{MMSE}}^{\mathrm{WLS}}$	[0 1 3 9 13 22 30 39 46 50 57 61 62 64]$^{\mathsf{T}}$

TABLE III. DESIGNED SAMPLING PATTERNS FOR $N = 128$,AND $M = 14$,

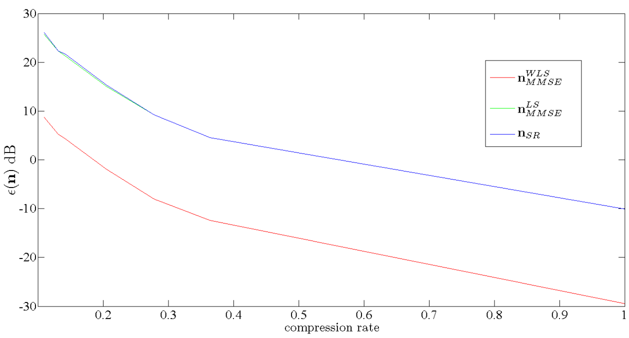

Fig. 2. Comparison of the MSE of the power spectrum estimate for different sampling patterns and estimators.

for $l_2 = l_1 + 1$, $l_1 \in \{1, 2, \ldots, 2L\}$, and $(l_1, l_2) = (2L+1, 1)$. Let $\boldsymbol{\alpha}(\boldsymbol{n}) \in \mathbb{R}^{N\times 1}$ be a vector, which is given by

$$\boldsymbol{\alpha}(\boldsymbol{n}) = [\alpha_1(\boldsymbol{n}) \quad \alpha_2(\boldsymbol{n}) \quad \cdots \quad \alpha_N(\boldsymbol{n})]^{\mathsf{T}}. \tag{41}$$

By noting that $\boldsymbol{\alpha}(\boldsymbol{n}) = \bar{\boldsymbol{\alpha}}(\boldsymbol{n}) + \tilde{\boldsymbol{\alpha}}(\boldsymbol{n})$, we have

$$\mathcal{I}\left(\boldsymbol{\Lambda}\left(\boldsymbol{\alpha}(\boldsymbol{n})\right)\right) \neq \emptyset \Longleftrightarrow \mathcal{I}\left(\boldsymbol{R}_{\mathrm{c}}^{\mathsf{H}} \hat{\boldsymbol{W}}_{\mathrm{MMSE}} \boldsymbol{R}_{\mathrm{c}}\right) \neq \emptyset. \tag{42}$$

(M, N)	WLS-MMSE sampling patterns
$(5, 18)$	[0 1 2 6 9]$^{\mathsf{T}}$ [0 1 4 7 9]$^{\mathsf{T}}$ [0 2 5 8 9]$^{\mathsf{T}}$ [0 3 7 8 9]$^{\mathsf{T}}$
$(8, 39)$	[0 1 2 5 10 13 17 19]$^{\mathsf{T}}$ [0 1 2 7 11 14 17 19]$^{\mathsf{T}}$ [0 1 3 7 9 14 18 19]$^{\mathsf{T}}$ [0 1 5 10 12 16 18 19]$^{\mathsf{T}}$ [0 2 5 8 12 17 18 19]$^{\mathsf{T}}$ [0 2 6 9 14 17 18 19]$^{\mathsf{T}}$
$(11, 78)$	[0 1 2 5 12 13 20 30 34 36 39]$^{\mathsf{T}}$ [0 1 2 6 7 16 24 27 35 37 39]$^{\mathsf{T}}$ [0 1 2 6 15 19 28 31 36 38 39]$^{\mathsf{T}}$ [0 1 2 6 17 18 26 29 32 36 39]$^{\mathsf{T}}$ [0 1 2 8 9 21 25 31 34 36 39]$^{\mathsf{T}}$ [0 1 2 9 13 14 18 29 33 36 39]$^{\mathsf{T}}$
(11,84)	[0 1 2 3 19 24 28 32 36 39 42]$^{\mathsf{T}}$ [0 3 6 10 14 18 23 39 40 41 42]$^{\mathsf{T}}$

TABLE IV. DESIGNED WLS MMSE PATTERNS FOR DIFFERENT VALUES OF (M, N).

The diagonal matrix $\boldsymbol{\Lambda}\left(\boldsymbol{\alpha}(\boldsymbol{n})\right)$ is invertible if and only if $\alpha_{n_r}(\boldsymbol{n})$ for all $n_r \in \{2, 3, \ldots, N\}$ is non-zero, or more strictly greater than or equal to one. Note that $\alpha_1(\boldsymbol{n})$ is not included into the condition in (35), because $\alpha_1(\boldsymbol{n}) = M$ for any $\boldsymbol{n}$.

REFERENCES

[1] G. Leus and D. D. Ariananda, "Power spectrum blind sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 443–446, Aug. 2011.

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

[3] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: Making software radios more personal," *IEEE Personal Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[4] F. C. Commission, "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," FCC 03-322, ET Docket No. 03-108, Dec.

[5] S. Haykin, D. J. Thompson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[6] D. D. Ariananda, M. K. Lakshmanan, and H. Nikookar, "A survey on spectrum sensing techniques for cognitive radio," in *Proc. 2nd Int. Works. Cognitive Radio and Advanced Spectrum Management (CogART 2009)*, Aalborg, Denmark, May 2009, pp. 74–79.

[7] 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 Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[8] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. 2009 (ICASSP 2009)*, Taipei, Taiwan, Apr. 2009, pp. 2337–2340.

[9] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, pp. 993–1009, Mar. 2009.

[10] M. E. Domínguez-Jiménez, N. González-Prelcic, G. Vazquez-Vilar, and R. López-Valcarce, "Design of universal multicoset sampling patterns for compressed sensing of multiband sparse signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. 2012 (ICASSP 2012)*, Kyoto, Japan, Mar. 2012, pp. 3337–3340.

[11] P. Pal and P. Vaidyanathan, "Coprime sampling and the music algorithm," in *IEEE Digit. Signal Process. Workshop*, 2011.

[12] P. Vaidyanathan and P. Pal, "Theory of sparse coprime sensing in multiple dimensions," *IEEE Trans. Signal Process.*, vol. 59, pp. 3592–3608, Aug. 2011.

[13] M. E. Domínguez-Jiménez and N. González-Prelcic, "A class of circular sparse rulers for compressive power spectrum estimation," in *Proc. of the European Signal Processing Conference (EUSIPCO) 2013*, Marrakech, Morocco, Sep. 2013.

[14] B. Tausiesakul and N. González-Prelcic, "Minimum mean square error multicoset sampling pattern for time-domain power spectrum blind sampling," Department of Signal Theory and Communications, University of Vigo, Vigo, Spain, Tech. Rep., Nov. 2013.