UWB Sparse/Diffuse Channels, Part II: Estimator Analysis and Practical Channels

Nicolò Michelusi, *Student Member, IEEE,* Urbashi Mitra, *Fellow, IEEE,* Andreas F. Molisch, *Fellow, IEEE,* and Michele Zorzi, *Fellow, IEEE*

***Abstract*—In this two-part paper, the problem of channel estimation in Ultra Wide-Band (UWB) systems is investigated. In Part I, a novel Hybrid Sparse/Diffuse (HSD) model is proposed for the UWB channel, and new channel estimation strategies are designed for this model. In this paper (Part II), a Mean-Squared Error (MSE) analysis of the *Generalized MMSE* and *Generalized Thresholding Estimators* developed in Part I is performed, for the asymptotic regimes of low and high SNR. The analysis quantifies the achievable MSE performance of these schemes over unstructured estimators. Specifically, we prove that it is beneficial to be conservative in the estimation of the sparse component, *i.e.*, to assume that the sparse component is sparser than it actually is. Moreover, we analyze the scenario with a non-orthogonal pilot sequence, and establish a connection between the Generalized Thresholding estimator and conventional sparse approximation algorithms proposed in the literature. In addition to the theoretical analysis, these channel estimation schemes are evaluated in a more realistic geometry-based channel emulator, for which the HSD model developed in Part I is an approximation. The numerical results are shown to match the expected asymptotic MSE behavior. Moreover, the proposed estimation techniques are shown to outperform conventional unstructured and purely sparse estimators, from both an MSE and a bit error rate perspectives, even for the realistic geometry-based channel model.**

I. Introduction

Ultra Wide-Band (UWB) signaling had been originally proposed as a technology for indoor mobile and multiple-access communications [1]–[3]. Due to its significant bandwidth, UWB offers high precision localization [4], robustness against multipath fading [5] and immunity to narrow-band interference [6], thus representing a compelling solution for applications such as short-range high-speed broadband access [7], Wireless Body Area Networks (WBANs) [8], covert communication links, through-wall imaging, high-resolution ground-penetrating radar and asset tracking [9]–[11]. However, the performance of coherent UWB transceivers relies on the availability of accurate channel estimates (*e.g.*, [12]–[14]). Thus, it is important to design channel estimation strategies that exploit the structural and statistical properties of UWB propagation to achieve the best estimation accuracy.

N. Michelusi and M. Zorzi are with the Department of Information Engineering, University of Padova, Italy, e-mail: {michelusi,zorzi}@dei.unipd.it.

U. Mitra and A.F. Molisch are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, USA, e-mail: {ubli,molisch}@usc.edu.

In the literature, different UWB channel models have been proposed, based on extensive measurement campaigns, *e.g.*, geometry-based, clustered Saleh-Valenzuela [10], [15] or dense models [16]. An overview of ray-tracing and statistical channel models is available in [17] from one of the authors. A generalization of the Saleh-Valenzuela model [15] is presented in [10] and provides a parametrization for different environments of interest, *e.g.*, residential indoor, office indoor, built-up outdoor, industrial indoor, farm environments, and WBANs. According to this model, the UWB channel is represented as a clustered pattern of arrivals of MultiPath Components (MPCs). Moreover, it presents a dense model for industrial environments and WBANs, where the richer interaction among MPCs gives rise to multipath fading. On the other hand, [18], [19] adopt an hybrid model, combining both a geometric approach for the resolvable individual specular components (echoes), arising from reflections from the scatterers in the environment, and a statistical approach to model the diffuse component of the channel. While in [18] the diffuse component is modeled as a diffuse tail associated with each MPC, [19] uses a different approach, and models the diffuse component as independent of the resolvable MPCs arrivals.

Due to the hybrid nature of UWB propagation, in Part I of this paper [20] (henceforth simply referred to as "Part I"), we have proposed a novel Hybrid Sparse/Diffuse (HSD) model, which is able to capture the main UWB propagation phenomena. Specifically, this model combines a *sparse component*, which models the fine grained delay resolution of UWB receivers, and a *diffuse component*, which models other effects not properly described by a sparse channel representation, namely, diffuse scattering from rough surfaces, unresolvable MPCs and frequency dispersion. Moreover, we have designed the *Generalized MMSE* (GMMSE) and the *Generalized Thresholding* (GThres) channel estimation schemes.

Our contributions are as follows: following Part I of this paper, here in Part II we present a Mean-Squared Error (MSE) analysis of the GMMSE and GThres estimators, in the asymptotic regimes of high and low Signal to Noise Ratios (SNR). In particular, we prove that, in these regimes, it is beneficial, from the perspective of minimizing the MSE, to use a conservative approach in the estimation of the sparse component of the channel, by assuming the sparse component to be sparser than it actually is. Also, we prove that the GMMSE estimator outperforms the GThres estimator in these regimes. Moreover, we analyze the scenario with a non-orthogonal pilot sequence, and establish a connection between the GThres estimator and conventional sparse approximation

algorithms proposed in the literature. Finally, we evaluate the simplified HSD model and the designed channel estimation strategies based on a realistic UWB channel emulator. In particular, we adopt the channel emulator developed in [18]. We observe that the model in [19] which assumes statistically independent sparse and diffuse components is closer to our proposed HSD model; however, we believe the model of [18] may be more physically accurate and thus we select [18] to validate the accuracy of our proposed schemes. Thus, using [18] we are able to evaluate the robustness and sensitivity of the HSD-based estimation strategies. We show that the new estimation schemes outperform conventional unstructured estimators, *e.g.*, Least Squares (LS), and purely sparse estimators, from both an MSE and a Bit Error Rate (BER) perspectives. The simulation results show that, despite its simplicity, the HSD model can effectively capture key UWB propagation mechanisms. Moreover, we argue that, due to its hybrid nature, this model easily accommodates a wide range of practical scenarios, where the channel exhibits a purely sparse, diffuse or hybrid nature.

The paper is organized as follows: in Section II, we present the system model, and review the HSD channel model developed in Part I. In Section III, we review the GMMSE and GThres estimators. In Section IV, we perform an asymptotic MSE analysis of these estimation schemes, and we discuss the results. In Section V, we analyze the case with a non-orthogonal pilot sequence. In Section VI, we present simulation results for both a channel that follows the HSD model, and a more realistic geometry-based stochastic channel emulator developed in [18]. Section VII concludes the paper. The Appendix provides the derivations of key lemmas which enable the analysis in Section IV.

Notation: We use lower-case and upper-case bold letters for column vectors ($\mathbf{a}$) and matrices ($\mathbf{A}$), respectively. The scalar $\mathbf{a}_k$ (or $\mathbf{a}(k)$) is the kth entry of vector $\mathbf{a}$, and $\mathbf{A}_{k,j}$ (or $\mathbf{A}(k,j)$) is the (k,j)th entry of matrix $\mathbf{A}$. A positive definite (positive semi-definite) matrix $\mathbf{A}$ is denoted by $\mathbf{A} \succ 0$ ($\mathbf{A} \succeq 0$). The transpose, complex conjugate of matrix $\mathbf{A}$ is denoted by $\mathbf{A}^*$. If $\mathbf{A} \succeq 0$ with eigenvalue decomposition $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, we define its square root as $\sqrt{\mathbf{A}} = \mathbf{U}\sqrt{\mathbf{D}}\mathbf{U}^*$. The matrix $\mathbf{I}_K$ is the $K \times K$ unit matrix. The trace operator is denoted by $\mathrm{tr}(\mathbf{A}) = \sum_k \mathbf{A}_{k,k}$. The vector $\mathbf{a} \odot \mathbf{b}$ is the component-wise (Schur) product of vectors $\mathbf{a}$ and $\mathbf{b}$. We use $p(\cdot)$ to indicate a continuous or discrete probability distribution. The expectation of random variable x, conditioned on y, is written as $\mathbb{E}[x|y]$. The circular Gaussian distribution with mean $\mathbf{m}$ and covariance $\mathbf{\Sigma}$ is denoted by $\mathcal{CN}(\mathbf{m}, \mathbf{\Sigma})$,[1] the Bernoulli distribution with parameter q by $\mathcal{B}(q)$, and the exponential distribution with mean m by $\mathcal{E}(m)$. The indicator function is denoted by $\mathcal{I}(\cdot)$.

II. System model

We adopt the same signal model as in Part I. Namely, we consider a single-user UWB system. The source transmits a sequence of $M = N + L - 1$ pilot symbols, $x(k), k = -(L-1), \ldots, N-1$, over a channel $h(l), l = 0, \ldots, L-1$ with delay spread $L \geq 1$. The received, discrete time, baseband signal over the corresponding observation interval of duration N is denoted by $y(k)$, $k = 0, \ldots, N-1$. Letting $\mathbf{y} = [y(0), y(1), \ldots, y(N-1)]^T$, $\mathbf{h} = [h(0), h(1), \ldots, h(L-1)]^T$, we have the matrix representation $\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w}$, where $\mathbf{X} \in \mathbb{C}^{N \times L}$ is the $N \times L$ Toeplitz matrix associated with the pilot sequence, having the vector $[x(-k), x(-k+1), \ldots, x(-k+N-1)]^T, k = 0, \ldots, L-1$, as its kth column, and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_L)$ is the noise vector.

The HSD model, developed in Part I [20] and [21] for the UWB channel $\mathbf{h}$, is given by

$$\mathbf{h} = \mathbf{a}_s \odot \mathbf{c}_s + \mathbf{h}_d, \tag{1}$$

where the term $\mathbf{a}_s \odot \mathbf{c}_s \in \mathbb{C}^L$ represents the *sparse component*,[2] and $\mathbf{h}_d \in \mathbb{C}^L$ is the *diffuse component*. In particular, $\mathbf{a}_s \sim \mathcal{B}(q)^L$ is the *sparsity pattern*, with $q \ll 1$ so as to enforce sparsity. The positions of the "1"s correspond to the delays of the resolvable MPCs. In the sequel, we refer to the non-zero entries of the sparse vector $\mathbf{a}_s \odot \mathbf{c}_s \in \mathbb{C}^L$ as *active sparse components*.

The vector $\mathbf{c}_s \in \mathbb{C}^L$ carries the *sparse coefficients*, drawn from the continuous probability distribution $p(\mathbf{c}_s)$, with second order moment $\mathbb{E}[\mathbf{c}_s \mathbf{c}_s^*] = \mathbf{\Lambda}_s$, where $\mathbf{\Lambda}_s$ is a diagonal matrix with diagonal elements given by the Power Delay Profile (PDP) of the active sparse components $\mathbf{\Lambda}_s(k,k) = \mathcal{P}_s(k), \forall k$.

Finally, we use the Rayleigh fading assumption for the diffuse component $\mathbf{h}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_d)$, where $\mathbf{\Lambda}_d$ is diagonal, with diagonal entries given by the PDP $\mathbf{\Lambda}_d(k,k) = \mathcal{P}_d(k), k = 0, \ldots, L-1$.

Since the LS estimate is a sufficient statistic of the channel (see, *e.g.*, [20]), we consider the effective model

$$\mathbf{h}_{\mathrm{LS}} = (\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*\mathbf{y} = \mathbf{h} + \sqrt{\mathbf{S}}^{-1}\mathbf{n}, \tag{2}$$

where we have defined the SNR matrix $\mathbf{S} = \frac{\mathbf{X}^*\mathbf{X}}{\sigma_w^2}$, and the noise vector $\mathbf{n} = \frac{1}{\sigma_w^2}\sqrt{\mathbf{S}}^{-1}\mathbf{X}^*\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$. We will refer to this sufficient representation of the observation model in the following treatment. In particular, with a slight abuse of notation, we will refer to $\mathbf{h}_{\mathrm{LS}}$ as the observed sequence.

In the sequel, as in Part I [20], we assume an orthogonal pilot sequence, so that $\mathbf{S}$ is diagonal. Under this assumption, the samples of the LS estimate are independent and a per-tap, rather than joint, estimation approach is optimal. We will discuss the non-orthogonal case in Section V, where we establish a connection between the GThres estimator and the sparse approximation algorithm proposed in [22].

