

Compressive Sampling-Based Multiple Symbol Differential Detection for UWB Communications

Shahzad Gishkori, *Student Member, IEEE*, Vincenzo Lottici, and Geert Leus, *Fellow, IEEE*

Abstract—Compressive sampling (CS) based multiple symbol differential detectors are proposed for impulse-radio ultra-wideband signaling, using the principles of generalized likelihood ratio tests. The CS based detectors correspond to two communication scenarios. One, where the signaling is fully synchronized at the receiver and the other, where there exists a symbol level synchronization only. With the help of CS, the sampling rates are reduced much below the Nyquist rate to save on the high power consumed by the analog-to-digital converters. In stark contrast to the usual compressive sampling practices, the proposed detectors work on the compressed samples directly, thereby avoiding a complicated reconstruction step and resulting in a reduction of the implementation complexity. To resolve the detection of multiple symbols, compressed sphere decoders are proposed as well, for both communication scenarios, which can further help to reduce the system complexity. Differential detection directly on the compressed symbols is generally marred by the requirement of an identical measurement process for every received symbol. Our proposed detectors are valid for scenarios where the measurement process is the same as well as where it is different for each received symbol.

Index Terms—Compressive sampling (CS), multiple symbol differential detection (MSDD), sphere decoding (SD), ultra-wideband impulse radio (UWB-IR).

I. INTRODUCTION

PROMISING the prospects of high data rates, fine time resolution, multipath immunity and coexistence with legacy services via frequency overlay, ultra-wideband (UWB) impulse radios (IRs) are deemed as strong candidates for short range connectivity, location-aware wireless sensor networks and low-rate communications with ranging capability [1], [2]. Owing to the ultra-large bandwidth, each transmitted pulse arrives at the receiver scattered over hundreds of separable paths with possible severe pulse distortion [3], [4]. Under these harsh propagation conditions, the rich diversity of UWB channels can be exploited by employing detection strategies based on Rake receivers, which however, require a large number of correlator-

based fingers combined with accurate channel estimation, thus resulting in an intensive computational load and a high power consumption [5], [6]. Such requirements are contrary to the UWB objectives that call for simple receiver processing units with moderate energy consumption. Therefore, efficient techniques are needed to overcome these impediments and facilitate a pervasive deployment of UWB-based networks.

Background and Prior Works: A number of viable yet sub-optimal receivers based on noncoherent detection have been proposed in the literature for efficient energy capture while avoiding channel estimation [7]. In the transmitted reference (TR) scheme [8], [10], an extra information-free reference pulse is used as a channel template by the correlator to detect the information data, thereby causing wastage of transmitted power and a decrease in data rate. These drawbacks can be avoided by adopting differential detection (DD) [9], [10]. Differentially encoding the information symbols allows employing the signal received within the previous symbol interval as a channel template for detection, thus enabling potentially low-complexity and energy-efficient receivers. However, the template waveform in both TR and DD schemes is neither noise-free nor interference-free, which contributes to a substantial performance degradation. This prompted the use of enhanced DD methods in the form of multiple symbol differential detection (MSDD) [11], [12]. Instead of correlating only the consecutive symbol-long received waveforms, a block of differentially encoded symbols is detected jointly, offering improved performance over both severe multipath fading and interference-limited scenarios. Still, accurate pulse level timing information has to be acquired, which in view of the low-power and ultra-short transmitted pulses, again requires a considerable computational effort; see e.g., [13]–[15]. Hence, a variant of the MSDD scheme has recently been proposed in [16] to reduce the timing restrictions, by limiting the timing accuracy from pulse or frame level to symbol level only, while maintaining a competitive performance.

Despite the considerable advantages offered by the symbol level synchronization (SLS) MSDD, the delay components required by the correlation units (on the order of tens or even hundreds of nanoseconds) lead to hardware implementation issues. Indeed, the long and accurate delay lines are hard to realize in the analog domain, and a digital implementation based on Nyquist rate (NR) sampling can heavily stress the receiver analog-to-digital converter (ADC), thereby causing a high power consumption [17]. To facilitate the ADC implementation, some attractive novel theories can be of effective help on reducing the sampling frequency below the corner-stone NR threshold, e.g., those based on sampling at the rate

Manuscript received May 26, 2013; revised November 25, 2013 and March 23, 2014; accepted April 3, 2014. Date of publication April 14, 2014; date of current version July 8, 2014. This work was supported in part by NWO-STW under the VICI program (Project 10382) and in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMMunications NEWCOM under Grant 318306. This work was presented in part at the IEEE International Conference on Ultra-Wideband (ICUWB), September 2012. The associate editor coordinating the review of this paper and approving it for publication was J. Wu.

S. Gishkori and G. Leus are with the Faculty of EEMCS, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: s.s.gishkori@tudelft.nl; g.j.t.leus@tudelft.nl).

V. Lottici is with the Department of Information Engineering, University of Pisa, 56126 Pisa, Italy (e-mail: vincenzo.lottici@iet.unipi.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2014.2317175

of innovation (SRI) [18], [19] or compressive sampling (CS) [20], [21]. Capitalizing on suitable properties of the signal, like the sparsity exhibited in the time domain by the UWB signals [3], [4], the key idea is to extract a reduced set of compressed samples from the analog received signal, or in other words, converting it into the compressed domain through a few measurements taken in the analog domain; see e.g., [22], [23]. Then, a reconstruction step from the compressed samples may follow by applying one of the algorithms proposed in [19]–[21], [24]–[26]. Alternatively, the reconstruction step is skipped and the receiver processing is based on the compressed samples directly.

The SRI technique is applied in [27], [28] to UWB receivers that work at sub-NR sampling but also require channel estimation (CE). On the other side, the CS framework supports a large variety of sampling kernels, e.g., random sampling, and hence allows for a higher flexibility [20], [21]. Practical applications of CS to the UWB scenario can be found in [29]–[33], mostly again for coherent receivers, thereby requiring CE. Apart from the overhead involved in the transmission of extra information such as pilot or training symbols in these works, one inevitably has to suffer from the complexity load required by the reconstruction of the channel template.

A simpler yet performance competitive implementation, consists of combining the CS framework with noncoherent detection, as illustrated in [34]–[37]. In [34], noncoherent receivers for differentially encoded UWB signals are designed exploiting the CS techniques. Besides introducing a joint reconstruction and detection scheme, a direct compressed DD (DC-DD) is also presented, which skips the reconstruction step, hence reducing the complexity. Building upon the DC-DD, the work in [35] merges the concepts of CS and decision feedback DD (DF-DD) [38]. A power-efficient and low-complexity receiver is enabled, named as CS based (sorted) DF-DD or csDF-DD in short, however it has to be emphasized that: *i*) its robustness to timing offsets is restricted to only a fraction of the symbol interval and, *ii*) the measurement matrix is required to be the same for all the symbols within each block.

Rationale of the Proposed Approach: The above facts indicate that CS-based noncoherent detection can lead to promising receiver schemes. Hence, the search for an effective way to reduce complexity while preserving performance, fully motivates the current paper to make a further contribution. The basic idea we pursue, in part traced back to [36], is threefold: *i*) instead of considering the DC-DD of a single information symbol as in [34], we cast the concept of MSDD into the CS framework, thus formalizing the CS-based MSDD (CMSDD) scheme at sub-NR sampling; *ii*) to relax the demanding prerequisite of sub-pulse level accuracy on the timing synchronization, we develop a modified version of the CMSDD which requires SLS only, in the sequel referred to as SLS-CMSDD; *iii*) aimed at skipping CE, we resort to the generalized likelihood ratio test (GLRT) principle [39] in line with [12] and [16], according to which the generalized log-likelihood metric (GLLM) is maximized not only over the information symbols but also over the unknown channel template. GLRT also helps alleviate the restrictions of the measurement matrices to be the same for all symbols.

Contributions: The main features of our approach are detailed as follows.

- 1) The proposed MSDD-like schemes are derived by avoiding the reconstruction step, i.e., they work directly on the compressed signal samples. The result is that the sampling rate as well as the implementation complexity related to the evaluation of the correlation coefficients needed by the objective function, are both kept at affordable levels, in accordance with the UWB requirements.
- 2) Unlike the CS-based noncoherent receivers illustrated so far, the measurement process can be either the same or different from symbol to symbol, thus offering an additional degree of freedom that can help the receiver better adapt to various scenarios.
- 3) As briefly touched above, resorting to the SLS concept, the robustness to timing errors of the proposed CS-based schemes is brought from pulse or frame level to symbol level. This feature relaxes the performance of the timing synchronizer, so further lowering the overall receiver complexity.
- 4) A particular effort is put on cutting back the complexity required to optimize the objective function over each data block for both the ideally-synchronized CMSDD and the SLS-CMSDD, which grows exponentially in the block size.¹ To this end, a modified sphere decoding (SD) algorithm is derived enabling the joint detection of blocks of tens of symbols at polynomial complexity.
- 5) Comprehensive numerical simulation results obtained over realistic UWB scenarios corroborate our analytical findings and demonstrate that the proposed noncoherent detectors can deliver efficient performance-versus-complexity trade-offs, and are capable of jointly relaxing the stringent requirements of both the high sampling rate and the accurate timing synchronization.

Organization: The rest of the paper is organized as follows. Section II describes the signal model. After reviewing the MSDD scheme with ideal timing synchronization, Section III introduces the CS-based version. Section IV extends the SLS variant of the MSDD to the CS framework, and Section V deals with a modified scheme of SD. The simulation results are illustrated in Section VI, and finally, in Section VII some concluding remarks are drawn.

Notations: Matrices are in upper case bold while column vectors are in lower case bold, $[\mathbf{a}]_i$ is the i th entry of the vector \mathbf{a} , \mathbf{I}_N is the identity matrix of size $N \times N$, $\mathbf{1}_{M \times N}$ is the $M \times N$ matrix with all components one, $\mathbf{0}_{M \times N}$ is the $M \times N$ matrix with all components zero, $(\cdot)^T$ denotes transpose, $(\cdot)^{-1}$ denotes inverse, \otimes stands for the Kronecker product, \star describes the convolution, $\text{diag}\{\cdot\}$ gives a block diagonal matrix having the arguments along its main diagonal, $\hat{\mathbf{a}}$ is the estimate of \mathbf{a} , $\lfloor \cdot \rfloor$ denotes the floor function, \triangleq defines an entity, the ℓ_p norm of a vector \mathbf{a} is denoted as $\|\mathbf{a}\|_p = (\sum_{i=0}^{N-1} |[\mathbf{a}]_i|^p)^{1/p}$, and $E\{\cdot\}$ denotes statistical expectation.

¹We recall from [12] that the block size plays a role in determining the performance improvements against the DD scheme, in the sense that the longer the block the better the performance.

