INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS
Int. J. Commun. Syst. (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/dac.2403

Cramer–Rao bound based mean-squared error and throughput analysis of superimposed pilots for semi-blind multiple-input multiple-output wireless channel estimation

Aditya K. Jagannatham[1,*,†] and Bhaskar D. Rao[2]

[1]*Electrical Engineering, I.I.T. Kanpur, Kanpur, UP 208016, India*
[2]*Electrical and Computer Engineering, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0407, USA*

SUMMARY

This work presents a study of the mean-squared error (MSE) and throughput performance of superimposed pilots (SP) for the estimation of a multiple-input multiple-output (MIMO) wireless channel. The Cramer–Rao bound (CRB) is derived for SP based estimation of the MIMO channel matrix. Employing the CRB analysis, it is proved that the asymptotic MSE bound is potentially 3 dB lower than the MSE performance of the existing SP mean based estimation (SPME) schemes. Motivated by this observation, a novel SP semi-blind scheme is presented for MIMO channel estimation. This scheme asymptotically achieves the CRB and hence has a lower MSE of estimation when compared with SPME schemes. We also derive closed form expressions for the optimal source-pilot power allocation in SP by maximizing the post-processing signal-to-noise power ratio at the receiver. In the final part, a new result is presented for the worst-case capacity of a communication channel with correlated information symbols and noise. This framework is employed to quantify the throughput performance of SP and also to demonstrate the bandwidth efficiency of SP compared with that of a conventional pilot based system. Copyright © 2012 John Wiley & Sons, Ltd.

Received 14 April 2012; Revised 12 June 2012; Accepted 12 June 2012

KEY WORDS: MIMO; superimposed; pilot; semi-blind

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems have gained popularity in recent years for application in upcoming 3G/4G wireless communication systems [1–4] owing to the increased throughput and diversity of reception features provided by such systems [5–7]. Availability of accurate channel knowledge in such systems can result in significant performance improvements, and hence channel estimation is an important area in current research [8, 9]. Traditionally, the wireless radio channel has been estimated by the transmission of a known sequence of pilot symbols [10] prior to the transmission of information bearing symbols in each estimation period. This scheme of time-multiplexed pilots, where exclusively pilot symbols are transmitted for a fraction of the communication period, is termed as conventional pilot (CP) based estimation. It offers the dual benefits of a robust estimate and low implementation complexity. However, a major concern in this estimation scheme is the potential wastage of bandwidth due to the transmission of pilot symbols that bear no information. Blind schemes [11–13] are an alternative to estimate a channel without wasting bandwidth. Frequently, such schemes cannot estimate the channel completely and leave a residual indeterminate 'phase factor', such as a complex phase for single-input multiple-output (SIMO) [14, 15] and a unitary matrix [16] for MIMO systems. Further, the optimization algorithms are often

*Correspondence to: Aditya K. Jagannatham, Electrical Engineering, I.I.T. Kanpur, Kanpur, UP 208016, India.
†E-mail: adityaj@iitk.ac.in

complex, involving second-order and higher-order statistics, and frequently result in sub-optimal performance from convergence to local minima.

Recent advances in signal processing have suggested an innovative scheme for channel estimation with superimposed pilot (SP) symbols, also termed as hidden or embedded pilots (Figure 1). SP based schemes employ additional power to transmit a repetitive sequence of pilot symbols superimposed over the information bearing data symbols and hence do not sacrifice bandwidth by exclusively transmitting pilots. Further, because they employ first-order (mean) statistics, compared with blind algorithms that traditionally employ second-order and higher-order statistics, they result in simplistic algorithms that obviate convergence problems. Thus, they offer the attractive benefit of bandwidth efficiency with moderate computational complexity. Early research on such channel estimation schemes has been reported in [17, 18]. Alternative schemes for SP based estimation have been explored in [19–23]. SP based channel estimation for orthogonal frequency-division multiplexing systems is discussed in [24].

In this work, the true Cramer–Rao bounds (CRB) for SP based estimation are derived where only approximate bounds have been derived previously in the literature [19]. It is demonstrated that the simplistic first-order statistic (mean) based SP estimation scheme or superimposed pilot mean estimate (SPME), proposed in works such as [19, 20], has a sub-optimal estimation performance compared with the CRB. Hence, an SP semi-blind (SPSB) estimation scheme is proposed, which asymptotically achieves the CRB [25, 26], thus improving the estimation performance over existing schemes. In the SIMO context, this estimate can be seen to have an asymptotic mean-squared error (MSE) that is 3 dB lower than the SPME. Further, unlike in CP based estimation where the channel estimate is independent of the data signal-to-noise power ratio (SNR) and depends only on pilot-to-noise ratio (PNR), in SP estimation, the transmitted data have a corrupting influence on the channel estimate. This aspect has been considered in the study in [20, 27]. Our work further addresses this issue of optimal pilot-source power allocation for a fixed total transmit power on the basis of maximizing the post-processing SNR (PSNR) for a Capon-like receive beamformer.