III. Review of the Estimators

In Part I, we have developed channel estimation strategies based on the HSD model. In particular, we have proposed the GMMSE and GThres estimators, for the scenario where the PDP of $\mathbf{h}_d$, $\mathbf{\Lambda}_d$, is known at the receiver, whereas the vector of sparse coefficients $\mathbf{c}_s$ and the sparsity level q are treated as

[1] For a vector $\mathbf{x} = \mathbf{x}_R + i\mathbf{x}_I \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{x}_R = \mathrm{Re}(\mathbf{x})$ and $\mathbf{x}_I = \mathrm{Im}(\mathbf{x})$, we define the covariance matrices of its real and imaginary parts as $\mathbb{E}[\mathbf{x}_R\mathbf{x}_R^*] = \mathbb{E}[\mathbf{x}_I\mathbf{x}_I^*] = \frac{\mathrm{Re}(\mathbf{\Sigma})}{2}$ and $\mathbb{E}[\mathbf{x}_I\mathbf{x}_R^*] = -\mathbb{E}[\mathbf{x}_R\mathbf{x}_I^*] = \frac{\mathrm{Im}(\mathbf{\Sigma})}{2}$.

[2] In the following, we use the terms *sparse*, *specular* and *resolvable MPCs* interchangeably. In fact, the physical specular components (resolvable MPCs) of the channel can be modeled and represented by a sparse vector.

deterministic and unknown parameters. These estimators are reviewed in this section.

Let $\tilde{q} \in (0,1)$ be the prior Bernoulli parameter for the generation of the sparsity pattern $\mathbf{a}_s$, assumed in the estimation phase, and $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)$. Note that we may have $\tilde{q} \neq q$, *i.e.*, the assumed prior may be different from the true prior generating the sequence $\mathbf{a}_s$.

The GMMSE and GThres estimators compute an LS estimate of the sparse coefficient vector $\mathbf{c}_s$, given by $\hat{\mathbf{c}}_s = \mathbf{h}_{\rm LS}$. Then, the sparsity pattern $\mathbf{a}_s$ is estimated by MMSE or MAP, respectively, yielding the estimate $\hat{\mathbf{a}}_s$. Finally, an MMSE estimate of the diffuse component $\mathbf{h}_d$ is computed based on the residual LS error, $\mathbf{h}_{\rm LS} - \hat{\mathbf{a}}_s \odot \hat{\mathbf{c}}_s$. The overall estimate is given by

$$\hat{\mathbf{h}}(k) = \hat{\mathbf{a}}_s(k)\mathbf{h}_{\rm LS}(k) + (1-\hat{\mathbf{a}}_s(k))\,\frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\mathbf{h}_{\rm LS}(k). \quad (3)$$

The two estimators differ in the estimate of the sparsity pattern $\mathbf{a}_s$. Namely, GMMSE and GThres performs an MMSE and an MAP estimate of $\mathbf{a}_s$, respectively, assuming the prior $\mathbf{a}_s \sim \mathcal{B}(\tilde{q})^L$. We have

$$\begin{cases} \hat{\mathbf{a}}_s^{(\rm GMMSE)}(k) = \left(1 + e^{\alpha}\exp\left\{-\frac{\mathbf{S}_{k,k}|\mathbf{h}_{\rm LS}(k)|^2}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\right\}\right)^{-1} \\ \hat{\mathbf{a}}_s^{(\rm GThres)}(k) = \mathcal{I}\left(|\mathbf{h}_{\rm LS}(k)|^2 \geq \alpha\left(1/\mathbf{S}_{k,k} + \mathcal{P}_d(k)\right)\right). \end{cases} \quad (4)$$

Note that the GThres estimator leads to a thresholding of the LS estimate, where the threshold $\alpha\left(1/\mathbf{S}_{k,k} + \mathcal{P}_d(k)\right)$ represents the effective noise background for the estimation of the sparse component, scaled by the parameter $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)$, which is related to the assumed sparsity level of the channel: in the positions where the LS samples are above the threshold, the channel is estimated as being a combination of specular and diffuse components; otherwise, it is estimated as being diffuse only.

It is worth noting that, in the derivation of these estimators and in the MSE analysis in Section IV, we assume that the PDP of $\mathbf{h}_d$ is known at the receiver. In practice, $\mathcal{P}_d(k)$ must be estimated. In Part I [20], we have developed a PDP estimator of $\mathbf{h}_d$ based on the Expectation-Maximization (EM) algorithm, assuming an exponential structure of the PDP [16], [19], [23], which is exploited to average the fading over the channel delay dimension rather than over subsequent realizations of the fading process.

The true value of q is unknown in practice, and therefore these estimators use an estimate or a guess $\tilde{q}$ of q. In the next section, we perform an asymptotic MSE analysis of the GMMSE and GThres estimators, in the high and low SNR regions, and we prove that using $\tilde{q} < q$ in the estimation process leads to improved MSE, thus suggesting that the knowledge of q is not crucial to achieve high accuracy.

IV. MSE ANALYSIS

Let $\hat{\mathbf{h}}^{(X)}$, $X \in \{\rm GMMSE, GThres, LS\}$ be either the GMMSE, the GThres or the LS estimator. We define the MSE of $\hat{\mathbf{h}}^{(X)}$ as a function of the SNR matrix $\mathbf{S}$ as

$$\mathrm{MSE}^{(X)}(\mathbf{S}) = \frac{1}{L}\mathbb{E}\left[\left\|\hat{\mathbf{h}}^{(X)} - \mathbf{h}\right\|_2^2\right] = \frac{1}{L}\sum_k \mathrm{MSE}_k^{(X)}(\mathbf{S}_{k,k}), \quad (5)$$

where, owing to the use of per-tap estimation approaches, the sum is over the MSE terms associated with the estimation of the kth channel coefficient, *i.e.*,

$$\mathrm{MSE}_k^{(X)}(\mathbf{S}_{k,k}) = \mathbb{E}\left[\left|\hat{\mathbf{h}}^{(X)}(k) - \mathbf{h}(k)\right|^2\right]. \quad (6)$$

The expectation is computed with respect to the joint probability distribution $p(\mathbf{a}_s)p(\mathbf{c}_s)p(\mathbf{h}_d)p(\mathbf{n})$. In this section, we study the asymptotic behavior of each term $\mathrm{MSE}_k^{(X)}(\mathbf{S}_{k,k})$, $k = 0,\ldots,L-1$, in the limit of high ($\mathbf{S}_{k,k} \to +\infty$) and low ($\mathbf{S}_{k,k} \to 0^+$) SNR.

For the sake of a more concise notation, we define $y = \mathbf{h}_{\rm LS}(k)$, $\hat{h}(y) = \hat{\mathbf{h}}(k)$, $a_s = \mathbf{a}_s(k)$, $c_s = \mathbf{c}_s(k)$, $h_d = \frac{1}{\sqrt{\mathcal{P}_d(k)}}\mathbf{h}_d(k)$ (normalized to have unit variance), $h = \mathbf{h}(k)$, $n = \mathbf{n}(k)$, $S = \mathbf{S}_{k,k}$ and $P_d = \mathcal{P}_d(k)$. From (1) and (2), we can then rewrite the observation model associated with the kth channel entry as

$$y = a_s c_s + \sqrt{P_d}h_d + \frac{1}{\sqrt{S}}n, \quad (7)$$

where $a_s \sim \mathcal{B}(q)$, $h_d \sim \mathcal{CN}(0,1)$, $n \sim \mathcal{CN}(0,1)$.

For the LS estimator, we have $\mathrm{mse}_k^{(\rm LS)}(S) \triangleq S\mathrm{MSE}_k^{(\rm LS)}(S) = \mathbb{E}\left[S\,|y-h|^2\right] = 1$. Hence, the normalized MSE, $\mathrm{mse}_k^{(\rm LS)}(S)$, is a constant, independent of the SNR. Herein, we show that the GMMSE and GThres estimators exhibit the same behavior in the asymptotic high and low SNR, *i.e.*, letting $\mathrm{mse}_k^{(X)}(S) \triangleq S\mathrm{MSE}_k^{(X)}(S)$, we have

$$\lim_{S\to 0(\infty)} \mathrm{mse}_k^{(X)}(S) = \text{const.} > 0, \;\; X \in \{\rm GMMSE, GThres\},$$

for a proper constant, which depends on the asymptotic regime and on the estimator. To this end, let

$$f^{(X)}\left(\sqrt{S}y, n\right) = S\left|\hat{h}(y) - h\right|^2. \quad (8)$$

Then, we have

$$\mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[f^{(X)}\left(\sqrt{S}h + n, n\right)\right], \quad (9)$$

where the expectation is calculated with respect to $h = a_s c_s + \sqrt{P_d}h_d$ and $n \sim \mathcal{CN}(0,1)$, which are independent of the SNR S. From Lemma 1 in the Appendix, we can exchange the limit operator with the expectation, yielding, for $S_{\rm lim} \in \{0, +\infty\}$,

$$\lim_{S\to S_{\rm lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S\to S_{\rm lim}} f^{(X)}\left(\sqrt{S}h + n, n\right)\right]. \quad (10)$$

We evaluate (10) for the GMMSE and GThres estimators in Sections IV-A and IV-B, respectively.

A. Generalized MMSE estimator

Substituting the expression of the GMMSE estimator (3) and (4) in (8), we obtain, after some algebraic manipulation,

$$f^{(\rm GMMSE)}\left(\sqrt{S}y, n\right) = \left| n - \frac{e^{\alpha}\exp\left\{\frac{-S|y|^2}{1+SP_d}\right\}\frac{\sqrt{S}y}{1+SP_d}}{1 + e^{\alpha}\exp\left\{\frac{-S|y|^2}{1+SP_d}\right\}}\right|^2. \quad (11)$$

We distinguish the three cases $S \to +\infty$ with $P_d = 0$, $S \to +\infty$ with $P_d > 0$, and $S \to 0$.

1) High SNR with no diffuse component: $S \to +\infty$, $P_d = 0$:
When $P_d = 0$, we have $\sqrt{S}y = \sqrt{S}a_s c_s + n$ and

$$f^{(\mathrm{GMMSE})}\left(\sqrt{S}a_s c_s + n, n\right) = \left| n - \frac{e^{\alpha}\exp\left\{-\left|\sqrt{S}a_s c_s + n\right|^2\right\}}{1 + e^{\alpha}\exp\left\{-\left|\sqrt{S}a_s c_s + n\right|^2\right\}}\left(\sqrt{S}a_s c_s + n\right)\right|^2.$$

In the limit of high SNR, we obtain

$$\begin{cases} \lim\limits_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(\sqrt{S}c_s + n, n\right) = |n|^2, \ a_s = 1, \ \text{a.e.}, \\ \lim\limits_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(n, n\right) = \frac{|n|^2}{\left(1+e^{\alpha}\exp\left\{-|n|^2\right\}\right)^2}, \ a_s = 0, \end{cases}$$

where a.e. stands for *almost everywhere*, *i.e.*, the limit holds except on a set with probability measure zero. In particular, this set is given by $\{c_s = 0\}$, which has probability measure zero since c_s is a continuous random variable. From (10), by averaging over $a_s \sim \mathcal{B}(q)$ and $n \sim \mathcal{CN}(0,1)$, we thus obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = q\mathbb{E}\left[|n|^2\right] + (1-q)\mathbb{E}\left[\frac{|n|^2}{\left(1+e^{\alpha}\exp\left\{-|n|^2\right\}\right)^2}\right] = q + (1-q)g(\alpha), \tag{12}$$

where we have defined $g(\alpha) = e^{-\alpha}\ln\left(1+e^{\alpha}\right)$ and we have used Lemma 2 in the Appendix.