II. SIGNAL MODEL

For the UWB-IR signal model, each symbol is represented by N_f frames with one pulse $q(t)$ per frame. The symbol, frame and pulse intervals are designated as T_s , T_f , and T_q , respectively, satisfying $T_s = N_f T_f$, $T_q \ll T_f$. Denoting the symbol level waveform² as

$$s(t) \triangleq \sum_{j=0}^{N_f-1} q(t - jT_f) \quad (1)$$

the transmitted signal corresponding to a block of $Q + 1$ consecutive symbols can be written as

$$u(t) = \sum_{k=0}^Q b_k s(t - kT_s) \quad (2)$$

where $b_k \in \{\pm 1\}$ are the transmitted symbols, which are differentially encoded according to the rule

$$b_k = b_{k-1} a_k \quad (3)$$

with $a_k \in \{\pm 1\}$ representing the information-bearing symbols. Without loss of generality, we consider $b_0 = 1$ as initial reference symbol.

The multipath channel is assumed to be time-invariant within an interval of length $(Q + 1)T_s$, which is required to transmit (2). The delay spread is smaller than T_f such that the overall channel fits within a single frame and hence inter-symbol interference (ISI) is avoided. Under the assumption that the channel impulse response (CIR) has L paths, the received pulse is given by

$$h(t) \triangleq \sum_{\ell=0}^{L-1} \alpha_\ell q(t - \tau_{\ell,0}) \star h_{LP}(t) \quad (4)$$

where $h_{LP}(t)$ is the low-pass filter at the receiver with bandwidth W , $\tau_{\ell,0} \triangleq \tau_\ell - \tau$, $0 \leq \ell \leq L - 1$, is the relative delay of the ℓ th path with respect to the timing offset $\tau \triangleq \tau_0$ of the first path due to signal propagation, τ_ℓ is the actual delay of the ℓ th path at the receiver and α_ℓ is the respective path gain. The symbol level received waveform can thus be expressed as

$$g(t) \triangleq \sum_{j=0}^{N_f-1} h(t - jT_f) \quad (5)$$

and correspondingly, after exploiting (2) and (4), (5), the received signal $r(t)$ is given by

$$r(t) = \underbrace{\sum_{k=0}^Q b_k g(t - kT_s - \tau)}_{\triangleq x(t)} + v(t) \quad (6)$$

²Our focus is on a single-user point-to-point link, so for simplicity of presentation, the time hopping (TH) code is not employed. Such an extension is easy to be included. However, frame averaging may not be possible in this case. Our model can also be extended to the multi-user scenario but it would necessitate a compressed user template to identify a specific user.

where $x(t)$ is the block level received signal and $v(t)$ is the zero-mean additive white Gaussian noise component with variance σ_v^2 .

III. MSDD WITH EXACT TIMING SYNCHRONIZATION

In this section, we consider the MSDD scheme when exact timing information is available at the receiver, or equivalently, when the timing offset is $\tau = 0$, and accordingly $\tau_{\ell,0} = \tau_\ell$. As a first step, we revisit the MSDD scheme presented in [12] for NR sampled UWB signals and derive it in an algebraic form (which is needed to build mathematical foundations for the compressed version), and denote it for simplicity as NMSDD. Then, we propose the MSDD based on the CS framework, referred to as CMSDD.

A. Nyquist-Rate MSDD

Denoting with $1/T \triangleq N/T_f$ the Nyquist sampling rate, the NR received signal (6) can be expressed as

$$\mathbf{r} \triangleq [\mathbf{r}_0^T, \mathbf{r}_1^T, \dots, \mathbf{r}_Q^T]^T \quad (7)$$

where $\mathbf{r}_k \triangleq [\mathbf{r}_k^{(0)T}, \mathbf{r}_k^{(1)T}, \dots, \mathbf{r}_k^{(N_f-1)T}]^T$, with

$$\mathbf{r}_k^{(j)} \triangleq [r(kT_s + jT_f), r(kT_s + jT_f + T), \dots, r(kT_s + jT_f + NT - T)]^T \quad (8)$$

collecting the N NR samples of the j th frame for the k th symbol. Similarly, we can define \mathbf{x} , \mathbf{x}_k and $\mathbf{x}_k^{(j)}$ based on $x(t)$, and \mathbf{v} , \mathbf{v}_k and $\mathbf{v}_k^{(j)}$ based on $v(t)$. From (6), we can then obtain that

$$\mathbf{r}_k = \mathbf{x}_k + \mathbf{v}_k, \quad 0 \leq k \leq Q \quad (9)$$

where $\mathbf{x}_k \triangleq b_k(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})$ is the signal part of \mathbf{r}_k , with

$$\mathbf{h} \triangleq [h(0), h(T), \dots, h(NT - T)]^T \quad (10)$$

made up of the NR samples of the received pulse waveform (4). Note that \mathbf{v}_k is a zero-mean Gaussian distributed noise vector with covariance matrix $\mathbf{C}_v \triangleq \mathbb{E}\{\mathbf{v}_k \mathbf{v}_k^T\} = \sigma_v^2 \mathbf{I}_{NN_f}$. Exploiting (7) and (9), the joint model for the block of $Q + 1$ symbols can now be written as

$$\mathbf{r} = (\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) + \mathbf{v} \quad (11)$$

where $\mathbf{b} \triangleq [b_0, b_1, \dots, b_Q]^T$ denotes the transmitted symbols. Hence, after defining the vector of the information symbols as $\mathbf{a} \triangleq [a_1, a_2, \dots, a_Q]^T$, the NMSDD scheme can be stated as follows.

Proposition 1 NMSDD: The GLRT NMSDD mixed-integer optimization problem (OP) is

$$\hat{\mathbf{a}}^{(\text{NMSDD})} = \arg \max_{\mathbf{a}} \left\{ \max_{\mathbf{h}} \Lambda(\mathbf{r} | \mathbf{a}, \mathbf{h}) \right\} \quad (12)$$

where the GLLM is

$$\Lambda(\mathbf{r}|\mathbf{a}, \mathbf{h}) \triangleq 2N_f \bar{\mathbf{r}}^T (\mathbf{b} \otimes \mathbf{I}_N) \mathbf{h} - (Q+1)N_f \mathbf{h}^T \mathbf{h} \quad (13)$$

with $\bar{\mathbf{r}} \triangleq [\bar{\mathbf{r}}_0^T, \bar{\mathbf{r}}_1^T, \dots, \bar{\mathbf{r}}_Q^T]^T$ and

$$\bar{\mathbf{r}}_k \triangleq \frac{1}{N_f} \sum_{j=0}^{N_f-1} \mathbf{r}_k^{(j)} \quad (14)$$

which represents the $N \times 1$ vector collecting the samples of the average frame for the k th symbol.

Proof: Under the joint NR sampled model (11), the GLLM can be written as

$$\begin{aligned} \Lambda(\mathbf{r}|\mathbf{a}, \mathbf{h}) &\triangleq 2\mathbf{r}^T (\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \\ &\quad - [(\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})]^T \\ &\quad \times [(\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})] \\ &= 2\mathbf{r}^T (\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \\ &\quad - (Q+1) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})^T (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \end{aligned} \quad (15)$$

which can be further simplified into (13). Since \mathbf{b} is a function of \mathbf{a} as described in (3), (12) can be solved into two steps according to the GLRT principle. First, the GLLM (13) is maximized over \mathbf{h} by setting the corresponding gradient to zero, and then, it is optimized over \mathbf{a} . ■

B. Compressive Sampling MSDD

For the CMSDD, we assume that each received frame vector $\mathbf{r}_k^{(j)}$ given by (8) is compressed using the $M \times N$ frame level fat measurement matrix Φ_k (i.e., $M < N$), such that $\Phi_k \Phi_k^T = \mathbf{I}_M$

$$\mathbf{y}_k^{(j)} \triangleq \Phi_k \mathbf{r}_k^{(j)}, \quad 0 \leq j \leq N_f - 1. \quad (16)$$

Note that the compression ratio $\mu \triangleq M/N$, with $0 < \mu \leq 1$, identifies how much one can economize the sampling rate, and accordingly, the computational load of the data detector.

Upon defining $\mathbf{y}_k \triangleq [\mathbf{y}_k^{(0)T}, \mathbf{y}_k^{(1)T}, \dots, \mathbf{y}_k^{(N_f-1)T}]^T$, the compressed received signal within the k th symbol can then be expressed by the $MN_f \times 1$ vector

$$\mathbf{y}_k = (\mathbf{I}_{N_f} \otimes \Phi_k) \mathbf{r}_k = (\mathbf{I}_{N_f} \otimes \Phi_k) \mathbf{x}_k + \boldsymbol{\xi}_k, \quad 0 \leq k \leq Q \quad (17)$$

where $\boldsymbol{\xi}_k \triangleq (\mathbf{I}_{N_f} \otimes \Phi_k) \mathbf{v}_k$ is the noise component with covariance matrix $\mathbf{C}_{\boldsymbol{\xi}} \triangleq E\{\boldsymbol{\xi}_k \boldsymbol{\xi}_k^T\} = \sigma_v^2 \mathbf{I}_{MN_f}$. It should be noted that the measurement process in (16) is performed in the compressed analog domain; see [22], [23] for details about possible analog implementations.

Now from (17), we can express the joint compressed model for the $Q+1$ symbols as

$$\mathbf{y} = \Psi (\mathbf{b} \otimes \mathbf{I}_{NN_f}) (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) + \boldsymbol{\xi} \quad (18)$$

where $\mathbf{y} \triangleq [\mathbf{y}_0^T, \mathbf{y}_1^T, \dots, \mathbf{y}_Q^T]^T$ and $\boldsymbol{\xi} \triangleq [\boldsymbol{\xi}_0^T, \boldsymbol{\xi}_1^T, \dots, \boldsymbol{\xi}_Q^T]^T$ are the compressed $(M < N)(Q+1)MN_f \times 1$ measurement and noise vectors, respectively, and

$$\Psi \triangleq \text{diag} \{ \mathbf{I}_{N_f} \otimes \Phi_0, \mathbf{I}_{N_f} \otimes \Phi_1, \dots, \mathbf{I}_{N_f} \otimes \Phi_Q \} \quad (19)$$

is the $(Q+1)MN_f \times (Q+1)NN_f$ block level measurement matrix, such that $\Psi \Psi^T = \mathbf{I}_{(Q+1)MN_f}$. Hence, the CMSDD can be formulated as follows.

Proposition 2 CMSDD: The GLRT CMSDD integer OP is

$$\hat{\mathbf{a}}^{(\text{CMSDD})} = \arg \max_{\mathbf{a}} \{ \Delta(\mathbf{y}|\mathbf{a}) \} \quad (20)$$

where the objective function is

$$\Delta(\mathbf{y}|\mathbf{a}) = \sum_{k=0}^Q \sum_{\ell=0}^Q b_k b_{\ell} \bar{\mathbf{y}}_k^T \Phi_k \Phi_{\ell}^T \bar{\mathbf{y}}_{\ell} \quad (21)$$

with

$$\bar{\mathbf{y}}_k \triangleq \frac{1}{N_f} \sum_{j=0}^{N_f-1} \mathbf{y}_k^{(j)} \quad (22)$$

being the $M \times 1$ vector collecting the samples of the average compressed frame for the k th symbol.

Proof: See Appendix A. ■

A number of remarks about the CMSDD can now be highlighted.

- 1) If the frame level measurement matrices Φ_k are all orthogonal to each other, i.e., $\Phi_k \Phi_{\ell}^T = \mathbf{0}_{M \times M}$, $\forall k, \ell$ with $0 \leq k, \ell \leq Q$, then $\Delta(\mathbf{y}|\mathbf{a})$ does not depend on \mathbf{a} , and accordingly the detector does not exist.
- 2) If the frame level measurement matrices Φ_k are all the same for each symbol, i.e., $\Phi_0 = \Phi_1 = \dots = \Phi_Q$, then taking into account (3), $\Delta(\mathbf{y}|\mathbf{a})$ turns into

$$\Delta(\mathbf{y}|\mathbf{a}) = \sum_{k=1}^Q \sum_{\ell=0}^{k-1} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{i+\ell} \bar{\mathbf{y}}_k^T \bar{\mathbf{y}}_{\ell} \quad (23)$$

whereas in the case they differ from symbol to symbol, $\Delta(\mathbf{y}|\mathbf{a})$ has the general form

$$\Delta(\mathbf{y}|\mathbf{a}) = \sum_{k=1}^Q \sum_{\ell=0}^{k-1} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{i+\ell} \bar{\mathbf{y}}_k^T \Phi_k \Phi_{\ell}^T \bar{\mathbf{y}}_{\ell}. \quad (24)$$

- 3) By virtue of the CS framework, the CMSDD relies on the evaluation of the average frame in (22), which is performed for each symbol in the compressed domain. This is less demanding than the implementation of (14) based on the NR sampling. As an additional strength, the detection process of the CMSDD avoids a reconstruction step, which further helps in keeping the complexity at an affordable level.
- 4) Concerning the performance limits of the CMSDD, if the frame level measurement matrices are orthogonal to each other, then the CMSDD does not work, whereas better performance is expected if they are the same for each

symbol. However, for applications where choosing identical measurement matrices is not feasible, the CMSDD can still offer compressed detection.

- 5) The performance-versus-complexity trade-off enabled by the CMSDD is expected to be governed by the compression ratio μ as well. Indeed, the higher the μ , the lower the performance loss, till the performance approaches that of the NMSDD as $\mu \rightarrow 1$. This can be established mathematically by noting that when $\mu = 1$ (i.e., $M = N$) then $\Phi_k^T \Phi_k = \mathbf{I}_N$ (which is a general property of orthogonal matrices). Thus,

$$\bar{\mathbf{y}}_k^T \Phi_k \Phi_\ell^T \bar{\mathbf{y}}_\ell = (\Phi_k \bar{\mathbf{r}}_k)^T \Phi_k \Phi_\ell^T (\Phi_\ell \bar{\mathbf{r}}_\ell) = \bar{\mathbf{r}}_k^T \bar{\mathbf{r}}_\ell$$

and the CMSDD in (24) reduces to the NMSDD.

IV. MSDD WITH SYMBOL LEVEL SYNCHRONIZATION

In Section III, we assumed ideal timing synchronization. This assumption means that the receiver can recover an accurate estimate of the timing offset at the pulse level. In this section, we will relax this computationally demanding constraint: first, we re-describe in algebraic form the MSDD scheme with synchronization at symbol level as proposed in [16] using NR sampling, denoted as the SLS-NMSDD in short. Then, we extend the above CMSDD approach to symbol level synchronization, thus formulating the SLS-CMSDD scheme. A coarse symbol level synchronization is thought to be available, so that the timing offset τ is less than a symbol duration, i.e., $\tau \in [0, T_s)$. Furthermore, the observation window is increased to $Q + 1$ symbols to accommodate the residual (unknown) timing offset.

The key idea of the MSDD with SLS is to partition the received symbol waveform $g(t)$ given by (5) into the two parts $g_0(t)$ and $g_1(t)$, such that

$$g_0(t) \triangleq \begin{cases} 0 & t \in [0, \tau) \\ g(t - \tau) & t \in [\tau, T_s) \end{cases} \quad (25)$$

$$g_1(t) \triangleq \begin{cases} g(t + T_s - \tau) & t \in [0, \tau) \\ 0 & t \in [\tau, T_s) \end{cases} \quad (26)$$

as depicted in Fig. 1, for a single frame per symbol, i.e., $N_f = 1$. It is apparent from (25) and (26) that $g_0(t)$ and $g_1(t)$ depend upon τ and are orthogonal to each other.

A. Nyquist-Rate MSDD With Symbol Level Synchronization

Denoting, $N_\tau \triangleq \lfloor \tau/T \rfloor$ and $\varepsilon \triangleq (\tau - N_\tau T)$, with $\varepsilon \in [0, T)$, the NR sampled symbol level versions of $g_0(t)$ and $g_1(t)$ are given by

$$\mathbf{g}_0 \triangleq [0_{N_\tau \times 1}^T, g(-\varepsilon), g(T - \varepsilon), \dots, g(NN_f T - N_\tau T - T - \varepsilon)]^T \quad (27)$$

$$\mathbf{g}_1 \triangleq [g(NN_f T - N_\tau T - \varepsilon), g(NN_f T - N_\tau T + T - \varepsilon), \dots, g(NN_f T - T - \varepsilon), \mathbf{0}_{(NN_f - N_\tau) \times 1}^T]^T \quad (28)$$

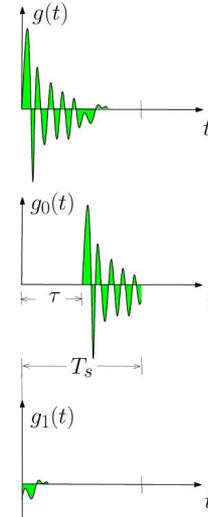


Fig. 1. Partitioning of $g(t)$ into $g_0(t)$ and $g_1(t)$ for $N_f = 1$, in the presence of a timing offset τ .

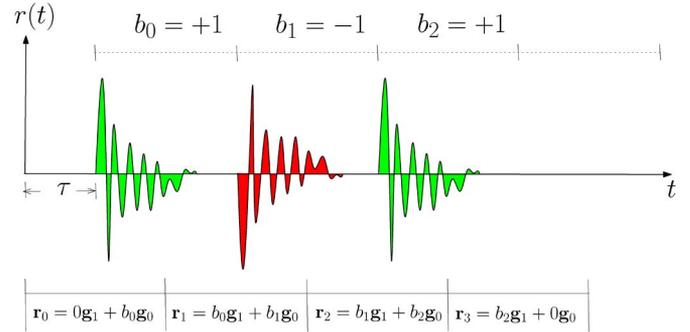


Fig. 2. SLS model in the noiseless case with $Q = 2$, $N_f = 1$ and timing offset τ .

Thus, the NR sampled version of the k th received symbol waveform can be represented by the $NN_f \times 1$ vector

$$\mathbf{r}_k = b_k \mathbf{g}_0 + b_{k-1} \mathbf{g}_1 + \mathbf{v}_k, \quad 0 \leq k \leq Q + 1 \quad (29)$$

where without loss of generality we assume $b_{-1} = b_{Q+1} = 0$. In view of (29), the joint SLS NR sampled model for the block of $Q + 2$ symbols can be put into the form

$$\hat{\mathbf{r}} = (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 + \hat{\mathbf{v}} \quad (30)$$

where $\mathbf{b}_0 \triangleq [b_0, b_1, \dots, b_Q, b_{Q+1}]^T$ and $\mathbf{b}_1 \triangleq [b_{-1}, b_0, b_1, \dots, b_Q]^T$ are the $(Q + 2) \times 1$ extended differential symbol vectors, while $\hat{\mathbf{r}} \triangleq [\mathbf{r}_0^T, \mathbf{r}_1^T, \dots, \mathbf{r}_{Q+1}^T]^T$ and $\hat{\mathbf{v}} \triangleq [\mathbf{v}_0^T, \mathbf{v}_1^T, \dots, \mathbf{v}_{Q+1}^T]^T$. Fig. 2 sketches out the SLS model for a simple noiseless example with one frame per symbol ($N_f = 1$). Due to the presence of the residual timing offset $\tau \in [0, T_s)$, to detect $Q = 2$ transmitted symbols, $Q + 2 = 4$ symbol intervals have to be collected, or equivalently, the sample vectors $\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. Hence, the SLS-NMSDD scheme can be formulated according to the following proposition.

Proposition 3 SLS-NMSDD: The GLRT SLS-NMSDD mixed-integer OP is

$$\hat{\mathbf{a}}^{(\text{SLS-NMSDD})} = \arg \max_{\mathbf{a}} \left\{ \max_{\mathbf{g}_0, \mathbf{g}_1} \Lambda_{\text{SLS}}(\hat{\mathbf{r}} | \mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) \right\} \quad (31)$$

where the GLLM is

$$\begin{aligned} \Lambda_{\text{SLS}}(\hat{\mathbf{r}} | \mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) &\triangleq 2\hat{\mathbf{r}}^T \left[(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right] \\ &\quad - 2\mathbf{g}_0^T (\mathbf{b}_0^T \mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \\ &\quad - \left[\mathbf{g}_0^T (\mathbf{b}_0^T \mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + \mathbf{g}_1^T (\mathbf{b}_1^T \mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right]. \quad (32) \end{aligned}$$

Proof: From the joint SLS NR sampled model (30), the GLLM can be expressed as

$$\begin{aligned} \Lambda_{\text{SLS}}(\hat{\mathbf{r}} | \mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) &= 2\hat{\mathbf{r}}^T \left[(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right] \\ &\quad - \left[(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right]^T \\ &\quad \times \left[(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right] \quad (33) \end{aligned}$$

which after some algebra gives (32). \blacksquare

B. Compressive Sampling MSDD With Symbol Level Synchronization

Bearing in mind the CMSDD and SLS-NMSDD schemes discussed in Sections III-B and IV-A, respectively, let us now combine the CS and SLS frameworks. Exploiting (17) and (29), the compressed waveform received within the k th symbol interval reads

$$\mathbf{y}_k = (\mathbf{I}_{N_f} \otimes \Phi_k) [b_k \mathbf{g}_0 + b_{k-1} \mathbf{g}_1] + \boldsymbol{\xi}_k, \quad 0 \leq k \leq Q+1. \quad (34)$$

Accordingly, the joint compressed model for the $Q+2$ symbols takes the form

$$\hat{\mathbf{y}} = \hat{\Psi} \left[(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \right] + \hat{\boldsymbol{\xi}} \quad (35)$$

where $\hat{\mathbf{y}} \triangleq [\mathbf{y}_0^T, \mathbf{y}_1^T, \dots, \mathbf{y}_{Q+1}^T]^T$ and $\hat{\boldsymbol{\xi}} \triangleq [\boldsymbol{\xi}_0^T, \boldsymbol{\xi}_1^T, \dots, \boldsymbol{\xi}_{Q+1}^T]^T$ are the extended $(Q+2)MN_f \times 1$ compressed measurement and noise vectors, respectively, and

$$\hat{\Psi} \triangleq \text{diag} \{ \mathbf{I}_{N_f} \otimes \Phi_0, \mathbf{I}_{N_f} \otimes \Phi_1, \dots, \mathbf{I}_{N_f} \otimes \Phi_{Q+1} \} \quad (36)$$

is the $(Q+2)MN_f \times (Q+2)NN_f$ extended block level measurement matrix, such that $\hat{\Psi} \hat{\Psi}^T = \mathbf{I}_{(Q+2)MN_f}$. Thus, based on the joint model (35), the MSDD version adopting both SLS and CS can be stated as follows.