Therefore, in the high SNR regime (*i.e.*, letting $\sigma_w^2 \to 0$, which scales the SNR matrix $\mathbf{S}$ to infinity) with no diffuse component, $\mathcal{P}_d(k) = 0, \forall k$, using (5), we obtain the following limiting MSE behavior:

$$\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) = \frac{1}{L}\sum_{k=0}^{L-1}\frac{\mathrm{mse}_k^{(\mathrm{GMMSE})}(\mathbf{S}_{k,k})}{\mathbf{S}_{k,k}} \simeq_{\infty} \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})\left(q + (1-q)g(\alpha)\right), \tag{13}$$

where we have defined $\simeq_{\infty}$ as the high SNR approximation, and we have denoted the MSE of the LS estimator as $\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}) = \frac{1}{L}\mathrm{tr}\left(\mathbf{S}^{-1}\right)$.

2) High SNR with diffuse component: $S \to +\infty$, $P_d > 0$:
From (11), we have $\lim_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(\sqrt{S}h + n, n\right) = |n|^2$. Then, from (10),

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = \mathbb{E}\left[|n|^2\right] = 1. \tag{14}$$

From (5), the limiting behavior of the overall MSE in the high SNR, with $\mathcal{P}_d(k) > 0, \forall k$, is given by

$$\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) \simeq_{\infty} \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}). \tag{15}$$

3) Low SNR: $S \to 0$:
From (11), we have

$$\lim_{S\to 0} f^{(\mathrm{GMMSE})}\left(\sqrt{S}h + n, n\right) = \left|\frac{n}{1 + e^{\alpha}\exp\left\{-|n|^2\right\}}\right|^2.$$

Then, using (10) and Lemma 2 in the Appendix, we obtain

$$\lim_{S\to 0} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = \mathbb{E}\left[\left|\frac{n}{1 + e^{\alpha}\exp\left\{-|n|^2\right\}}\right|^2\right] = g(\alpha).$$

Then, from (5), the overall MSE in the low SNR regime behaves like

$$\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) \simeq_0 \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})g(\alpha), \tag{16}$$

where we have defined $\simeq_0$ as the low SNR approximation.

B. Generalized Thresholding estimator

Substituting the expression of the GThres estimator (3) and (4) in (8), we obtain, after some algebraic manipulation,

$$f^{(\mathrm{GThres})}\left(\sqrt{S}h + n, n\right) = \mathcal{I}\left(\left|\sqrt{S}h + n\right|^2 \geq \alpha(1+SP_d)\right)|n|^2 + \mathcal{I}\left(\left|\sqrt{S}h + n\right|^2 < \alpha(1+SP_d)\right)\left|\frac{\sqrt{S}h - SP_d n}{1+SP_d}\right|^2. \tag{17}$$

Note that, if $\alpha \leq 0$, then we have a trivial thresholding operation, and the estimator is equivalent to LS. This case is of no interest. In the following, therefore, we study the case $\alpha > 0$.

Similarly to the GMMSE estimator, we distinguish the three cases $S \to +\infty$ with $P_d = 0$, $S \to +\infty$ with $P_d > 0$, and $S \to 0$.

1) High SNR with no diffuse component: $S \to +\infty$, $P_d = 0$:
When $P_d = 0$ we have $y = a_s c_s + \sqrt{S}^{-1} n$ and

$$f^{(\mathrm{GThres})}\left(\sqrt{S}a_s c_s + n, n\right) = \mathcal{I}\left(\left|\sqrt{S}a_s c_s + n\right|^2 \geq \alpha\right)|n|^2 + \mathcal{I}\left(\left|\sqrt{S}a_s c_s + n\right|^2 < \alpha\right)\left|\sqrt{S}a_s c_s\right|^2. \tag{18}$$

We have

$$\begin{cases} \lim\limits_{S\to+\infty} f^{(\mathrm{GThres})}\left(\sqrt{S}c_s + n, n\right) = |n|^2, \ a_s = 1, \ \text{a.e.}, \\ \lim\limits_{S\to+\infty} f^{(\mathrm{GThres})}\left(n, n\right) = \mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2, \ a_s = 0, \end{cases}$$

where the first limit holds a.e., *i.e.*, except on the set with zero probability measure $\{c_s = 0\}$. From (10), we then obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GThres})}(S) = q\mathbb{E}\left[|n|^2\right] + (1-q)\mathbb{E}\left[\mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2\right] = q + (1-q)w(\alpha), \tag{19}$$

where in the last step we have used the fact that $|n|^2 \sim \mathcal{E}(1)$ to compute the second expectation term, and we have defined $w(\alpha) = e^{-\alpha}(1+\alpha)$.

Then, from (5), the overall MSE in the high SNR regime with $\mathcal{P}_d(k) = 0, \forall k$, behaves like

$$\mathrm{MSE}^{(\mathrm{GThres})}(\mathbf{S}) \simeq_{\infty} \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})\left(q + (1-q)e^{-\alpha}(1+\alpha)\right).$$

2) High SNR with diffuse component: $S \to +\infty$, $P_d > 0$:
From (17), we have $\lim_{S\to+\infty} f^{(\mathrm{GThres})}\left(\sqrt{S}h+n,n\right) = |n|^2$. Then, from (10), we obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GThres})}(S) = \mathbb{E}\left[|n|^2\right] = 1. \tag{20}$$

Therefore, in the high SNR regime with $\mathcal{P}_d(k) > 0, \forall k$, the GThres estimator performs like

$$\mathrm{MSE}^{(\mathrm{GThres})}(\mathbf{S}) \simeq_\infty \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}). \tag{21}$$

3) Low SNR: $S \to 0$:
From (17), we have

$$\lim_{S\to 0} f^{(\mathrm{GThres})}\left(\sqrt{S}h+n,n\right) = \mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2. \tag{22}$$

Then, from (10), we obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GThres})}(S) = \mathbb{E}\left[\mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2\right] = w(\alpha).$$

Therefore, in the low SNR regime, the GThres estimator performs like

$$\mathrm{MSE}^{(\mathrm{GThres})}(\mathbf{S}) \simeq_0 \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})\, e^{-\alpha}(1+\alpha). \tag{23}$$

C. Discussion

The asymptotic MSE behavior of the GMMSE and GThres estimators is summarized in Table I. A plot is given in Figure 1. We compare their limiting behavior with the (unstructured) LS estimator and with the Oracle estimator, which assumes the HSD model, perfect knowledge of $\mathbf{a}_s$, and treats $\mathbf{c}_s$ as a deterministic unknown vector.

The latter, by knowing $\mathbf{a}_s$, performs an LS estimate of $\mathbf{c}_s$ and an MMSE of $\mathbf{h}_d$. Its MSE as a function of the SNR matrix $\mathbf{S}$ is given by

$$\mathrm{MSE}^{(\mathrm{Oracle})}(\mathbf{S}) = q\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}) + \frac{1-q}{L}\sum_{k=0}^{L-1}\frac{\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}.$$

The limiting MSE behavior in the table is normalized to $\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})$. Then, a value smaller than 1 indicates that the estimation accuracy, in the corresponding regime, improves over LS. Moreover, the smaller the value, the better the asymptotic MSE accuracy.

TABLE I
ASYMPTOTIC MSE BEHAVIOR OF LS, ORACLE, GMMSE AND GTHRES ESTIMATORS.
$\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right), g(\alpha) = e^{-\alpha}\ln(1+e^{\alpha}), w(\alpha) = e^{-\alpha}(1+\alpha).$

$\frac{\mathrm{MSE}^{(\mathbf{X})}(\mathbf{S})}{\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})}$	**High SNR, $\mathbf{\Lambda}_d = \mathbf{0}$**	**High SNR, $\mathbf{\Lambda}_d \succ 0$**	**Low SNR**
LS, GThres, $\alpha \leq 0$	1	1	1
Oracle	q	1	q
GMMSE	$q+(1-q)g(\alpha)$	1	$g(\alpha)$
GThres, $\alpha > 0$	$q+(1-q)w(\alpha)$	1	$w(\alpha)$

Notice that, in the high SNR with diffuse component, all estimators achieve the LS MSE accuracy. In fact, in the high SNR the diffuse component is strong compared to the noise level, *i.e.*, $\mathcal{P}_d(k) \gg 1/\mathbf{S}_{k,k}$, hence the observed channel exhibits a dense structure, yielding the same accuracy as the LS estimator.

On the other hand, in the high SNR with no diffuse component, the GMMSE and GThres estimators achieve a better estimation accuracy than LS. Their limiting behavior can be explained as follows. When $\mathbf{a}_s(k) = 1$ (with probability q), the active sparse coefficients $\mathbf{c}_s(k)$, which are much stronger than the noise background in the high SNR, are always correctly detected, and are estimated with the same estimation accuracy as LS. On the other hand, when $\mathbf{a}_s(k) = 0$ (with probability $1-q$), the GMMSE (respectively, GThres) estimator incurs a mis-detection error $\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})g(\alpha)$ ($\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})w(\alpha)$), due to strong noise samples which are mis-detected as active sparse components.

Moreover, since $g(\alpha)$ and $w(\alpha)$ are decreasing functions of $\alpha \in \mathbb{R}$ (*i.e.*, increasing functions of $\tilde{q} \in (0,1)$), with $\lim_{\alpha\to-\infty} g(\alpha) = w(0) = 1$ and $\lim_{\alpha\to+\infty} g(\alpha) = \lim_{\alpha\to+\infty} w(\alpha) = 0$, the MSE is a decreasing function of α (*i.e.*, an increasing function of $\tilde{q}$). In particular, for small values of α, the estimates of $\mathbf{a}_s$ in (4) approach 1 for both the GMMSE and the GThres estimators, hence the overall HSD estimate (3) approaches the LS solution, yielding the same LS accuracy. Conversely, for increasing values of α, the GMMSE and GThres estimators approach the MSE accuracy of the Oracle estimator. Note that, the larger α, the larger the threshold level of the GThres estimator in (4), hence the fewer noise samples are mis-detected as active sparse components, and the smaller the overall mis-detection error and MSE accuracy (a similar interpretation holds for the GMMSE estimator).

Similarly, in the low SNR, the MSE of the GMMSE and GThres estimators is a decreasing function of α. In particular, a better MSE than the Oracle estimator is achieved for α sufficiently large. In fact, the main source of error is associated with the LS estimates of the sparse coefficients. On the other hand, the MMSE estimate of the diffuse component is forced to zero at small SNR values, hence the resulting MSE

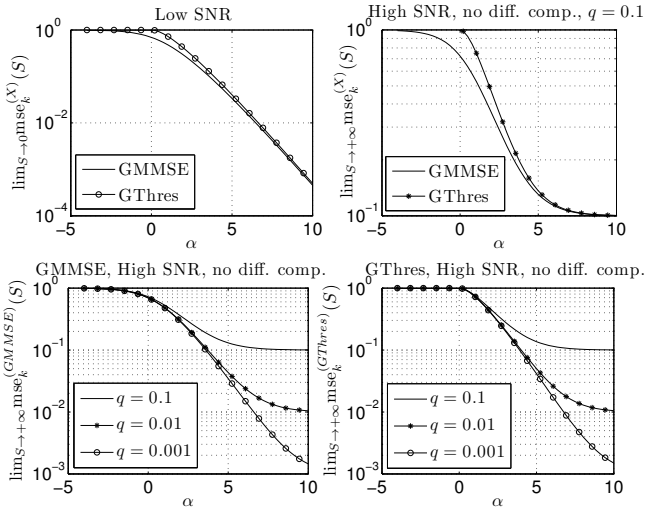

Fig. 1. High and Low asymptotic SNR behavior of the GMMSE and GThres estimators as a function of $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)$.

approaches the channel energy floor. Therefore, the larger α (alternatively, the smaller $\tilde{q}$), the smaller the weight given to the LS estimates of the sparse coefficients in (3) with respect to the MMSE estimates of the diffuse coefficients, and the better the estimation accuracy. In the limit $\alpha \to +\infty$ (*i.e.*, $\tilde{q} \to 0^+$), the GMMSE and GThres estimators treat the channel as being purely diffuse, hence the MMSE estimate of the channel is forced to zero and the MSE approaches the channel energy floor.