Proposition 4 SLS-CMSDD: The GLRT SLS-CMSDD integer OP is

$$\hat{\mathbf{a}}^{(\text{SLS-CMSDD})} = \arg \max_{\mathbf{a}} \left\{ \Delta_{\text{SLS}}(\hat{\mathbf{y}} | \mathbf{a}) \right\} \quad (37)$$

where the cost function is expressed as

$$\begin{aligned} \Delta_{\text{SLS}}(\hat{\mathbf{y}} | \mathbf{a}) &\triangleq \sum_{k=0}^Q \sum_{\ell=0}^Q b_k b_\ell \left[\mathbf{y}_k^T (\mathbf{I}_{N_f} \otimes \Phi_k \Phi_\ell^T) \mathbf{y}_\ell \right. \\ &\quad \left. + \mathbf{y}_{k+1}^T (\mathbf{I}_{N_f} \otimes \Phi_{k+1} \Phi_{\ell+1}^T) \mathbf{y}_{\ell+1} \right]. \quad (38) \end{aligned}$$

Proof: See Appendix B. \blacksquare

Some remarks about the SLS-CMSDD scheme are now in order.

- 1) When the frame level measurement matrices Φ_k are all orthogonal to each other, i.e., $\Phi_k \Phi_\ell^T = \mathbf{0}_{M \times M}$, $\forall k, \ell$ with $0 \leq k, \ell \leq Q$, the detector again does not exist.
- 2) When the frame level measurement matrices are the same for all the symbols, i.e., $\Phi_0 = \Phi_1 = \dots = \Phi_Q$, the cost function (38) to be optimized takes the following simpler form

$$\Delta_{\text{SLS}}(\hat{\mathbf{y}} | \mathbf{a}) = \sum_{k=1}^Q \sum_{\ell=0}^{k-1} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{i+\ell} \left(\mathbf{y}_k^T \mathbf{y}_\ell + \mathbf{y}_{k+1}^T \mathbf{y}_{\ell+1} \right) \quad (39)$$

whereas in the case they differ from symbol to symbol its general form is

$$\begin{aligned} \Delta_{\text{SLS}}(\hat{\mathbf{y}} | \mathbf{a}) &= \sum_{k=1}^Q \sum_{\ell=0}^{k-1} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{i+\ell} \left[\mathbf{y}_k^T (\mathbf{I}_{N_f} \otimes \Phi_k \Phi_\ell^T) \mathbf{y}_\ell \right. \\ &\quad \left. + \mathbf{y}_{k+1}^T (\mathbf{I}_{N_f} \otimes \Phi_{k+1} \Phi_{\ell+1}^T) \mathbf{y}_{\ell+1} \right]. \quad (40) \end{aligned}$$

- 3) Similar to the CMSDD, the SLS-CMSDD shows the advantage of enabling data detection while skipping the reconstruction step, and its performance is basically dictated by the choice on both the measurement matrices and the compression ratio μ .
- 4) In view of relaxing the demanding constraints not only on the sampling rate but also on the timing synchronization accuracy, it is expected that SLS-CMSDD offers more competitive performance-versus-complexity trade-offs when compared to both the CMSDD and the SLS-NMSDD, which require either a higher timing accuracy or a higher sampling rate, respectively.

V. COMPRESSED SPHERE DECODER

Despite the major advantages of CMSDD and SLS-CMSDD as noncoherent differential detectors working directly on sub-NR sampled signals, it can be argued from the Propositions 2 and 4 that maximizing the objective functions (24) and (40) over all the possible realizations of \mathbf{a} involves an exhaustive search that exhibits combinatorial complexity. Accordingly, such a route turns to be quite unfeasible even for short block sizes Q . To gain a manageable OP we resort to the SD.

Basics on SD: SD is an effective iterative decoding algorithm originally proposed to efficiently solve the shortest vector problem (SVP) in a lattice [40]–[45], i.e.,

$$\hat{\mathbf{s}}^{(\text{SVP})} = \arg \min_{\mathbf{s} \in \mathbb{Z}^{N \times 1}} \{ \|\mathbf{U}\mathbf{s}\|_2 \} \quad (41)$$

where \mathbf{U} is the $M \times N$ full-rank generator matrix, whereas the lattice is defined as the set of $M \times 1$ vectors $\mathcal{L}(\mathbf{U}) \triangleq \{ \mathbf{U}\mathbf{s} \mid \mathbf{s} \in \mathbb{Z}^{N \times 1} \}$. In the SD, only those lattice points are searched iteratively that lie within a sphere of radius ρ centered at $\mathbf{0}_{M \times 1}$, i.e., only the subset of $\hat{\mathbf{s}} \in \mathbb{Z}^{N \times 1}$ satisfying the condition $\|\mathbf{U}\mathbf{s}\|_2 \leq \rho$. Iteration after iteration, ρ is progressively made smaller and smaller, so that the search space is greatly reduced compared

with a naive method based on exhaustive search. As a result, the SVP, which is NP hard as shown in [43], can be iteratively solved at low-degree polynomial complexity (cubic or higher) in the length N of the optimal vector to be searched for.

The SD algorithm was proposed for MSDD in [46], for frequency-flat Rayleigh fading channels to improve the performance over DF-DD [47], and successively, was extended to UWB detection in the MSDD scheme proposed in [12]. In the sequel, we will illustrate how to apply the SD framework to the CMSDD and SLS-CMSDD proposed in Sections III-B and IV-B, respectively, leading thus to the concept of CS-based SD, or CSD for short.

CS-Based SD: To make our problem SD-compatible, let us reformulate the objective functions in (24) and (40) in an easy-to-evaluate form. In the case of the CMSDD, the maximum value of the objective function amounts to

$$\Delta_{\text{Max}}(\mathbf{y}|\mathbf{a}) = \sum_{k=1}^Q \sum_{\ell=0}^{k-1} |\bar{\mathbf{y}}_k^T \Phi_k \Phi_\ell^T \bar{\mathbf{y}}_\ell| \quad (42)$$

and subtracting (24) from (42) yields an equivalent objective function (to be minimized)

$$\check{\Delta}(\mathbf{y}|\mathbf{a}) = \sum_{k=1}^Q \sum_{\ell=0}^{k-1} |\bar{\mathbf{y}}_k^T \Phi_k \Phi_\ell^T \bar{\mathbf{y}}_\ell| \times \left[1 - \text{sign} \left\{ \bar{\mathbf{y}}_k^T \Phi_k \Phi_\ell^T \bar{\mathbf{y}}_\ell \right\} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{I+i} \right] \quad (43)$$

where, depending upon the sign of $\prod_{i=1}^{k-\ell} [\mathbf{a}]_{I+i}$, each term inside the square brackets takes a value in $\{0, 2\}$. Similarly, in the case of SLS-CMSDD, an equivalent objective function can be defined as

$$\begin{aligned} \check{\Delta}_{\text{SLS}}(\mathbf{y}|\mathbf{a}) &= \sum_{k=1}^Q \sum_{\ell=0}^{k-1} |\mathbf{y}_k^T (\mathbf{I}_{N_f} \otimes \Phi_k \Phi_\ell^T) \mathbf{y}_\ell \\ &\quad + \mathbf{y}_{k+1}^T (\mathbf{I}_{N_f} \otimes \Phi_{k+1} \Phi_{\ell+1}^T) \mathbf{y}_{\ell+1}| \\ &\times \left[1 - \text{sign} \left\{ \mathbf{y}_k^T (\mathbf{I}_{N_f} \otimes \Phi_k \Phi_\ell^T) \mathbf{y}_\ell \right. \right. \\ &\quad \left. \left. + \mathbf{y}_{k+1}^T (\mathbf{I}_{N_f} \otimes \Phi_{k+1} \Phi_{\ell+1}^T) \mathbf{y}_{\ell+1} \right\} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{I+i} \right]. \end{aligned} \quad (44)$$

For the ease of notation, let us now define

$$Z_{\ell,k} \triangleq \begin{cases} \bar{\mathbf{y}}_k^T \Phi_k \Phi_\ell^T \bar{\mathbf{y}}_\ell, & \text{CMSDD} \\ \mathbf{y}_k^T (\mathbf{I}_{N_f} \otimes \Phi_k \Phi_\ell^T) \mathbf{y}_\ell \\ \quad + \mathbf{y}_{k+1}^T (\mathbf{I}_{N_f} \otimes \Phi_{k+1} \Phi_{\ell+1}^T) \mathbf{y}_{\ell+1}, & \text{SLS - CMSDD.} \end{cases} \quad (45)$$

Hence, the OP related to the CMSDD or SLS-CMSDD results in the general form

$$\hat{\mathbf{a}}_{\text{opt}} = \arg \min_{\mathbf{a}} \{ \Xi(\mathbf{y}|\mathbf{a}) \} \quad (46)$$

where

$$\Xi(\mathbf{y}|\mathbf{a}) \triangleq \sum_{k=1}^Q \sum_{\ell=0}^{k-1} \eta_{\ell,k} |Z_{\ell,k}| \quad (47)$$

with

$$\eta_{\ell,k} \triangleq \left[1 - \text{sign} \left\{ Z_{\ell,k} \right\} \prod_{i=1}^{k-\ell} [\mathbf{a}]_{I+i} \right] \quad (48)$$

and $Z_{\ell,k}$ given by (45). From (46)–(48), the following remarks can be obtained: *i)* the objective function (47) consists of the sum of the non-negative coefficients $|Z_{\ell,k}|$, weighted by the unknowns $\eta_{\ell,k} \in \{0, 2\}$; *ii)* the partial objective

$$\Xi_j(\mathbf{y}|\mathbf{a}_j) \triangleq \sum_{k=1}^j \sum_{\ell=0}^{k-1} \eta_{\ell,k} |Z_{\ell,k}|, \quad 1 \leq j \leq Q \quad (49)$$

depends only on $\mathbf{a}_j \triangleq [[\mathbf{a}]_1, [\mathbf{a}]_2, \dots, [\mathbf{a}]_j]^T$ and given \mathbf{a}_{j-1} , \mathbf{a}_j depends only on $[\mathbf{a}]_j$; *iii)* in light of features *i)* and *ii)*, (47) defines a sphere in the Q -dimensional lattice of the vectors $\mathbf{a} \in \{\pm 1\}^Q$ [43]. Therefore, (46)–(48) combined with remarks *i)–iii)* fully comply with the SD framework, and as a consequence our OP is amenable to be solved. It is worth mentioning that the above formulation of our objective function is not the same as the conventional SD since it is a nonlinear function of \mathbf{a} . Nonetheless, the possibility of estimating an element of \mathbf{a} based on the previously estimated elements in a sequential manner, makes it solvable as an SD problem.

Implementation of CSD: Concerning the implementation of the iterative algorithm, at the generic n th SD iteration, a *necessary condition* for any tentative estimate $\hat{\mathbf{a}}^{(n)}$ to lie inside the sphere of radius $\rho^{(n)} > 0$ is given by

$$\Xi_j(\mathbf{y} | \hat{\mathbf{a}}_j^{(n)}) \leq \rho^{(n)}, \quad 1 \leq j \leq Q. \quad (50)$$

Based on condition (50), the CSD can be computationally arranged according to the pseudo-code outlined in Table I. We note that the CSD algorithm is initialized by the solution $\hat{\mathbf{a}}^{\text{DC-DD}}$ obtained by applying the low-complexity DC-DD scheme proposed in [34], which also gives the initial radius $\rho^{(0)}$ by evaluating (47). The iterations go on with a smaller and smaller sphere as search space, with the candidate $\hat{\mathbf{a}}_Q^{(n)}$ found at the previous iterations lying on its surface. When at a given iteration, for a certain j , condition (50) is satisfied for both values of $[\hat{\mathbf{a}}^{(n)}]_j$, i.e., $\{\pm 1\}$, a random value is taken from the candidate set $\mathcal{A}_j^{(n)}$, and if none of the values satisfies (50), j is decreased by 1 and $[\hat{\mathbf{a}}^{(n)}]_{j-1}$ is tried with the other value from the candidate set. Eventually, the algorithm stops when the candidate set $\mathcal{A}_1^{(n)}$ results to be empty, i.e., all the conditions on the candidate sets have been checked without reducing the sphere radius, thus meaning that the objective has safely reached its minimum value. It is worth mentioning that the set of coefficients $Z_{\ell,k}$ can be precomputed before the iterations, or even can be hard-quantized to two levels, and the unknowns $\eta_{\ell,k}$ take non-negative integer-values so checking the Q conditions at each iteration in Table I requires only real or integer format additions combined with logical operations, thus

TABLE I
PSEUDO-CODE FOR CSD

Pseudo-Code for CSD	
Input:	$Z_{\ell,k}$, for $k = 1, \dots, Q$, $\ell = 0, \dots, k-1$
Initialize:	$n = 0$, $\hat{\mathbf{a}}^{(0)} = \hat{\mathbf{a}}^{\text{DC-DD}}$, $\rho^{(0)} = \Xi(\mathbf{y} \hat{\mathbf{a}}^{\text{DC-DD}})$
Repeat	
	Candidate set for $[\hat{\mathbf{a}}^{(n)}]_1$:
	$\mathcal{A}_1^{(n)} = \{[\hat{\mathbf{a}}^{(n)}]_1 \in \{\pm 1\} \Xi_1(\mathbf{y} \hat{\mathbf{a}}_1^{(n)}) \leq \rho^{(n)}\}$
	Choose a tentative estimate of $[\hat{\mathbf{a}}^{(n)}]_1$ from $\mathcal{A}_1^{(n)}$
	Candidate set for $[\hat{\mathbf{a}}^{(n)}]_2$ given $[\hat{\mathbf{a}}^{(n)}]_1$:
	$\mathcal{A}_2^{(n)} = \{[\hat{\mathbf{a}}^{(n)}]_2 \in \{\pm 1\} \Xi_2(\mathbf{y} \hat{\mathbf{a}}_2^{(n)}) \leq \rho^{(n)}\}$
	Choose a tentative estimate of $[\hat{\mathbf{a}}^{(n)}]_2$ from $\mathcal{A}_2^{(n)}$
	\vdots
	Candidate set for $[\hat{\mathbf{a}}^{(n)}]_Q$ given $[\hat{\mathbf{a}}^{(n)}]_1, \dots, [\hat{\mathbf{a}}^{(n)}]_{Q-1}$:
	$\mathcal{A}_Q^{(n)} = \{[\hat{\mathbf{a}}^{(n)}]_Q \in \{\pm 1\} \Xi_Q(\mathbf{y} \hat{\mathbf{a}}_Q^{(n)}) \leq \rho^{(n)}\}$
	Choose a tentative estimate of $[\hat{\mathbf{a}}^{(n)}]_Q$ from $\mathcal{A}_Q^{(n)}$
	$\hat{\mathbf{a}}_{\text{opt}} \leftarrow \hat{\mathbf{a}}^{(n)}$
	$\rho^{(n+1)} \leftarrow \Xi_Q(\mathbf{y} \hat{\mathbf{a}}^{(n)}) = \Xi(\mathbf{y} \hat{\mathbf{a}}_{\text{opt}})$
	Set $n = n + 1$
Until	$\mathcal{A}_1^{(n)} = \emptyset$
Output:	$\hat{\mathbf{a}}_{\text{opt}}$

contributing in keeping the complexity at affordable levels in solving the OP (46)–(48).

VI. SIMULATION RESULTS

In this section, the proposed sub-NR MSDD schemes are tested through numerical simulations over realistic multipath environments. In particular, the bit error rate (BER) metric is quantified as a function of either the mean-bit-energy-to-noise-spectral-density ratio defined as $E_b/N_0 \triangleq N_f \|\mathbf{h}\|_2^2 / \sigma_v^2$ or the compression ratio μ , for different values of the block size Q and frame number N_f , with ideal pulse level or coarse symbol level timing synchronization.

A. Simulation Setup

The transmitted signal consists of a number of bursts including Q consecutive differentially encoded binary symbols according to rule (3). In each symbol interval, the frame length is chosen to be $T_f = 50$ ns, whereas the transmitted pulse per frame $q(t)$ is selected as the second derivative of a Gaussian shape with width $T_q = 1$ ns. The slow-fading channel is assumed to be time-invariant within each burst, but randomly varying from burst to burst according to the IEEE 802.15.3a CM1 model [4], whose maximum delay spread is 25 ns. The bandwidth of the receive low-pass filter is taken as $W = 2$ GHz, and consequently, the NR is 4 GHz., i.e., $N = 200$ samples per frame. Therefore, assuming a compression ratio of μ means that only $M = \mu N$ samples are employed by the detection algorithm. Further, we consider frame level measurement matrices Φ_k , $0 \leq k \leq Q$. We initially generate them as having zero-mean equi-distributed Gaussian entries and later orthonormalize the rows. Two different options are considered for compressing each symbol within the burst: *i*) same measurement matrix (SMM), i.e., $\Phi_k = \Phi_{k+1}$, $0 \leq k \leq Q-1$; *ii*) different measurement matrix (DMM), i.e., $\Phi_k \neq \Phi_{k+1}$, $0 \leq k \leq Q-1$.

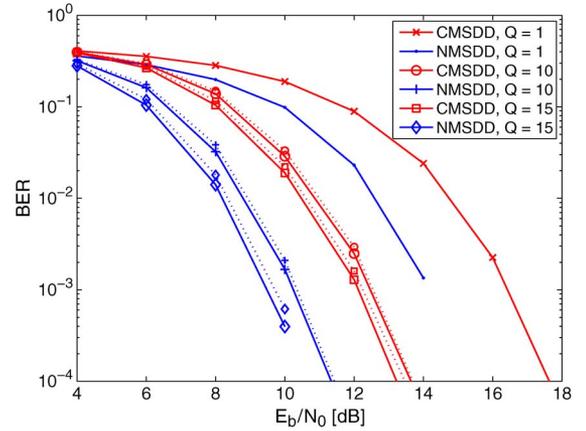


Fig. 3. BER comparison of NMSDD and CMSDD with SMM, along with sbDF-DD and csDF-DD (dotted lines), different block sizes, $N_f = 1$ and $\mu = 0.5$.

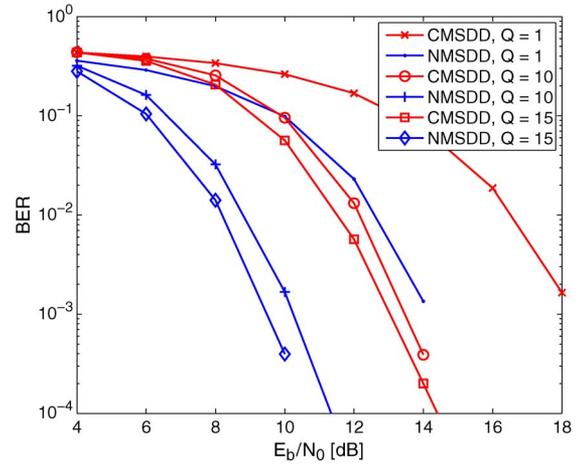


Fig. 4. BER comparison of NMSDD and CMSDD with DMM, different block sizes, $N_f = 1$ and $\mu = 0.5$.

TABLE II
BER PERFORMANCE OF CMSDD WITH VARYING N_f AND $Q = 10$

E_b/N_0 [dB]	$N_f = 1$	$N_f = 5$	$N_f = 10$
4	0.4009	0.4031	0.4038
6	0.3053	0.3074	0.3072
8	0.1558	0.1587	0.1582
10	0.0376	0.0384	0.0365
12	0.0034	0.0032	0.0032

B. BER With Ideal Timing Synchronization

Figs. 3 and 4 depict the BER metric versus the E_b/N_0 ratio for the SMM and DMM options, respectively, for the compression ratio $\mu = 0.5$, and block sizes $Q = 1, 10, 15$. The number of frames per symbol is set to $N_f = 1$ since for ideal timing synchronization the frame averaging in (14) or (22) is such that higher values are expected not to affect the performance, as confirmed by Table II. For both figures, increasing Q gives reasonably better performance when compared with $Q = 1$, namely the conventional DD, regardless of choosing SMM or DMM. Indeed, at the BER of 10^{-3} , when moving from $Q = 1$ to $Q = 15$ both the NMSDD and CMSDD gain around 4 dB, regardless whether we choose SMM or DMM. Given that the channel stays invariant at least within the block

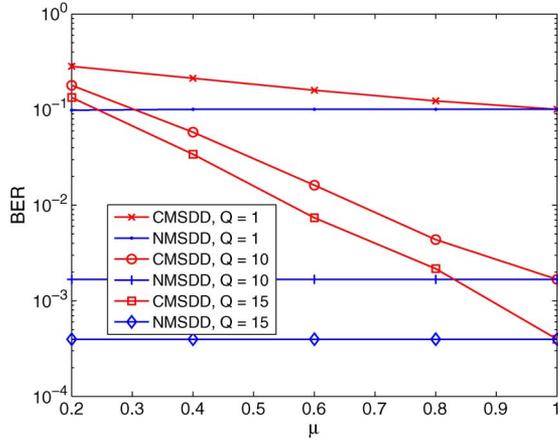


Fig. 5. BER comparison of NMSDD and CMSDD with SMM, different block sizes, $N_f = 1$, different values of μ and $E_b/N_0 = 10$ dB.