We conclude that, in the asymptotic SNR regimes, using $\alpha > \ln \frac{1-q}{q}$ (*i.e.*, $\tilde{q} < q$) improves the performance of the GMMSE and GThres estimators compared to assuming the true sparsity prior q. Hence, it is beneficial to use a conservative approach, *i.e.*, to assume the sparse component to be sparser than it actually is. However, this behavior does not always hold for medium SNR, where in fact a larger α (*i.e.*, a smaller $\tilde{q}$) may induce a larger MSE. This behavior can be seen by studying the two extreme cases $\alpha \to -\infty$ and $\alpha \to +\infty$, *i.e.*, $\tilde{q} \to 1$ and $\tilde{q} \to 0$, respectively. In the first case ($\alpha \to -\infty$, $\tilde{q} \to 1$), the two estimators are equivalent to LS, yielding the same MSE accuracy as LS. Conversely, when $\alpha \to +\infty$ (*i.e.*, $\tilde{q} \to 0^+$), the channel is treated as being diffuse only and is estimated via MMSE. The MSE in this case is given by

$$\begin{aligned}\mathrm{MSE}^{(\mathrm{Diff})}(\mathbf{S}) &= \frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\left[\left|\frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\mathbf{h}_{LS}(k)-\mathbf{h}(k)\right|^2\right] \\ &= \frac{1}{L}\sum_{k=0}^{L-1}\left(q\frac{\mathcal{P}_s(k)}{\left(1+\mathbf{S}_{k,k}\mathcal{P}_d(k)\right)^2}+\frac{\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\right), \quad (24)\end{aligned}$$

which performs worse than LS, for any value of the SNR matrix $\mathbf{S}$, for sufficiently large values of $\mathcal{P}_s(k), k = 0, \ldots, L-1$. Hence, in medium SNR we expect a trade-off between large values of α (*i.e.*, small values of $\tilde{q}$), which induce sparsity in the estimate of the sparse component, and small values of α, which, on the other hand, induce a less sparse solution and privilege the diffuse channel component.

It is worth noting that the MMSE estimator of the channel, which assumes perfect knowledge of q, $\mathbf{\Lambda}_s$ and $\mathbf{\Lambda}_d$, minimizes the MSE when the true sparsity level $\tilde{q} = q$ is employed. We conclude that the uncertainty about the sparse coefficients, which are treated as deterministic and unknown under the GMMSE and GThres estimators, is compensated by employing a conservative approach in the estimation of the sparse component.

Finally, for a given value of α, the GMMSE estimator achieves a better MSE accuracy than the GThres estimator, in the asymptotic regimes. In fact, the MMSE estimate of $\mathbf{a}_s(k)$ (4), *i.e.*, the posterior probability of $\mathbf{a}_s(k) = 1$, incorporates also the reliability associated with an active sparse component, and therefore, the closer the estimate to one, the more likely an active sparse component. On the other hand, the MAP estimate of $\mathbf{a}_s(k)$, by allowing only the two extreme values of $\hat{\mathbf{a}}_s(k) \in \{0, 1\}$, completely discards the reliability associated with these estimates, thus incurring a performance degradation.

V. Orthogonality vs non-Orthogonality of the pilot sequence

Thus far, we have assumed an orthogonal pilot sequence, which results in the optimality of per-tap estimation approaches versus joint estimation methods. In this section, we consider the non-orthogonal pilot scenario. We follow two approaches. In Section V-A, we examine the impact of using an estimator designed under the assumption of an orthogonal pilot sequence on received signals where the pilots are in fact non-orthogonal. We show that, from an MSE perspective, the effect of this mismatch can be characterized via an effective SNR loss. In Section V-B, we establish a connection between the GThres estimator and the classical sparse approximation algorithms [22], [24]–[26].

A. GMMSE *and* GThres *estimators with non-orthogonal pilot sequence*

Note that in the non-orthogonal case the SNR matrix $\mathbf{S}$ is non-diagonal. In this case, the observation model associated with the kth delay bin is given by $\mathbf{h}_{\mathrm{LS}}(k) = \mathbf{h}(k) + \left[\sqrt{\mathbf{S}}^{-1}\mathbf{n}\right]_k$, where the noise term $\left[\sqrt{\mathbf{S}}^{-1}\mathbf{n}\right]_k \sim \mathcal{CN}\left(0, \left[\mathbf{S}^{-1}\right]_{k,k}\right)$. Since the GMMSE and GThres estimators, designed under the assumption of orthogonal pilot sequence, operate on a per-tap basis, the non-orthogonal case is obtained by replacing $\mathbf{S}_{k,k}$ with $1/\left[\mathbf{S}^{-1}\right]_{k,k}$ in (3) and (4).

We now evaluate the MSE performance loss induced by a non-orthogonal pilot sequence. Let $\mathbf{X}$ be the corresponding Toeplitz matrix. Then, the SNR matrix $\mathbf{S} = \frac{\mathbf{X}^*\mathbf{X}}{\sigma_w^2}$ has some non-zero off-diagonal elements. The effective SNR at the kth delay bin is $S_k^{(\mathrm{NO})} \triangleq 1/\left[\mathbf{S}^{-1}\right]_{k,k}$. Therefore, using (5) and (6), in the non-orthogonal case we have, for $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$,

$$\mathrm{MSE}^{(X)}(\mathbf{S}) = \sum_{k=0}^{L-1}\mathrm{MSE}_k^{(X)}\left(1/\left[\mathbf{S}^{-1}\right]_{k,k}\right). \quad (25)$$

Now, consider a second scenario where the pilot sequence is orthogonal. Letting $\tilde{\mathbf{X}}$ be the associated Toeplitz matrix, and assuming that the pilot sequence has the same energy budget as in the non-orthogonal case, we have the SNR matrix $\tilde{\mathbf{S}} = \mathrm{diag}(\mathbf{S})$, where $\mathrm{diag}(\mathbf{B})$ is a diagonal matrix with the same diagonal elements as $\mathbf{B}$. The SNR at the kth delay bin is $S_k^{(\mathrm{O})} \triangleq \tilde{\mathbf{S}}_{k,k} = \mathbf{S}_{k,k}$, and the resulting MSE is given by

$$\mathrm{MSE}^{(X)}\left(\tilde{\mathbf{S}}\right) = \sum_{k=0}^{L-1}\mathrm{MSE}_k^{(X)}\left(\mathbf{S}_{k,k}\right). \quad (26)$$

We now prove that the effective SNRs in the non-orthogonal and orthogonal cases satisfy $S_k^{(\mathrm{O})} \geq S_k^{(\mathrm{NO})}$, $\forall k$. We can rewrite $\mathbf{S}$ as

$$\mathbf{S} = \mathbf{U}\begin{bmatrix} S_k^{(\mathrm{O})} & \mathbf{b} \\ \mathbf{b}^* & \Delta \end{bmatrix}\mathbf{U}^*, \quad (27)$$

for a proper $\Delta \succ 0$, row vector $\mathbf{b}$, and permutation matrix $\mathbf{U}$, where we have used the fact that $\mathbf{S}_{k,k} = S_k^{(\mathrm{O})}$. Then, from the inversion formula for 2×2 block-matrices, we have

$$S_k^{(\mathrm{NO})} = \frac{1}{\left[\mathbf{S}^{-1}\right]_{k,k}} = \left[\mathbf{U}^*\mathbf{S}^{-1}\mathbf{U}\right]_{1,1}^{-1} = S_k^{(\mathrm{O})} - \mathbf{b}\Delta^{-1}\mathbf{b}^*.$$

Finally, since $\Delta \succ 0$, we obtain $\mathbf{b}\Delta^{-1}\mathbf{b}^* \geq 0$ (with equality if and only if $\mathbf{b} = \mathbf{0}$), which proves the inequality $S_k^{(\mathrm{O})} \geq S_k^{(\mathrm{NO})}$, $\forall k$. Therefore, imperfect orthogonality of the pilot sequence yields a decrease of the effective SNR experienced on each channel delay bin, thus impairing the estimation performance.

We can quantify the loss in the estimation accuracy in the high and low SNR regimes where, as shown in Section IV, for the GMMSE and GThres estimators we have $\lim_{S\to 0(+\infty)} S\mathrm{MSE}_k^{(X)}(S) = \text{constant} > 0$, for a proper constant, as given in Table I. To this end, we define the *orthogonality coefficient* of the pilot sequence associated with the kth delay bin as the ratio between the effective SNR experienced in the non-orthogonal case and the SNR experienced in the orthogonal case, under the same pilot energy budget, *i.e.*,

$$\eta_k = \frac{S_k^{(\mathrm{NO})}}{S_k^{(\mathrm{O})}} = \frac{1}{[\mathbf{S}^{-1}]_{k,k}\,\mathbf{S}_{k,k}} \leq 1. \tag{28}$$

Then, in the high and low SNR regimes, the ratio between the MSE in the orthogonal case and the MSE in the non-orthogonal case, in the kth channel bin, is given by

$$\frac{\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{O})}\right)}{\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{NO})}\right)} = \frac{S_k^{(\mathrm{NO})}}{S_k^{(\mathrm{O})}} \times \frac{S_k^{(\mathrm{O})}\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{O})}\right)}{S_k^{(\mathrm{NO})}\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{NO})}\right)} \simeq \eta_k,$$

where we have used the fact that $\lim_{S\to 0(+\infty)} S\mathrm{MSE}_k^{(X)}(S) = \text{constant}$ and the definition (28).

B. Exploiting the non-orthogonality of the pilot sequence

We next investigate estimators designed for the non-orthogonal case, by establishing a connection between the GThres estimator and classical sparse approximation algorithms [22], [24], [25].

In particular, we show that the GThres estimator solves

$$\left\{\hat{\mathbf{c}}_s, \hat{\mathbf{a}}_s, \hat{\mathbf{h}}_d\right\} = \arg \max_{\mathbf{c}_s, \mathbf{a}_s, \mathbf{h}_d} p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d | \mathbf{c}_s\right). \tag{29}$$

We have

$$p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d | \mathbf{c}_s\right) = p\left(\mathbf{h}_{\mathrm{LS}} | \mathbf{a}_s, \mathbf{h}_d, \mathbf{c}_s\right) p\left(\mathbf{a}_s\right) p\left(\mathbf{h}_d\right), \tag{30}$$

where

$$\mathbf{h}_{\mathrm{LS}} | \left\{\mathbf{a}_s, \mathbf{h}_d, \mathbf{c}_s\right\} \sim \mathcal{CN}\left(\mathbf{a}_s \odot \mathbf{c}_s + \mathbf{h}_d, \mathbf{S}^{-1}\right), \tag{31}$$

$$p\left(\mathbf{a}_s\right) = \left(\frac{q}{1-q}\right)^{\|\mathbf{a}_s\|_0} (1-q)^L = \left(\frac{q}{1-q}\right)^{\|\mathbf{h}_s\|_0} (1-q)^L,$$

$$\mathbf{h}_d \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Lambda}_d\right),$$

where $\|\mathbf{x}\|_0$ is the $\mathcal{L}_0$-norm of vector $\mathbf{x}$, and $\mathbf{h}_s = \mathbf{a}_s \odot \mathbf{c}_s$ is the sparse component.

Then, from (29) and (31), we have

$$\begin{aligned}\left\{\hat{\mathbf{c}}_s, \hat{\mathbf{a}}_s, \hat{\mathbf{h}}_d\right\} &= \arg\max_{\mathbf{c}_s, \mathbf{a}_s, \mathbf{h}_d} \ln p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d | \mathbf{c}_s\right) \\ &= \arg\min_{\mathbf{h}_s = \mathbf{a}_s \odot \mathbf{c}_s, \mathbf{h}_d} \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s - \mathbf{h}_d\right)^* \mathbf{S}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s - \mathbf{h}_s\right) \\ &\qquad + \alpha \|\mathbf{h}_s\|_0 + \mathbf{h}_d^* \mathbf{\Lambda}_d^{-1} \mathbf{h}_d,\end{aligned} \tag{32}$$

where $\alpha = \ln\left(\frac{1-q}{q}\right)$. This can be viewed as an LS regression problem, with a $\mathcal{L}_0$ regularization term associated with $\mathbf{h}_s$, enforcing sparseness of the solution, and a $\mathcal{L}_2$ regularization term associated with $\mathbf{h}_d$, enforcing its Gaussian nature.