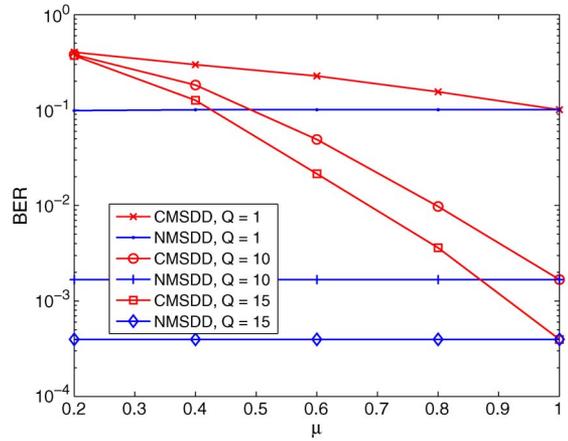


Fig. 6. BER comparison of NMSDD and CMSDD with DMM, different block sizes, $N_f = 1$, different values of μ and $E_b/N_0 = 10$ dB.

interval, i.e., $(Q + 1)N_f T_f$, the above behavior is basically due to the multi-symbol structure of both the algorithms, which advantageously exploit the signal correlation not only between adjacent symbols as the DD does, but also between many other symbols up to the block size apart. Further, in spite of the 2 dB loss suffered by the CMSDD against the NMSDD in case of SMM, the former presents the advantage of halving the sampling rate, thus reducing the computational load required to detect each data burst. It is further to be remarked that changing the setup from SMM to DMM, i.e., passing from Figs. 3 to 4, causes the performance of CMSDD to deteriorate by 3 dB. It can be imagined that the limiting case of this scenario will be in line with the first remark made both in Sections III-B and IV-B, explaining that frame level orthogonal measurement matrices can make the detector independent of the differential symbols, and thus ineffective. Note that for the sake of comparison, we also plot in Fig. 3 the results of using sorted block-wise DF-DD (sbDF-DD) [38] and its compressed version CS based DF-DD (csDF-DD) [35] (both in dotted lines). The results point out that the proposed CSD-based detector has a slight edge over the csDF-DD. Although, both require ideal timing recovery, the latter is further limited to the SMM scenario. On the other side, as quantified in Section VI-C, the SLS-CMSDD is the only scheme that can considerably relax the timing accuracy, thereby enabling good performance-versus-complexity trade-off solutions. However, it is worth mentioning that our proposed schemes, CMSDD and SLS-CMSDD are not restricted to be used only with CSD as an alternative to exhaustive search, but other strategies, e.g., DF can also be opted. Figs. 5 and 6 show the BER versus the compression ratio μ at $E_b/N_0 = 10$ dB, for both the NMSDD and CMSDD, with $Q = 14, 10, 15$, and adopting the SMM and DMM options, respectively. As expected, increasing μ , the CMSDD performance improves till it approaches that of the NMSDD when $\mu = 1$.

C. BER With Coarse Symbol Level Timing Synchronization

Concerning the SLS-based detectors, we choose $N_f = 10$ frames per symbol since in this configuration the timing offset is acquired with a coarse accuracy at symbol level, and thus, the value of N_f is expected to affect performance (as will be

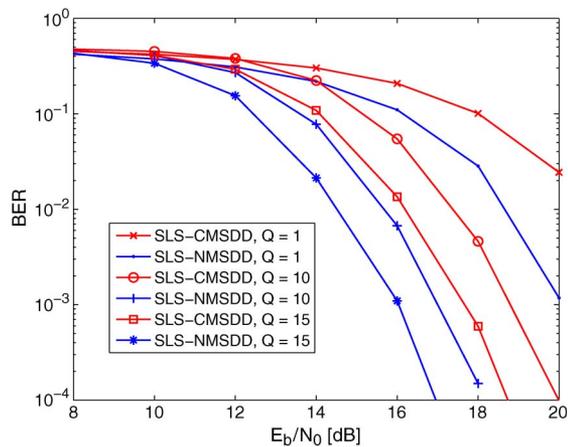


Fig. 7. BER comparison of SLS-NMSDD and SLS-CMSDD with SMM, different block sizes, $N_f = 10$, $\mu = 0.5$ and $\tau \in [0.1T_s, 0.9T_s]$.

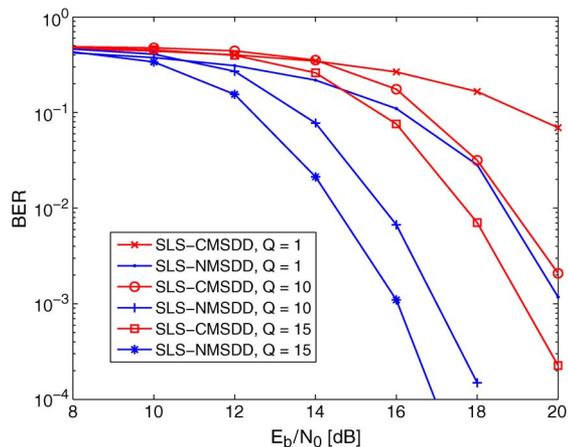


Fig. 8. BER comparison of SLS-NMSDD and SLS-CMSDD with DMM, different block sizes, $N_f = 10$, $\mu = 0.5$ and $\tau \in [0.1T_s, 0.9T_s]$.

shown in a while). Figs. 7 and 8 quantify the BER in case the SMM and DMM options are adopted, respectively, with each figure referring to both SLS-NMSDD and SLS-CMSDD schemes, with block sizes $Q = 1, 10, 15$, and compression ratio $\mu = 0.5$. Given that the timing offset of each received burst is uniformly distributed over the symbol interval to comply

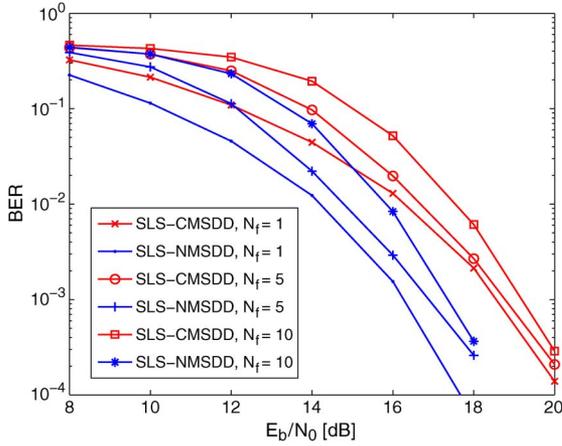


Fig. 9. BER comparison of SLS-NMSDD and SLS-CMSDD with SMM, $Q = 10$, $\mu = 0.5$, different values of N_f and $\tau \in [0.1T_s, 0.9T_s]$.

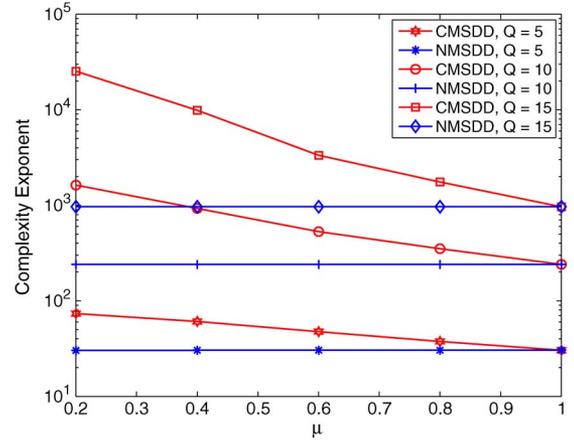


Fig. 11. Complexity comparison of SD against compressed and Nyquist rate symbols, different block sizes, varying μ , SMM, $E_b/N_0 = 10$ dB, $N_f = 1$.

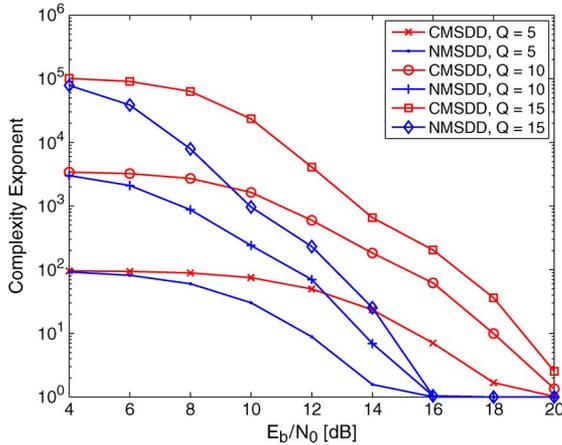


Fig. 10. Complexity comparison of SD against compressed and Nyquist rate symbols, different block sizes, SMM, $N_f = 1$.

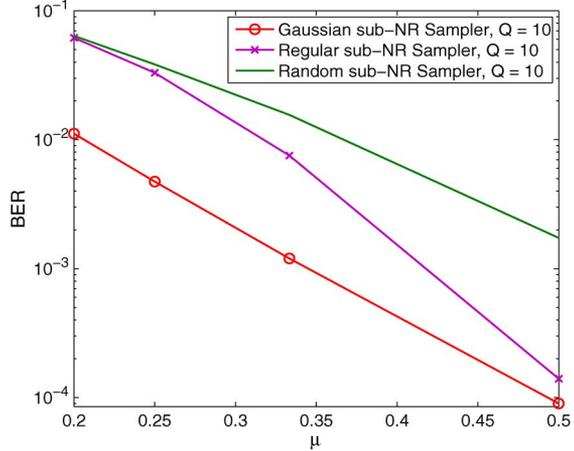


Fig. 12. BER comparison of CMSDD with Gaussian, regular and random sub-NR sampler, different block sizes, SMM, $N_f = 1$, different values of μ and $E_b/N_0 = 14$ dB.

with the condition of asynchronous access to the channel and in line with the assumption that timing synchronization is performed at symbol level only, the BER curves are averaged over the uniformly distributed timing offset $\tau \in [0.1T_s, 0.9T_s]$. Similar to the NMSDD and CMSDD, it is apparent that the performance of the SLS detectors at both NR and CS sampling improves using a larger block size Q , whereas the DMM incurs again a loss of around 3 dB with respect to the SMM option. It is worth emphasizing that the advantages of the SLS-CMSDD are twofold, in the sense that it can relax the stringent requirements on both the sampling rate and the timing accuracy at an affordable performance loss against the more demanding NMSDD and CMSDD schemes. In addition, similar to Figs. 5 and 6, it can be proved that as $\mu \rightarrow 1$ the SLS-CMSDD and SLS-NMSDD meet at the same BER level. Fig. 9 shows the averaged BER for the SLS-NMSDD and SLS-CMSDD, with SMM, $Q = 10$ and different values of the frame number, namely $N_f = 1, 5, 10$. It can be argued that the performance improves when N_f decreases given the corresponding decrease in noise accumulation in the absence of frame averaging.

In Figs. 10 and 11, we give the complexity performance of CSD against NR SD, for varying SNR and μ , respectively. We define the complexity metric as ‘Complexity Exponent’

which basically is the total number of sum operations consumed during a search (since there are no multiplications in our cost functions). As expected, the CSD has a comparatively higher Complexity Exponent but decreases with increasing SNR and/or μ , thereby indicating a trade-off between performance and complexity.

Finally, in Fig. 12, we show a BER performance of CMSDD when using different types of samplers (i.e., measurement matrices). Although, we use a Gaussian sub-NR sampler in general but other samplers can also be used. Fig. 12 shows the BER performance when the Gaussian, regular and random sub-NR samplers are used, respectively. We see that the Gaussian sampler shows better performance than the regular sub-NR sampler especially at lower values of μ , whereas the random sub-NR sampler lags behind the other two.

VII. CONCLUSION

In this paper, we have presented compressive sampling-based multiple symbol differential detectors using the GLRT approach, both in the presence of full timing information as well as with symbol-level synchronization only. The detectors

avoid an explicit reconstruction step and operate on the compressed samples directly. The detectors perform better when the measurement matrices are the same for each symbol within the block but have the ability to work even when they are different. The detectors do not exist for the case of orthogonal measurement matrices. Combined with sphere decoding, the proposed detectors offer very low complexity and power efficient detection possibilities.

APPENDIX A PROOF OF PROPOSITION 2

From the joint compressed model (18), the GLLM given \mathbf{a} and \mathbf{h} can be written as

$$\begin{aligned} \Omega(\mathbf{y}|\mathbf{a}, \mathbf{h}) &\triangleq 2\mathbf{y}^T \Psi(\mathbf{b} \otimes \mathbf{I}_{NN_f})(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \\ &\quad - [(\mathbf{b} \otimes \mathbf{I}_{NN_f})(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})]^T \Psi^T \\ &\quad \times \Psi [(\mathbf{b} \otimes \mathbf{I}_{NN_f})(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})] \end{aligned} \quad (51)$$

which, in view of the structure of \mathbf{y} , can be rearranged as

$$\begin{aligned} \Omega(\mathbf{y}|\mathbf{a}, \mathbf{h}) &= 2\mathbf{y}^T \Psi(\mathbf{b} \otimes \mathbf{I}_{NN_f})(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \\ &\quad - (\mathbf{1}_{N_f \times 1} \otimes \mathbf{h})^T (\mathbf{b} \otimes \mathbf{I}_{NN_f})^T \Psi^T \Psi \\ &\quad \times (\mathbf{b} \otimes \mathbf{I}_{NN_f})(\mathbf{1}_{N_f \times 1} \otimes \mathbf{h}) \\ &= 2N_f \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) \mathbf{h} \\ &\quad - N_f \mathbf{h}^T (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N) \mathbf{h} \end{aligned} \quad (52)$$

where $\Phi \triangleq \text{diag}\{\Phi_0, \Phi_1, \dots, \Phi_Q\}$ is a $(Q+1)M \times (Q+1)N$ block-diagonal matrix, $\bar{\mathbf{y}} \triangleq [\bar{\mathbf{y}}_0^T, \bar{\mathbf{y}}_1^T, \dots, \bar{\mathbf{y}}_Q^T]^T$, with $\bar{\mathbf{y}}_k$ given by (22).

Following the GLRT principle, the first step is to maximize (52) over \mathbf{h} . Thus, setting the gradient with respect to \mathbf{h} to zero yields

$$2N_f \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) - 2N_f \mathbf{h}^T [(\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N)] = \mathbf{0}^T \quad (53)$$

which leads to the estimate

$$\hat{\mathbf{h}} = \mathbf{H} \bar{\mathbf{y}} \quad (54)$$

where

$$\mathbf{H} \triangleq [(\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N)]^{-1} [\Phi(\mathbf{b} \otimes \mathbf{I}_N)]^T. \quad (55)$$

Then, after plugging (54) into (52), we obtain the cost function

$$\begin{aligned} \Gamma(\mathbf{y}|\mathbf{a}) &\triangleq 2N_f \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) \mathbf{H} \bar{\mathbf{y}} \\ &\quad - N_f [\mathbf{H} \bar{\mathbf{y}}]^T (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N) \mathbf{H} \bar{\mathbf{y}}. \end{aligned} \quad (56)$$

Considering that

$$\begin{aligned} &- N_f [\mathbf{H} \bar{\mathbf{y}}]^T (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N) \mathbf{H} \bar{\mathbf{y}} \\ &= -N_f \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) [(\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N)]^{-1} \\ &\quad \times [(\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N)] \mathbf{H} \bar{\mathbf{y}} \\ &= -N_f \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) \mathbf{H} \bar{\mathbf{y}} \end{aligned} \quad (57)$$

after some algebra and dropping the immaterial factor N_f , (56) can be reformulated as

$$\Gamma[\mathbf{y}|\mathbf{a}] = \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) \mathbf{S}^{-1} (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \bar{\mathbf{y}} \quad (58)$$

where

$$\mathbf{S} \triangleq (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \Phi (\mathbf{b} \otimes \mathbf{I}_N) = \sum_{k=0}^Q \Phi_k^T \Phi_k \quad (59)$$

is a positive (semi-)definite matrix³ depending only on the measurement matrices Φ_k , $0 \leq k \leq Q$. Intensive numerical simulations have shown that the presence of \mathbf{S} in (58) affects the differential detection of \mathbf{a} only in a weak way, i.e., a specific $\hat{\mathbf{a}}$ maximizing (58) also (approximately) maximizes

$$\Delta[\mathbf{y}|\mathbf{a}] = \bar{\mathbf{y}}^T \Phi(\mathbf{b} \otimes \mathbf{I}_N) (\mathbf{b} \otimes \mathbf{I}_N)^T \Phi^T \bar{\mathbf{y}}. \quad (60)$$

Hence, after rearranging (60) according to $\bar{\mathbf{y}}$ and Φ , the objective function of the CMSDD OP takes the form of (21), which concludes the proof.

APPENDIX B PROOF OF PROPOSITION 4

From the joint compressed model (35), the GLLM given \mathbf{a} , \mathbf{g}_0 and \mathbf{g}_1 for the SLS-CMSDD can be put into the form

$$\begin{aligned} \Omega_{\text{SLS}}(\hat{\mathbf{y}}|\mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) &\triangleq 2\hat{\mathbf{y}}^T \hat{\Psi} [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1] \\ &\quad - [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1]^T \hat{\Psi}^T \\ &\quad \times \hat{\Psi} [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1]. \end{aligned} \quad (61)$$

After some algebra, (61) can be rearranged as

$$\begin{aligned} \Omega_{\text{SLS}}(\hat{\mathbf{y}}|\mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) &= 2\hat{\mathbf{y}}^T \hat{\Psi} [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1] \\ &\quad - 2\mathbf{g}_0^T (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \\ &\quad - \mathbf{g}_0^T (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 \\ &\quad - \mathbf{g}_1^T (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 \end{aligned} \quad (62)$$

where $\hat{\mathbf{y}}$ and $\hat{\Psi}$ are the extended measurement vector and block level measurement matrix, respectively, defined in Section IV-B. It is worth observing in (62) that

$$\mathbf{g}_0^T (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1 = \sum_{\ell=1}^Q [\mathbf{a}]_{\ell} \varpi_{\ell} \quad (63)$$

where $\varpi_{\ell} \triangleq \mathbf{g}_0^T (\mathbf{I}_{N_f} \otimes \Phi_{\ell}^T \Phi_{\ell}) \mathbf{g}_1$. Note that due to the orthogonality of \mathbf{g}_0 and \mathbf{g}_1 , ϖ_{ℓ} will have very few addends.⁴ Now

³As detailed in [34], the positive (semi-)definite property of \mathbf{S} can be easily shown through the eigenvalue decomposition (EVD).

⁴If Φ_{ℓ} are the same, for $\ell = 1, \dots, Q$, then ϖ_{ℓ} s would also be the same, and (63) will result in a summation over $[\mathbf{a}]_{\ell}$ s scaled by a constant value. If Φ_{ℓ} are different, for $\ell = 1, \dots, Q$, then ϖ_{ℓ} s would produce a scrambling effect over $[\mathbf{a}]_{\ell}$ s.

given that it is equally probable for a_i to be $+1$ or -1 , we can expect that the result can (on the average) be considered as vanishing for a sufficiently large block size Q . Hence, the objective function in (62) can be further simplified as

$$\begin{aligned} \Omega_{\text{SLS}}(\hat{\mathbf{y}}|\mathbf{a}, \mathbf{g}_0, \mathbf{g}_1) & \\ & \simeq 2\hat{\mathbf{y}}^T \hat{\Psi} [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})\mathbf{g}_0 + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})\mathbf{g}_1] \\ & - \mathbf{g}_0^T (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_0 \\ & - \mathbf{g}_1^T (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{g}_1. \end{aligned} \quad (64)$$

In accordance with the GLRT principle, setting the gradient of (64) to zero with respect to \mathbf{g}_0 and \mathbf{g}_1 gives

$$\hat{\mathbf{g}}_i = \mathbf{G}_i \hat{\mathbf{y}}, \quad i = 0, 1 \quad (65)$$

where

$$\begin{aligned} \mathbf{G}_i & \triangleq [(\mathbf{b}_i \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_i \otimes \mathbf{I}_{NN_f})]^{-1} \\ & \times [\hat{\Psi} (\mathbf{b}_i \otimes \mathbf{I}_{NN_f})]^T, \quad i = 0, 1. \end{aligned} \quad (66)$$

Thus, upon plugging (65) into (64), after some algebra we obtain

$$\begin{aligned} \Gamma_{\text{SLS}}(\hat{\mathbf{y}}|\mathbf{a}) & \triangleq \hat{\mathbf{y}}^T \hat{\Psi} (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \mathbf{S}_0^{-1} (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\mathbf{y}} \\ & + \hat{\mathbf{y}}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \mathbf{S}_1^{-1} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\mathbf{y}} \end{aligned} \quad (67)$$

where \mathbf{S}_0 and \mathbf{S}_1 are defined, respectively, as

$$\begin{aligned} \mathbf{S}_0 & \triangleq (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_0 \otimes \mathbf{I}_{NN_f}) \\ & = \mathbf{I}_{N_f} \otimes \sum_{k=0}^Q \Phi_k^T \Phi_k \end{aligned} \quad (68)$$

$$\begin{aligned} \mathbf{S}_1 & \triangleq (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})^T \hat{\Psi}^T \hat{\Psi} (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f}) \\ & = \mathbf{I}_{N_f} \otimes \sum_{k=1}^{Q+1} \Phi_k^T \Phi_k. \end{aligned} \quad (69)$$

From (68) and (69), it can be remarked that: *i*) \mathbf{S}_0 and \mathbf{S}_1 are independent of both \mathbf{b}_0 and \mathbf{b}_1 ; *ii*) applying the EVD, it can be proved that \mathbf{S}_0 and \mathbf{S}_1 are positive (semi-)definite matrices; *iii*) it can be shown that the inverses of \mathbf{S}_0 and \mathbf{S}_1 affect the maximization of (67) in a weak way (in terms of \mathbf{a}). Hence, collecting together the above results, we are left with the approximate cost function

$$\begin{aligned} \Delta_{\text{SLS}}(\hat{\mathbf{y}}|\mathbf{a}) & \triangleq \hat{\mathbf{y}}^T \hat{\Psi} [(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})(\mathbf{b}_0 \otimes \mathbf{I}_{NN_f})^T \\ & + (\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})(\mathbf{b}_1 \otimes \mathbf{I}_{NN_f})^T] \hat{\Psi}^T \hat{\mathbf{y}}. \end{aligned} \quad (70)$$

Finally, similar to the approach pursued for the CMSDD, (70) can be reformulated in the equivalent form given by (38), thus concluding the proof.