Solving with respect to $\mathbf{h}_d$ first, as a function of $\mathbf{h}_s$, we have

$$\hat{\mathbf{h}}_d\left(\mathbf{h}_s\right) = \mathbf{\Lambda}_d\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right), \tag{33}$$

and substituting this solution into the cost function, we obtain the following optimization problem for the sparse component:

$$\hat{\mathbf{h}}_s = \hat{\mathbf{a}}_s \odot \hat{\mathbf{c}}_s = \tag{34}$$
$$\arg\min_{\mathbf{h}_s} \alpha \|\mathbf{h}_s\|_0 + \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right)^* \left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right).$$

In the orthogonal pilot case, the SNR matrix $\mathbf{S}$ is diagonal and the optimization problem (34) factorizes into L separate problems, one for each channel delay bin, yielding the same solution as the GThres estimator (4). Conversely, in the non-orthogonal case, the optimal solution requires a combinatorial search over the 2^L realizations of $\mathbf{a}_s$. This is circumvented by the use of sparse approximation algorithms [22], [27].

An equivalent problem has been addressed in [22], namely

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}\in\mathbb{C}^L} \|\mathbf{w} - \Phi\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_0, \tag{35}$$

where $\mathbf{w}$ is a noisy version of $\Phi\mathbf{z}$, and Φ is known, with $\mathbf{I}_L - \Phi^*\Phi \succ 0$. Equation (34) is equivalent to (35) by letting $\mathbf{w} = \sqrt{\rho}\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-\frac{1}{2}} \mathbf{h}_{\mathrm{LS}}$, $\Phi = \sqrt{\rho}\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-\frac{1}{2}}$, $\lambda = \rho\alpha$, and $\mathbf{z} = \mathbf{h}_s$, where $\rho > 0$ is chosen so as to guarantee $\mathbf{I}_L - \Phi^*\Phi \succ 0$. The *Iterative Thresholding Algorithm* proposed in [22] may then be used to estimate $\mathbf{h}_s$, and equation (33) to estimate the diffuse component $\mathbf{h}_d$.

Alternatively, in [24], [25] the $\mathcal{L}_0$ cost associated with $\mathbf{h}_s$ is relaxed and the $\mathcal{L}_1$ regularization norm is used instead, thus yielding the convex problem

$$\hat{\mathbf{h}}_s = \arg\min_{\mathbf{h}_s} \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right)^* \left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1} \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right) + \alpha \|\mathbf{h}_s\|_1,$$

where we define the $\mathcal{L}_1$-norm $\|\mathbf{h}_s\|_1 = \sum_k |\mathbf{h}_s(k)|$.

As justified by the MSE analysis (Section IV), a conservative $\tilde{q} < q$ may be assumed in the estimation of the sparse component, by using $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right) > \ln\left(\frac{1-q}{q}\right)$.

VI. Simulation Results

In this section, we present the simulation results, and evaluate the performance achievable with the above estimation strategies, from both an MSE (as defined in (5)) and a BER perspectives. In particular, we use two different approaches. In Section VI-A, we evaluate the performance of the GMMSE and GThres estimators in a system whose channel perfectly follows the HSD model, and compare it with the asymptotic MSE behavior derived in Section IV. Moreover, we evaluate the performance of the estimators under a non-orthogonal pilot sequence, as discussed in Section V. In Section VI-B, we evaluate the BER and MSE performance of an OFDM-UWB system in a more realistic UWB channel emulator developed in [18], which we refer to as *K&P model* in the following. This approach is important as a validation of the HSD model, of the

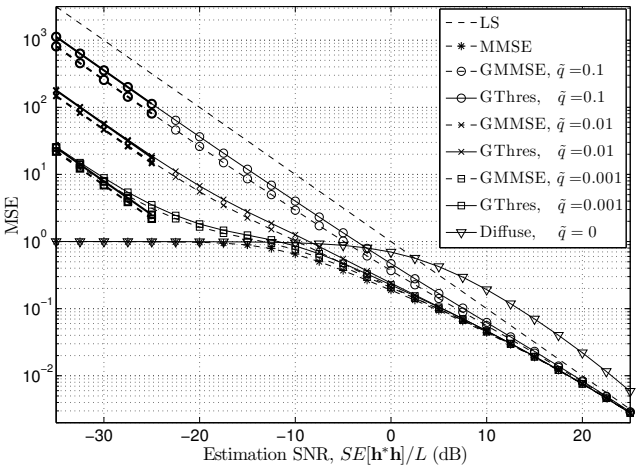

Fig. 2. MSE of the GMMSE and GThres estimators, for the HSD channel model. The bold lines with the corresponding markers represent the low SNR MSE behavior. The high SNR behavior is given by the LS estimate. $\beta = 0.01$, $q = 0.1$.

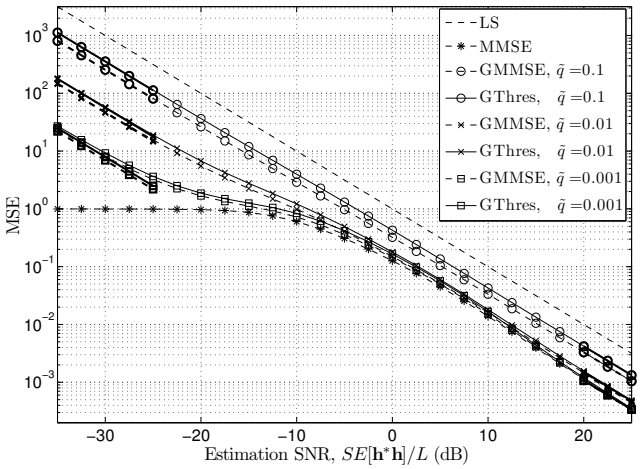

Fig. 3. MSE of the GMMSE and GThres estimators, for the HSD channel model. The bold lines with the corresponding markers represent the high/low SNR MSE behavior. $\beta = 0$ (no diffuse component), $q = 0.1$.

GMMSE and GThres estimators and of the analysis we have developed. We argue that the K&P model is more suitable than the model in [19] to evaluate the robustness and sensitivity of the proposed HSD channel estimation strategies to deviations from the HSD model. In fact, as explained in more detail in Section VI-B, K&P models the diffuse component as a diffuse tail associated with each specular component, whereas in the HSD model the diffuse and sparse components are assumed to be independent. Therefore, it represents a deviation from the HSD model. In contrast, the model developed in [19] exhibits a better fit to the HSD model, since the diffuse component is generated independently of the specular MPCs arrivals.

Unless otherwise stated, we assume an orthogonal pilot sequence, so that $\mathbf{S}$ is diagonal. For simplicity, we assume that $\mathbf{S} = S \cdot \mathbf{I}_L$, for some $S > 0$, so that we can rewrite the observation model (2) as

$$\mathbf{h}_{\text{LS}} = \mathbf{h} + \sqrt{S}^{-1}\mathbf{n}. \tag{36}$$

Moreover, we define the *estimation SNR* as the average estimation SNR per channel entry, $S\mathbb{E}[\mathbf{h}^*\mathbf{h}]/L$.

A. Hybrid Sparse/Diffuse channel model

In this section, we compare the MSE accuracy of the GMMSE and GThres estimators with the asymptotic behavior derived in Section IV, for a channel drawn according to the HSD model developed in Part I. We refer the interested reader to Part I for a more comprehensive MSE and BER analysis.

For the simulation results, we generate a channel $\mathbf{h} \in \mathbb{C}^L$ with delay spread $L = 100$. The sparsity pattern $\mathbf{a}_s \sim \mathcal{B}(q)^L$, with parameter $q = 0.1$. The vector $\mathbf{c}_s \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_s)$, where the covariance matrix $\mathbf{\Lambda}_s$ is diagonal, with exponential PDP $\mathbf{\Lambda}_s(k,k) = \mathcal{P}_s(k) = P_s e^{-\omega k}$, and $\omega = 0.05$. The diffuse component $\mathbf{h}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_d)$, where the covariance matrix $\mathbf{\Lambda}_d$ is diagonal, with exponential PDP $\mathbf{\Lambda}_d(k,k) = \mathcal{P}_d(k) = \beta P_s e^{-\omega k}$. The parameter $P_s > 0$ is a normalization factor, and is chosen so that the average channel energy is L, *i.e.*, $\sum_{k=0}^{L-1} \mathbb{E}\left[|\mathbf{h}(k)|^2\right] = P_s \sum_{k=0}^{L-1} (\beta + q) e^{-\omega k} = L$.

We compare the GMMSE and GThres estimators, for different values of the assumed sparsity level $\tilde{q} \in \{0.1, 0.01, 0.001\}$ (*i.e.*, $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right) \in \{2.2, 4.6, 6.9\}$), with the asymptotic MSE behavior derived in Section IV. For performance comparison, we also consider the unstructured LS estimator and the MMSE estimator [20], which assumes perfect knowledge of q, $\mathbf{\Lambda}_d$ and $\mathbf{\Lambda}_s$, and thus performs an MMSE estimate of the channel. Note that the latter represents a lower bound to the estimation accuracy. This is primarily used as a reference.

In Figure 2, we plot the MSE of the estimators as a function of the estimation SNR, and their asymptotic MSE behavior (bold lines, with the corresponding markers for the different values of α), for the case where the diffuse component is present with relative power $\beta = 0.01$. We note that there is a perfect match between the MSE in the high and low SNR regimes, and the asymptotic analysis developed in Section IV. In particular, from an MSE perspective, it is confirmed that it is beneficial to use a conservative approach in the estimation process, *i.e.*, by assuming the sparse component to be sparser than it actually is. Also, as predicted by the MSE analysis, the GMMSE estimator outperforms the GThres estimator, in the asymptotic regimes.

Notice that, for the set of values of $\tilde{q}$ considered in these previous figures, the optimality of a conservative approach holds also in the medium SNR range. However, this is not always true, as we have discussed in Section IV and as we can observe for the "only diffuse" estimator, which ignores the sparse component. This can be interpreted as the limiting estimator for GMMSE and GThres when $\alpha \to +\infty$ ($\tilde{q} \to 0$).

In Figure 3, we plot the MSE of the estimators as a function of the SNR S, for the case with no diffuse component, $\beta = 0$. Even in this case, we notice a perfect match between the MSE in the high and low SNR regimes, and the asymptotic analysis in Section IV. In particular, the larger the factor α used (the smaller $\tilde{q}$), the better the estimation accuracy. Unlike Figure 2, where the MSE approaches the LS estimate for high SNR, in this case we note a performance improvement.

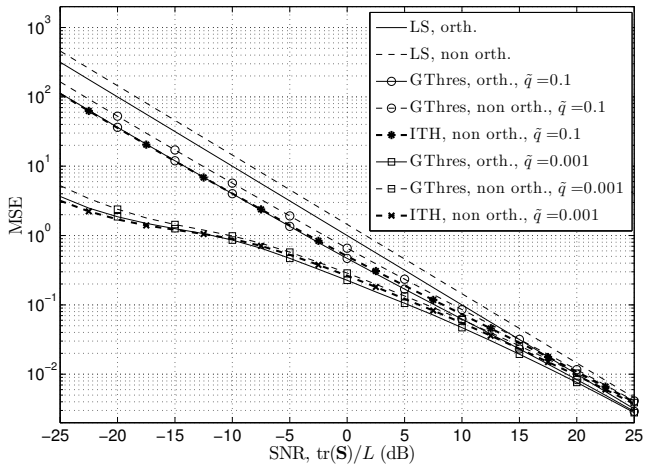

Fig. 4. Comparison between the non-orthogonal and orthogonal pilot sequence cases. $\beta = 0.01$, $q = 0.1$.

In fact, when $\beta = 0$, the estimate of $\mathbf{h}_d$ is forced to zero. Therefore, whenever the GThres estimator correctly detects $\hat{\mathbf{a}}_s(k) = \mathbf{a}_s(k) = 0$, the channel component $\mathbf{h}(k)$ is estimated with no error. On the other hand, when $\beta > 0$, a residual MMSE estimation error is incurred.