REFERENCES

- [1] M. Z. Win and R. A. Scholtz, "Impulse radio: How it works," *IEEE Commun. Lett.*, vol. 2, no. 2, pp. 36–38, Feb. 1998.
- [2] L. Yang and G. B. Giannakis, "Ultra-wideband communications: An idea whose time has come," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 26–54, Nov. 2004.
- [3] A. F. Molisch, J. R. Foerster, and M. Pendergrass, "Channel models for ultrawideband personal area networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 14–21, Dec. 2003.
- [4] A. F. Molisch, "Ultra-wide-band propagation channels," *Proc. IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.
- [5] J. D. Choi and W. E. Stark, "Performance of ultra-wideband communications with suboptimal receivers in multipath channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 9, pp. 1754–1766, Dec. 2002.
- [6] V. Lottici, A. N. D'Andrea, and U. Mengali, "Channel estimation for ultra-wideband communications," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 9, pp. 1638–1645, Dec. 2002.
- [7] K. Witrisal *et al.*, "Noncoherent ultra-wideband systems," *IEEE Signal Process. Mag.*, vol. 26, no. 4, pp. 48–66, Jul. 2009.
- [8] R. Hoctor and H. Tomlinson, "Delay-hopped transmitted-reference RF communications," in *Proc. IEEE UWBST*, May 2002, pp. 265–269.
- [9] M. Ho, S. Somayazulu, J. Foerster, and S. Roy, "A differential detector for UWB communications system," in *Proc. IEEE VTC*, May 2002, vol. 4, pp. 1896–1900.
- [10] Y. L. Chao and R. Scholtz, "Optimal and suboptimal receivers for ultra-wideband transmitted reference systems," in *Proc. IEEE Globecom*, Dec. 2003, pp. 759–763.
- [11] N. Guo and R. C. Qiu, "Improved autocorrelation demodulation receivers based on multiple-symbol detection for UWB communications," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2026–2031, Aug. 2006.
- [12] V. Lottici and Z. Tian, "Multiple symbol differential detection for UWB communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1656–1666, May 2008.
- [13] Z. Tian and V. Lottici, "Low-complexity ML timing acquisition for UWB communications in dense multipath channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 3031–3038, Nov. 2005.
- [14] Z. Tian and G. B. Giannakis, "A GLRT approach to data-aided timing acquisition in UWB radios—Part I: Algorithms," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2956–2967, Nov. 2005.
- [15] L. Yang and G. B. Giannakis, "Timing ultra-wideband signals with dirty templates," *IEEE Trans. Wireless Commun.*, vol. 53, no. 11, pp. 1952–1963, Nov. 2005.
- [16] V. Lottici, Z. Tian, and G. Leus, "A novel approach to UWB data detection with symbol-level synchronization," *Phys. Commun.*, vol. 2, no. 4, pp. 296–305, Dec. 2009.
- [17] P. Schvan, D. Pollex, S.-C. Wang, C. Falt, and N. Ben-Hamida, "A 22 GS/s 5 b ADC in 0.130 μm SiGe BiCMOS," in *Proc. IEEE ISSCC*, Feb. 2006, pp. 572–573.
- [18] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
- [19] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets StrangDFix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [20] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [21] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [22] S. Kirolos *et al.*, "Analog-to-information conversion via random demodulation," in *Proc. IEEE DCAS*, Oct. 2006.
- [23] T. Ragheb *et al.*, "A prototype hardware for random demodulation based compressive analog-to-digital conversion," in *Proc. MWSCAS*, Aug. 2008, pp. 37–40.
- [24] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
- [25] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [26] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [27] J. Kusuma, I. Maravic, and M. Vetterli, "Sampling with finite rate of innovation: Channel and timing estimation for UWB and GPS," in *Proc. IEEE ICC*, May 2003, pp. 3540–3544.

- [28] Y. Vanderperren, G. Leus, and W. Dehaene, "Performance analysis of a flexible subsampling receiver for pulsed UWB signals," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4134–4142, Aug. 2009.
- [29] J. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: Channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 383–395, Oct. 2007.
- [30] T. C.-K. Liu, X. Dong, and W.-S. Lu, "Compressed sensing maximum likelihood channel estimation for ultra-wideband impulse radio," in *Proc. IEEE ICC*, Jun. 2009, pp. 1–5.
- [31] Z. Wang, G. R. Arce, J. L. Paredes, and B. M. Sadler, "Compressed detection for ultra-wideband impulse radio," in *Proc. IEEE SPAWC*, Jun. 2007, pp. 1–5.
- [32] Z. Wang, G. R. Arce, B. M. Sadler, J. L. Paredes, and X. Ma, "Compressed detection for pilot assisted ultra-wideband impulse radio," in *Proc. IEEE ICUBW*, Sep. 2007, pp. 393–398.
- [33] A. Oka and L. Lampe, "A compressed sensing receiver for UWB impulse radio in bursty applications like wireless sensor networks," *Phys. Commun.*, vol. 2, no. 4, pp. 248–264, Dec. 2009.
- [34] S. Gishkori, G. Leus, and V. Lottici, "Compressive sampling based differential detection for UWB impulse radio signals," *Phys. Commun.*, vol. 5, no. 2, pp. 185–195, Jun. 2012.
- [35] A. Schenk and R. F. H. Fischer, "Compressed-sensing (decision-feedback) differential detection in impulse-radio ultra-wideband systems," in *Proc. ICUBW*, Sep. 2011, pp. 121–125.
- [36] S. Gishkori, G. Leus, and V. Lottici, "Compressive sampling based multiple symbol differential detection for UWB IR signals," in *Proc. IEEE ICUBW*, Sep. 2012, pp. 130–134.
- [37] S. Gishkori and G. Leus, "Compressive sampling based energy detection of ultra-wideband pulse position modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3866–3879, Aug. 2013.
- [38] A. Schenk and R. F. H. Fischer, "Decision-feedback differential detection in impulse-radio ultra-wideband systems," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1604–1611, Jun. 2011.
- [39] S. M. Kay, *Fundamentals of Statistical signal processing, Detection theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [40] U. Finche and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comput.*, vol. 44, pp. 463–471, Apr. 1985.
- [41] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.
- [42] O. Damen, H. El Gamal, and G. Caire, "On maximum likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.
- [43] B. Hassibi and H. Vikalo, "On sphere decoding algorithm. I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.
- [44] O. Goldreich, D. Micciancio, S. Safra, and J.-P. Seifert, "Approximating shortest lattice vectors is not harder than approximating closest lattice vectors," *Inform. Process. Lett.*, vol. 71, no. 2, pp. 55–61, Jul. 1999.
- [45] A. Burg *et al.*, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.
- [46] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Trans. Wireless Commun.*, vol. 53, no. 12, pp. 1981–1985, Dec. 2005.
- [47] R. Schober, W. H. Gerstacker, and J. B. Huber, "Decision-feedback differential detection of MPDSK for flat Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1025–1035, Jul. 1999.



(compressed sensing), signal processing for communications and wireless communications.



Shahzad Gishkori (S'14) received the B.Sc. degree in electrical engineering from the University of Engineering and Technology Lahore, Lahore, Pakistan, in 2002 and the M.Sc. degree (*cum laude*) in electrical engineering from the Delft University of Technology, Delft, The Netherlands, in 2009. He has been pursuing the Ph.D. degree at the Faculty of Electrical Engineering, Mathematics and Computer Science of Delft University of Technology, Delft, since 2009.

From 2002 to 2007, he worked in the industry.

His research interests include compressive sampling

Vincenzo Lottici received the Dr. Ing. (*cum laude*) and the B.T.A. degrees in electrical engineering from the University of Pisa, Pisa, Italy, in 1985 and 1986, respectively.

From 1987 to 1993 he was engaged in the research of sonar digital signal processing algorithms. Since 1993 he has been with the Department of Information Engineering of the University of Pisa, Pisa, where he is currently an Assistant Professor in Communication Systems. He participated in several international and national research projects, and as TPC member, in numerous IEEE conferences in wireless communications and signal processing, such as Globecom2013, WCNC2013, ICUBW2013, WCNC2012, ICUBW2011, ICC2011, WCNC2011, ICASSP2010, PIMRC2010, WCNC2010, CIP2010, Globecom2009, SPAWC2009, EUSIPCO2006, Globecom2006. He joined the Editorial Board of EURASIP Advances on Signal Processing. His research interests include the broad area of signal processing for communications, with emphasis on synchronization, dynamic resource allocation, cognitive radio and compressive sensing.

Dr. Lottici received the Best Paper Award in 2006 for the work "A Theoretical Framework for Soft Information Based Synchronization in Iterative (Turbo) Receivers", EURASIP Journal on Wireless Communications and Networking, April 2005, by the EU-funded project Network of Excellence in Wireless Communications (NEWCOM).



Geert Leus (F'12) received the electrical engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in June 1996 and May 2000, respectively.

Currently, he is an "Antoni van Leeuwenhoek" Full Professor at the Faculty of Electrical Engineering, Mathematics and Computer Science of the Delft University of Technology, Delft, The Netherlands. His research interests are in the area of signal processing for communications.

Dr. Leus received a 2002 IEEE Signal Processing Society Young Author Best Paper Award and a 2005 IEEE Signal Processing Society Best Paper Award. Geert Leus was the Chair of the IEEE Signal Processing for Communications and Networking Technical Committee, and an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE SIGNAL PROCESSING LETTERS, and the EURASIP Journal on Advances in Signal Processing. Currently, he is a member of the IEEE Sensor Array and Multichannel Technical Committee and serves as the Editor-in-Chief of the EURASIP Journal on Advances in Signal Processing.