In Figure 4, we compare the MSE of the GThres estimator for the non-orthogonal and orthogonal pilot sequence cases, under the same pilot energy budget, as discussed in Section V-A. Moreover, we plot the curves associated with the modified *Iterative Thresholding Algorithm* (ITH), designed in Section V-B based on a variation of [22] which takes into account the presence of the diffuse component. The non-orthogonal pilot sequence is generated from a CAZAC sequence of length $M = 50 = L/2$ [28]. As expected, we observe a performance loss in the non-orthogonal case, compared to the orthogonal pilot scenario with the same pilot energy budget. In fact, the GThres estimator, by employing a per-tap estimation approach, neglects any correlation among the channel taps, thus incurring a performance degradation. We measured that the *orthogonality coefficient* (28) ranges in the interval $\eta_k \in [0.625, 0.765]$ (note that this is a function of the delay $k \in \{0, \ldots, L-1\}$), corresponding to an SNR loss in the range $[1.16, 2.05]$ dB. These values are confirmed by simulation, where the SNR loss induced by GThres under a non-orthogonal pilot sequence (by averaging over all channel delay taps, as in (5)) is approximately $[1.5, 2]$ dB. Interestingly, the performance degradation incurred by the GThres estimator is partially recovered (fully, in the low SNR regime) by the ITH algorithm, which exploits the correlation introduced by the non-orthogonal pilot sequence by estimating the channel taps jointly.

B. Realistic UWB channel model

We now evaluate and compare the MSE and the BER performance of the GMMSE and GThres estimators in a more realistic UWB channel, which does not follow the HSD model. This allows us to evaluate how sensitive the estimation performance is to the type of channel we are operating in.

We evaluate the BER performance induced by channel estimation errors in an OFDM-UWB system, with $N_{\rm dft} = 2048$ sub-carriers, 4-QAM constellation and transmission bandwidth B. In the estimation phase, we use an orthogonal pilot sequence. This may be achieved, for example, by allocating an OFDM symbol with a constant modulus pilot sequence. Since we want to evaluate the impact of channel estimation errors on the BER performance, we consider a noise-free setting, *i.e.*, no noise is added to the information symbols, whereas noise is added in the estimation phase, so as to induce channel estimation errors. In particular, let $X(n)$ be the 4-QAM symbol transmitted on the nth sub-carrier, and $H(n) = \sum_{l=0}^{L-1} \mathbf{h}(n) e^{-i2\pi \frac{ln}{N_{\rm dft}}}$ be the $N_{\rm dft}$-points DFT of the channel, where $n \in \{0, \ldots, N_{\rm dft} - 1\}$. Then, the received symbol is $Y(n) = H(n)X(n)$. This is equalized by using the estimate $\hat{H}(n)$ of $H(n)$, *i.e.*, $\tilde{X}(n) = \frac{\hat{H}(n)}{\hat{H}(n)} X(n)$, and the decision is based on a minimum distance criterion, *i.e.*, $\hat{X}(n) = \min_{x \in 4-\rm QAM} |\tilde{X}(n) - x|^2$.

We use the K&P model [18], which is suited to indoor environments. This model combines both a geometric approach for the resolvable individual specular components (echoes), arising from reflections from the scatterers in the environment, and a statistical approach for the dense multipath clusters associated with each echo. The model also includes a frequency dependent gain decay, so that the overall channel transfer function is expressed as

$$H(f) = \sum_l A_l(\tau_l) \left(1 + \mathcal{D}_l(f)\right) e^{-i2\pi f \tau_l} \times \\ \times \left(1 + \frac{f}{f_0}\right)^{-\nu} \mathcal{I}\left(|f| \leq \frac{B}{2}\right), \qquad (37)$$

The sum is over the individual echoes, with the lth echo having amplitude $A_l(\tau_l)$ and delay τ_l. $\mathcal{D}_l(f)$ is the multipath cluster associated with the lth echo, with exponential PDP and circularly symmetric Gaussian distribution in the time-domain, ν is the frequency domain decay exponent, f_0 is the center frequency, and $B < R$ is the transmission bandwidth.

The time-domain baseband representation of the channel is obtained by performing an inverse Fourier transform of (37), and by sampling at rate R samples per ns. We further clip the channel in the delay domain, so that only the channel window carrying most of the energy is kept. This step determines the delay spread of the channel ($L = 600$). The channel snapshot is finally normalized to have energy L, *i.e.*, $\sum_{l=0}^{L-1} |\mathbf{h}(l)|^2 = L$.

It is worth noting that τ_l is quantized to discrete values, and equals an integer number of the sampling interval R^{-1} ns. This is a simplification, which guarantees that the MPC arrival matches exactly the sampling period. Therefore, in general, the K&P model [18] does not cope with the side-lobes of the sinc pulse, which arise when the MPCs arrive in the middle of two sampling times. However, the bandwidth limitation $B < R$ introduces a sinc filtering of the channel, so that the side-lobes of the sinc pulse affect the channel impulse response.

We choose the *Office LOS* scenario in [18] for our simulations. A summary of the main parameters are given in Table II. In particular, 10000 channel snapshot are generated, each corresponding to a different position of the mobile receiver

TABLE II
MAIN PARAMETERS FOR THE *Office LOS* SCENARIO IN [18]

N_{dft}	2048	Number of channel samples in the delay domain
R	$12.8\,\text{ns}^{-1}$	Sampling rate in the delay domain
B	10 GHz	Bandwidth of the UWB system
f_0	6 GHz	Center frequency
d_0	0.8 m	Reference distance for individual echo power law
δ	3	Path loss exponent for individual echo power law
G_{MP}	-20 dB	Cluster gain with respect to associated individual echo
G_{MP-LOS}	-13 dB	Additional cluster gain for LOS individual echo
γ	10 ns	Multipath cluster exponential decay parameter
ν	1.1	Frequency domain decay exponent
(x_t, y_t, z_t)	$(1.78, 4, 1.5)$ m	Coordinates of transmitter position
(x_r, y_r, z_r)	from $(3.0, 1.5, 1.5)$ m to $(4.5, 1.8, 1.5)$ m	Coordinates of mobile receiver position

along the line connecting the points $(3.0, 1.5, 1.5)\,\text{m}$ and $(4.5, 1.8, 1.5)\,\text{m}$ (where $(x, y, z)\,\text{m}$ represents a point in the three dimensional space with coordinates x, y and z, measured in m relative to the origin). For each position of the mobile receiver, the arrival pattern of the resolvable MPCs, *i.e.*, their gain and delay, is determined by the relative positions of the mobile receiver, transmitter and scatterers (these are positioned on a grid in the three dimensional space). Moreover, for each channel snapshot, we generate an independent realization of the diffuse component (Rayleigh fading) and of the additive noise. We refer the interested reader to [18] for further details.

It is worth noting that the sparsity level q of the HSD model is not defined for the K&P model. This parameter may be roughly estimated as the ratio between the number of active scatterers and the delay spread L. For the *Office LOS* scenario defined in [18], we have 6 (virtual) scatterers and $L = 600$, which gives $\hat{q} \simeq 0.01$. Moreover, the PDP estimator developed in Part I for the HSD model assumes an exponential PDP for the diffuse component, which is not defined for the K&P model. In Figure 5, we plot the PDP of a channel snapshot as an example, and the exponential PDP fitting, estimated using the EM algorithm developed in Part I [20]. We note a good fitting of the exponential PDP model to the PDP of the channel realization. Remarkably, although the K&P model

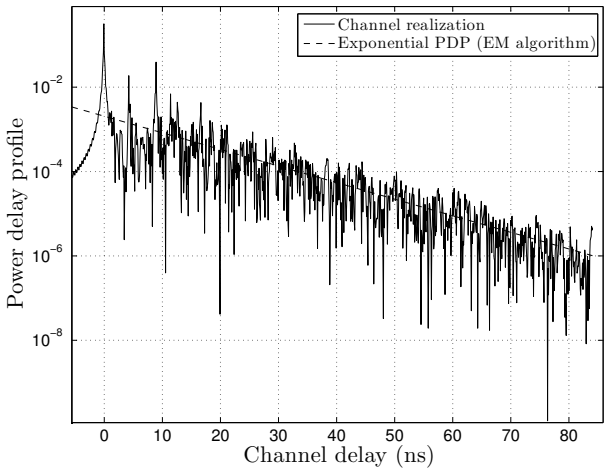

Fig. 5. PDP of one realization of the K&P model (with parameters given in Table II) and exponential PDP, estimated using the EM algorithm developed in Part I [20].

defines the diffuse component as a diffuse tail associated with each specular component, the overall effect, by summing the contribution from all MPCs, is that of a unique PDP tail, which fits well the exponential shape.

The channel and the PDP of the diffuse component are estimated based on a single snapshot of the channel. In particular, the PDP of the diffuse component is estimated using the EM algorithm developed in Part I. Hence, the MSE and BER results are not affected by the structure of the spatio-temporal correlation of the channel.

Figures 6 and 7 plot the MSE of the GMMSE, GThres and purely sparse and diffuse estimators, for different values of the assumed sparsity level $\tilde{q}$. Since a per-tap approach is optimal in this case, for the sparse estimator we choose a variation of the GThres estimator, which assumes no diffuse component ($\mathbf{h}_d = \mathbf{0}$). The diffuse estimator assumes a purely diffuse channel, and performs a linear MMSE estimate based on the estimated PDP of the diffuse component.

In Figure 6, we observe that, the smaller $\tilde{q}$ (*i.e.*, the larger α), the better the estimation accuracy of the GMMSE and GThres estimators. Moreover, the GMMSE estimator outperforms the GThres estimator, for a given value of $\tilde{q}$. This is the same behavior, predicted by the MSE analysis in Section IV, that we have observed in the case where the channel follows the HSD model (Figure 2). Remarkably, we notice a perfect match between the simulation results and the low/high SNR asymptotic behavior of the estimators (bold lines). This is a surprising result, if we consider that the K&P channel emulator deviates from the HSD model, and the PDP of the diffuse component is unknown and estimated from the data. However, note that the value of the channel delay spread, $L = 600$, allows sufficient averaging over the small scale fading in the delay dimension, so that the PDP is accurately estimated.

Moreover, we notice that the diffuse estimator outperforms the HSD estimators in the low SNR ($< -12.5\,\text{dB}$). This is an expected result, which is coherent with the simulation results based on the HSD model (Figure 2) and with the asymptotic analysis in Section IV, where we have proved that, in the low SNR, the smaller $\tilde{q}$, the better the estimation accuracy (note that the diffuse estimator corresponds to the limit case $\tilde{q} \to 0$). In fact, the diffuse estimator forces the channel estimate to zero in the low SNR, thus approaching the channel energy floor. Conversely, a performance degradation is observed for higher SNR values, with respect to the HSD estimators with

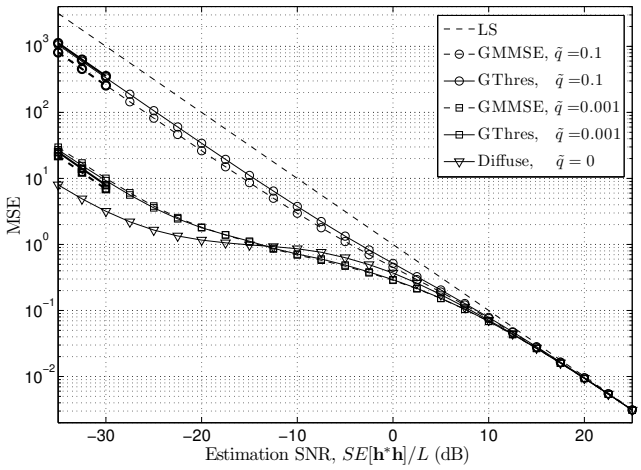

Fig. 6. MSE of the GMMSE and GThres estimators, for the K&P channel model, LOS-Office scenario. The bold lines with the corresponding markers represent the low SNR MSE behavior. The high SNR behavior is given by the LS estimate.

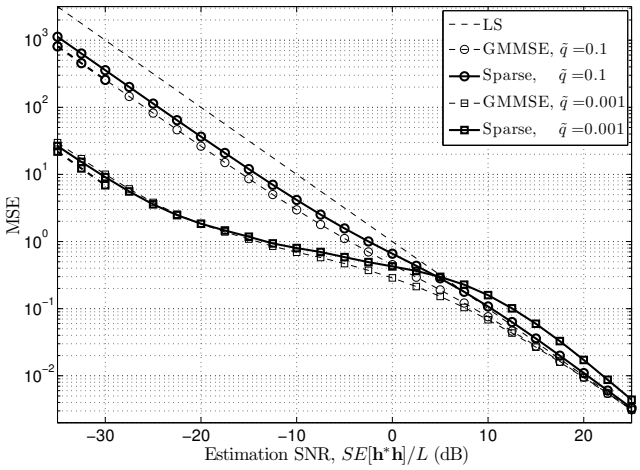

Fig. 7. MSE of the GMMSE and Sparse estimators, for the K&P channel model, LOS-Office scenario.

$\tilde{q} = 0.001$, which achieve the best performance.

In Figure 7, we notice that the GMMSE estimator achieves better performance than the sparse estimator, for the same values of $\tilde{q}$. In fact, the sparse estimator does not effectively capture the diffuse component of the channel, thus incurring a performance degradation, mainly in medium and high SNR (in particular, in the high SNR range ($> 5\,\mathrm{dB}$), it performs even worse than LS).

Figure 8 plots the BER associated with the GMMSE, LS, purely sparse and diffuse estimators, for different values of $\tilde{q}$. The SNR is referred to the output of an ideal Rake receiver with perfect channel knowledge, where the estimation noise is treated as additive Gaussian noise at the receiver. This is defined as $\mathrm{SNR}_{\mathrm{rake}} = S\mathbf{h}^*\mathbf{h}$. A 4-QAM constellation is employed with Gray mapping, and the bit sequence is uncoded. Note that some sub-carriers may have a very small channel amplitude, thus inducing high BER. Therefore, in Figure 8 we plot the BER averaged over only the "good" sub-carriers, which are chosen based on the heuristic carrier selection scheme

$$\left\{k : |H(k)|^2 \geq \lambda \max_n |H(n)|^2\right\}, \tag{38}$$

where $\lambda \in (0, 1)$ is a threshold value. In particular, we choose λ such that 30% of the sub-carriers are classified as "good". The rationale behind this choice is that, in a practical system, the "bad" sub-carriers would never be used, since they are not suitable to carry information.

Generally, we observe that the better the MSE estimation accuracy, the smaller the BER. In particular, the best performance is achieved by the GMMSE estimator with $\tilde{q} = 0.001$. Moreover, similarly to the MSE, also the BER benefits from a conservative approach in the estimation of the sparse component, *i.e.*, it is beneficial to use small values of $\tilde{q}$. We notice that a poor BER performance is incurred by the purely sparse estimator which, in the high SNR, performs even worse than LS. Similarly, the diffuse estimator performs worse than GMMSE with $\tilde{q} = 0.001$ in the medium SNR range. As in the MSE case, the purely sparse and diffuse estimators are unable to exploit both the diffuse and sparse components of the channel jointly, thus incurring a performance degradation. Finally, we observe an irregular behavior of the GMMSE and sparse estimators with $\tilde{q} = 0.001$ around $18\,\mathrm{dB}$ SNR. We argue that this is a consequence of the fact that we do not average over independent realizations of the surrounding environment, *i.e.*, we use the particular *Office LOS* in [18], which specifies the relative positions of the scatterers, and of the transmitter/receiver pair as well.

These results show that the GMMSE and GThres estimators effectively capture the main UWB propagation phenomena, *e.g.*, the resolvable MPCs of the channel, modeled by a sparse component, unresolvable MPCs, scattering from rough surfaces and frequency dispersion, which are better modeled by a diffuse component. Also, we observe that a small performance degradation is incurred by the diffuse estimator. However, we argue that one of the strengths of the proposed HSD model and channel estimation strategies relies in their robustness and adaptability to different scenarios of interest, where the channel exhibits a sparse, diffuse or hybrid nature. Conversely, a diffuse (respectively, sparse) estimator is expected to perform poorly in sparse (diffuse) channels.

VII. Conclusions

In this paper, we have further investigated the Hybrid Sparse/Diffuse (HSD) model developed in Part I of this paper [20]. Specifically, we have carried out an MSE analysis of the GMMSE and GThres estimators, for the scenario where the statistics of the sparse component are unknown at the receiver, in the asymptotic regimes of high and low SNR. Simulation results are provided, for a channel following the HSD model, showing that the MSE performance of these estimators agrees with the expected asymptotic behavior. This analysis suggests that it is beneficial, from an MSE perspective, to use a conservative approach in the estimation of the sparse component, *i.e.*, to assume the sparse component to be sparser than it actually is. While this result cannot be extended to medium SNR, simulation results show that a

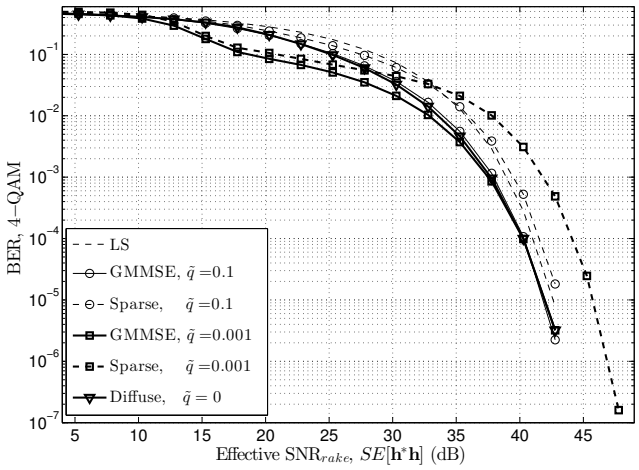

Fig. 8. BER induced by channel estimation errors, for the K&P channel model, LOS-Office scenario. The BER is averaged over the "good" subcarriers only (38).

similar behavior often holds in this regime. Moreover, we have analyzed the case with a non-orthogonal pilot sequence, and shown that the GThres estimator can be recast as a modification of a sparse approximation algorithm proposed in the literature. Finally, we have evaluated these estimation schemes using a more realistic geometry-based stochastic UWB channel emulator, developed in [18]. Simulation results for this case show that the GMMSE and GThres estimators achieve better performance, in terms of both MSE and BER, than conventional unstructured (Least Squares) and purely sparse or diffuse estimators, thus suggesting that, although simplified (*e.g.*, compared to [14]), the HSD model is able to capture key UWB propagation mechanisms, such as resolvable MPCs, diffuse scattering from rough surfaces, unresolvable MPCs, and frequency dispersion.

ACKNOWLEDGMENT

This research has been funded in part by the following grants and organizations: ONR N00014-09-1-0700, NSF CNS-0832186, NSF CNS-0821750 (MRI), Aldo Gini Foundation (Padova, Italy).

APPENDIX

Lemma 1 (Exchanging the limit with the expectation)**.** For the GMMSE and GThres estimators of the kth channel delay bin $\hat{h}\,(y)$, where $y = h + \sqrt{S}^{-1} n$ is the observation, $h = a_s c_s + \sqrt{P_d} h_d$ is the HSD channel bin, n is the noise in the kth delay bin, we have, for $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$,

$$\lim_{S \to S_{\lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S \to S_{\lim}} f^{(X)}\left(\sqrt{S}y, n\right)\right],$$

where $S_{\lim} \in \{0, +\infty\}$, and $\mathrm{mse}_k^{(X)}(S)$ and $f^{(X)}\left(\sqrt{S}y, n\right)$ are defined in (9) and (8), respectively.

Proof. Note from (3) that, for $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$, we can write

$$\hat{h}(y) = r^{(X)}\left(\sqrt{S}|y|\right) y, \tag{39}$$

where $r^{(X)}\,(z)$, for $z \geq 0$, is given by

$$r^{(X)}\,(z) = \phi^{(X)}\,(z) + \left(1 - \phi^{(X)}\,(z)\right) \frac{SP_d}{1 + SP_d}. \tag{40}$$

The function $\phi^{(X)}\,(z)$ is the estimate of the sparsity bit a_s conditioned on $|y| = \sqrt{S}^{-1} z$, and its expression depends on the chosen estimator $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$, specifically, from (4),

$$\phi^{(X)}\,(z) = \begin{cases} \frac{1}{1 + e^{\alpha} \exp\left\{-\frac{z^2}{1+SP_d}\right\}}, & X = \mathrm{GMMSE}, \\ \mathcal{I}\left(z^2 \geq \alpha\,(1 + SP_d)\right), & X = \mathrm{GThres}. \end{cases} \tag{41}$$

Let $\{S_j > 0, j = 0, \ldots, +\infty\}$ be a generic SNR sequence, indexed by j, such that $\lim_{j \to +\infty} S_j = S_{\lim}$. From Lebesgue's Dominated Convergence Theorem [29], if there exists a function $g^{(X)}(h, n)$ such that

$$\begin{cases} \left|f^{(X)}\left(\sqrt{S_j}h + n, n\right)\right| \leq g^{(X)}(h, n) \quad \text{a.e.}, \quad \forall j \\ \mathbb{E}\left[g^{(X)}(h, n)\right] < +\infty, \end{cases} \tag{42}$$

where a.e. stands for *almost everywhere*, *i.e.*, the inequality holds except on a set with probability measure zero (with respect to the random variables $h_d \sim \mathcal{CN}(0, 1)$, $n \sim \mathcal{CN}(0, 1)$, $a_s \sim \mathcal{B}(q)$ and c_s), then

$$\lim_{j \to +\infty} \mathrm{mse}_k^{(X)}(S_j) = \mathbb{E}\left[\lim_{j \to +\infty} f^{(X)}(\sqrt{S_j}h + n, n)\right].$$

If this property holds for any SNR sequence such that $\lim_{j \to +\infty} S_j = S_{\lim}$, then

$$\lim_{S \to S_{\lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S \to S_{\lim}} f^{(X)}(\sqrt{S}y, n)\right],$$

and the Lemma is proved.

We now prove the existence of such a function $g^{(X)}\,(\cdot)$. Let $x = \sqrt{S}y$. Then, from (8) and (39), we have

$$f^{(X)}(x, n) = \left|r^{(X)}(|x|)\,x - \sqrt{S}h\right|^2 = \left|\left(1 - r^{(X)}(|x|)\right)x - n\right|^2,$$

where in the last step we used the fact that $\sqrt{S}h = x - n$. Using the inequality $|A + B|^2 \leq 2|A|^2 + 2|B|^2$, we have

$$f^{(X)}(x, n) \leq 2\left(1 - r^{(X)}\,(|x|)\right)^2 |x|^2 + 2\,|n|^2. \tag{43}$$

Moreover, from (40), we have, $\forall x \in \mathbb{C}$,

$$1 - r^{(X)}\,(|x|) = \left(1 - \phi^{(X)}(|x|)\right) \frac{1}{1 + SP_d} \leq 1 - \phi^{(X)}(|x|)\,.$$

Letting $m^{(X)}(|x|) = \left(1 - \phi^{(X)}\,(|x|)\right)|x|$, we finally obtain $f^{(X)}(x, n) \leq 2m^{(X)}(|x|)^2 + 2\,|n|^2$.

In order to proceed, we distinguish between the estimators.

1) Generalized MMSE Estimator:

For the GMMSE estimator, using the expression of $\phi^{(\mathrm{GMMSE})}\,(|x|)$ in (41), we have

$$m^{(\mathrm{GMMSE})}(|x|) = \frac{e^{\alpha} \exp\left\{-|x|^2\right\}|x|}{1 + e^{\alpha} \exp\left\{-|x|^2\right\}} \leq e^{\alpha} \exp\left\{-|x|^2\right\}|x|.$$

The right hand side is maximized at $|x| = \frac{1}{\sqrt{2}}$, and therefore we obtain the bound $m^{(\mathrm{GMMSE})}(|x|) \leq e^{\alpha} \frac{1}{\sqrt{2e}}$. Then, we have the following bound on $f^{(\mathrm{GMMSE})}(x, n)$:

$$f^{(\mathrm{GMMSE})}(x, n) \leq e^{2\alpha - 1} + 2\,|n|^2 \triangleq g^{(\mathrm{GMMSE})}(h, n). \tag{44}$$

$g^{(\mathrm{GMMSE})}(h, n)$ is an integrable function, in fact $\mathbb{E}\left[g^{(\mathrm{GMMSE})}(h, n)\right] = e^{2\alpha - 1} + 2 < +\infty$.

2) Generalized Thresholding Estimator:

For the GThres estimator, using the expression of $\phi^{(\mathrm{GThres})}(|x|)$ in (41), we have $m^{(\mathrm{GThres})}(|x|) = \mathcal{I}\left(|x|^2 < \alpha\right)|x|$. For $|x| \geq \sqrt{\alpha}$, we have $m^{(\mathrm{GThres})}(|x|) = 0$. On the other hand, for $|x| < \sqrt{\alpha}$, we have $m^{(\mathrm{GThres})}(|x|) = |x| \leq \sqrt{\alpha}$. In general, $m^{(\mathrm{GThres})}(|x|) \leq \sqrt{\alpha}, \quad \forall |x| \geq 0$, and therefore

$$f^{(\mathrm{GThres})}(x,n) \leq 2\alpha + 2\,|n|^2 \triangleq g^{(\mathrm{GThres})}(h,n).$$

$g^{(\mathrm{GThres})}(h,n)$ is an integrable function, in fact we have

$$\mathbb{E}\left[g^{(\mathrm{GThres})}(h,n)\right] = 2\alpha + 2 < +\infty. \tag{45}$$

The Lemma is thus proved. □

Lemma 2. We have, for $n \in \mathcal{CN}(0,1)$,

$$\mathbb{E}\left[\frac{|n|^2}{\left(1+e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = e^{-\alpha}\ln\left(1+e^{\alpha}\right). \tag{46}$$

Proof. We have

$$\mathbb{E}\left[\frac{|n|^2}{\left(1+e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = \int_0^{+\infty} \frac{x}{\left(1+e^{\alpha-x}\right)^2} e^{-x}\mathrm{d}x \tag{47}$$

$$= \lim_{t\to+\infty} \int_0^t \frac{x}{\left(1+e^{\alpha-x}\right)^2} e^{-x}\mathrm{d}x,$$

where we have used the substitution $x = |n|^2$, and the fact that, since $n \sim \mathcal{CN}(0,1)$, $x \sim \mathcal{E}(1)$.

Let $B(x) = \frac{e^{-\alpha}}{1+e^{\alpha-x}}$ and $B'(x) \triangleq \frac{\mathrm{d}B(x)}{\mathrm{x}} = \frac{e^{-x}}{(1+e^{\alpha-x})^2}$. Then, from (47) we have

$$\mathbb{E}\left[\frac{|n|^2}{\left(1+e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = \lim_{t\to+\infty} \int_0^t xB'(x)\mathrm{d}x. \tag{48}$$

By solving the integral in the limit by parts, we have

$$\int_0^t xB'(x)\mathrm{d}x = tB(t) - \int_0^t B(x)\mathrm{d}x \tag{49}$$

$$= tB(t) - e^{-\alpha}\ln\left(e^t + e^{\alpha}\right) + e^{-\alpha}\ln\left(1+e^{\alpha}\right),$$

where in the last step we used the fact that $B(x) = e^{-\alpha}\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(e^x + e^{\alpha}\right)$. Finally, the result is straightforwardly obtained by substituting the expression above in (48), and by letting $t \to +\infty$. □

REFERENCES

[1] M. Win and R. Scholtz, "Impulse radio: how it works," *IEEE Communications Letters*, vol. 2, no. 2, pp. 36–38, Feb. 1998.

[2] R. Scholtz, "Multiple access with time-hopping impulse modulation," in *IEEE Military Communications Conference*, vol. 2, Oct. 1993, pp. 447–450.

[3] R. Qiu, H. Liu, and X. Shen, "Ultra-wideband for multiple access communications," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 80–87, Feb. 2005.

[4] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via Ultra-Wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, July 2005.

[5] M. Win and R. Scholtz, "On the robustness of ultra-wide bandwidth signals in dense multipath environments," *IEEE Communications Letters*, vol. 2, no. 2, pp. 51–53, Feb. 1998.

[6] M. Chiani and A. Giorgetti, "Coexistence Between UWB and Narrow-Band Wireless Communication Systems," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 231–254, Feb. 2009.

[7] A. Batra, J. Balakrishnan, G. Aiello, J. Foerster, and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, Sep. 2004.

[8] T. Zasowski, G. Meyer, F. Althaus, and A. Wittneben, "Propagation effects in UWB body area networks," in *IEEE International Conference on Ultra-Wideband (ICU)*, Sep. 2005, pp. 16–21.

[9] L. Yang and G. Giannakis, "Ultra-Wideband Communications: an idea whose time has come," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 26–54, Nov. 2004.

[10] A. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. Schantz, K. Siwiak, and M. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 11, pp. 3151–3166, Nov. 2006.

[11] G. Aiello and G. Rogerson, "Ultra-Wideband Wireless Systems," *IEEE Microwave Magazine*, vol. 4, no. 2, June 2003.

[12] C. Carbonelli and U. Mitra, "Clustered Channel Estimation for UWB Multiple Antenna Systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 970–981, Mar. 2007.

[13] ——, "Clustered ML Channel Estimation for Ultra-Wideband Signals," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2412–2416, July 2007.

[14] P. Schniter, "A Message-Passing Receiver for BICM-OFDM Over Unknown Clustered-Sparse Channels," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1462–1474, Dec. 2011.

[15] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[16] D. Cassioli, M. Win, and A. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1247–1257, Aug. 2002.

[17] A. Molisch, "Ultra-Wide-Band Propagation Channels," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.

[18] J. Kunisch and J. Pamp, "An ultra-wideband space-variant multipath indoor radio channel model," in *IEEE Conference on Ultra Wideband Systems and Technologies*, Nov. 2003, pp. 290–294.

[19] T. Santos, F. Tufvesson, and A. Molisch, "Modeling the Ultra-Wideband Outdoor Channel: Model Specification and Validation," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1987–97, June 2010.

[20] N. Michelusi, U. Mitra, A. Molisch, and M. Zorzi, "UWB Sparse/Diffuse Channels, Part I: Channel Models and Bayesian Estimators," *IEEE Transactions on Signal Processing*, 2012, to be published.

[21] N. Michelusi, U. Mitra, and M. Zorzi, "Hybrid Sparse/Diffuse UWB channel estimation," in *IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2011, pp. 201–205.

[22] T. Blumensath and M. Davies, "Iterative Thresholding for Sparse Approximations," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 629–654, 2008, 10.1007/s00041-008-9035-z.

[23] J. Hansen, "An analytical calculation of power delay profile and delay spread with experimental verification," *IEEE Communications Letters*, vol. 7, no. 6, pp. 257–259, June 2003.

[24] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[25] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[26] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[27] K. Herrity, A. Gilbert, and J. Tropp, "Sparse Approximation Via Iterative Thresholding," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, May 2006.

[28] Y. Wen, W. Huang, and Z. Zhang, "CAZAC sequence and its application in LTE random access," in *IEEE Information Theory Workshop, ITW*, Oct. 2006, pp. 544–547.

[29] R. G. Bartle, *The Elements of Integration and Lebesgue Measure*, 1st ed. Wiley-Interscience, Jan. 1995.

Nicolò Michelusi (S'09) received the B.S. (Electronics Engineering) and M.S. (Telecommunications Engineering) degrees summa cum laude from the University of Padova, Italy, in 2006 and 2009, respectively, and the M.S. degree in Telecommunications Engineering from Technical University of Denmark, Copenhagen, Denmark, in 2009, as part of the T.I.M.E. double degree program. Since January 2009, he is a Ph.D. student at University of Padova, Italy. In 2011, he was on leave at the University of Southern California, Los Angeles, United States, as a visiting Ph.D. student. His research interests include ultrawideband communications, wireless networks, cognitive networks, optimal control, energy harvesting for communications.

Urbashi Mitra received the B.S. and the M.S. degrees from the University of California at Berkeley and her Ph.D. from Princeton University. After a six year stint at the Ohio State University, she joined the Department of Electrical Engineering at the University of Southern California, Los Angeles, where she is currently a Professor. Dr. Mitra has been an Associate Editor for the following IEEE publications: Transactions on Information Theory (2007-2011), Journal of Oceanic Engineering (2006-2011), and Transactions on Communications (1996-2001). She was a member of the IEEE Information Theory Society's Board of Governors (2002-2007) and began a third term in 2012. Dr. Mitra is also a member of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking (2012-2014). She is the recipient of: 2012 NAE Lillian Gilbreth Lectureship, 2011 USC Zumberge Interdisciplinary Innovation Fund (with M. El-Naggar), USC Center for Excellence in Research Fellowship (2010-2013), the Viterbi School of Engineering Dean's Faculty Service Award (2009), USC Mellon Mentoring Award (2008), IEEE Fellow (2007), Texas Instruments Visiting Professor (Fall 2002, Rice University), 2001 Okawa Foundation Award, 2000 Lumley Award for Research (OSU College of Engineering), 1997 MacQuigg Award for Teaching (OSU College of Engineering), and a 1996 National Science Foundation (NSF) CAREER Award. Dr. Mitra has held visiting appointments at: the Delft University of Technology, Stanford University, Rice University, and the Eurecom Institute. She served as co-Director of the Communication Sciences Institute at the University of Southern California from 2004-2007.

Andreas F. Molisch (S'89-M'95-SM'00-F'05) received the Dipl. Ing., Ph.D., and habilitation degrees from the Technical University of Vienna, Vienna, Austria, in 1990, 1994, and 1999, respectively. He subsequently was with AT&T (Bell) Laboratories Research (USA); Lund University, Lund, Sweden, and Mitsubishi Electric Research Labs (USA). He is now a Professor of electrical engineering with the University of Southern California, Los Angeles. His current research interests are the measurement and modeling of mobile radio channels, ultra-wideband communications and localization, cooperative communications, multiple-input-multiple-output systems, and wireless systems for healthcare. He has authored, coauthored, or edited four books (among them the textbook Wireless Communications, Wiley-IEEE Press), 14 book chapters, some 140 journal papers, and numerous conference contributions, as well as more than 70 patents and 60 standards contributions. Dr. Molisch has been an Editor of a number of journals and special issues, General Chair, Technical Program Committee Chair, or Symposium Chair of multiple international conferences, as well as Chairman of various international standardization groups. He is a Fellow of the IET, an IEEE Distinguished Lecturer, and a member of the Austrian Academy of Sciences. He has received numerous awards, most recently the 2011 James Evans Avant-Garde award of the IEEE Vehicular Technology Society, the Donald Fink Prize of the IEEE, and the Eric Sumner Award of the IEEE.

Michele Zorzi (S'89, M'95, SM'98, F'07) was born in Venice, Italy, on December 6th, 1966. He received the Laurea and the PhD degrees in Electrical Engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the Academic Year 1992/93, he was on leave at the University of California, San Diego (UCSD) as a visiting PhD student, working on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the University of Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty at the Information Engineering Department of the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, energy constrained communications protocols, broadband wireless access and underwater acoustic communications and networking.

Dr. Zorzi was the Editor-In-Chief of the IEEE WIRELESS COMMUNICATIONS MAGAZINE from 2003 to 2005 and the Editor-In-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2008 to 2011, and currently serves on the Editorial Board of the WILEY JOURNAL OF WIRELESS COMMUNICATIONS AND MOBILE COMPUTING. He was also guest editor for special issues in the IEEE PERSONAL COMMUNICATIONS MAGAZINE (Energy Management in Personal Communications Systems) IEEE WIRELESS COMMUNICATIONS MAGAZINE (Cognitive Wireless Networks) and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Multi-media Network Radios, and Underwater Wireless Communications Networks). He served as a Member-at-large of the Board of Governors of the IEEE Communications Society from 2009 to 2011.