Another aspect of our work is the development of a framework for the throughput performance analysis of SP. A similar study has been presented in [28, 29] for frequency selective SIMO channels. A novel contribution of our work is to derive an expression for the capacity lower bound of channels with source–noise correlation to analyze the throughput performance of SP. This framework is more general and can be used to characterize the throughput performance of any estimator and is not limited to the minimum mean-squared error (MMSE) estimate as in [28, 30]. Specifically, this work focuses on the maximum-likelihood (ML) estimate, which is commonly employed in practice. This expression for worst-case capacity is also utilized to demonstrate the throughput gains of SP over a system employing CPs. From simulation studies employing this framework, SP can be seen to potentially outperform CP in terms of overall system throughput, especially in scenarios where the block length is small so that exclusive transmission of pilot symbols results in a significant bandwidth overhead. This typically arises in ad hoc and sensor networks [31, 32], where the information

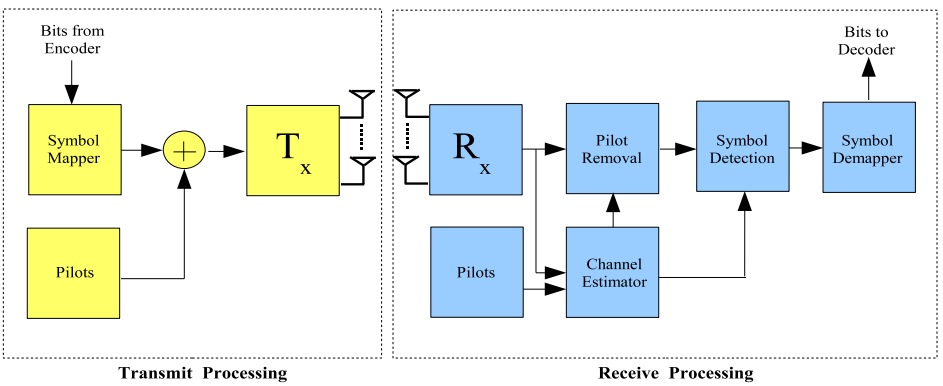

Figure 1. Schematic of a superimposed pilot system.

DOI: 10.1002/dac

is communicated in short bursts over a large number of channels. It also arises in mobile wireless scenarios where a short coherence time renders repeated training inefficient.

The rest of the paper is organized as follows. In the next section, we formulate the problem. The SPME and expressions for its MSE performance are derived in Section 2.2. The CRB analysis for SP and the SPSB are derived in Section 2.3. Optimum power allocation for SP is elaborated in Section 3. The expression for worst-case capacity with correlation is presented in Section 4 followed by performance comparison of SP with CP in Section 4.2. Finally, results from simulation studies are presented in Section 5 followed by our conclusions.

2. SUPERIMPOSED PILOT BASED MULTIPLE-INPUT MULTIPLE-OUTPUT ESTIMATION

Consider a MIMO wireless system with r receive antennas, t transmit antennas, and $r \geqslant t$, that is, at least as many receive antennas as transmit antennas. Let $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_r] \in \mathbb{C}^{r\times t}$ denote the flat-fading MIMO channel, where $\mathbf{h}_j \triangleq \left[h_{1j}, h_{2j}, \ldots, h_{rj}\right]^T$ represents the vector of complex fading coefficients between the jth transmit antenna and the receiver. The equivalent discrete-time baseband MIMO system model after matched filtering and sampling is given as

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \boldsymbol{\eta}(k), \quad 1 \leqslant k \leqslant N_b \tag{1}$$

where the index k denotes the time instant and $\mathbf{y}(k) \in \mathbb{C}^{r\times 1}, \mathbf{x}(k) \in \mathbb{C}^{t\times 1}$ denote the kth received and transmitted symbol vectors, respectively. The vector $\boldsymbol{\eta}(k) \in \mathbb{C}^{r\times 1}$ is complex circularly symmetric spatio-temporally uncorrelated additive white Gaussian noise of power σ_n^2, that is, $\mathrm{E}\left\{\boldsymbol{\eta}(k)\boldsymbol{\eta}(l)^H\right\} = \sigma_n^2\,\delta(k-l)\mathbf{I}_r$, where $\delta(k) = 1$, if $k = 0$ and 0 otherwise. The SP transmission scheme can be described as follows. Let each frame of contiguous transmitted symbols contain N_f sub-frames of length L_p symbols where $N_b \triangleq N_f L_p$ denotes the block length. The transmitted data symbols $\mathbf{x}_d^s(k)$ are assumed to be stochastic in nature with $\mathrm{E}\left\{\mathbf{x}_d^s(k)\right\} = 0$ and power P_d^s, that is, $\mathrm{E}\left\{\mathbf{x}_d^s(k)\mathbf{x}_d^s(l)^H\right\} = P_d^s\,\delta(k-l)\mathbf{I}_t$. Let $\mathbf{X}_d \triangleq \left[\mathbf{x}_d^s(1), \mathbf{x}_d^s(2), \ldots, \mathbf{x}_d^s(N_b)\right] \in \mathbb{C}^{t\times N_b}$ be the transmitted information symbol sequence. Each such sub-frame consists of independent data symbols with the pilot sequence $\mathbf{X}_p \in \mathbb{C}^{t\times L_p}$, defined as $\mathbf{X}_p \triangleq \left[\mathbf{x}_p(1), \mathbf{x}_p(2), \ldots, \mathbf{x}_p(L_p)\right]$, of length L_p symbols and pilot power P_t^s superimposed over the data symbols, that is, $\mathrm{tr}\left(\mathbf{X}_p\mathbf{X}_p^H\right) = tP_t^s L_p$. Also, let $\rho_d^s \triangleq \left(\sigma_t^s/\sigma_n^2\right)$ and $\rho_t^s \triangleq \left(P_t^s/\sigma_n^2\right)$ be the SNR and pilot-to-noise power ratio (PNR), respectively. A schematic diagram of this SP frame structure is given in Figure 2. The actual transmitted symbol at the kth instant, $\mathbf{x}^s(k)$, is therefore given as

$$\mathbf{x}^s(k) \triangleq \mathbf{x}_d^s(k) + \mathbf{x}_p^s(k) = \mathbf{x}_d^s(k) + \mathbf{x}_p\left(\mathrm{mod}(k-1, L_p) + 1\right). \tag{2}$$

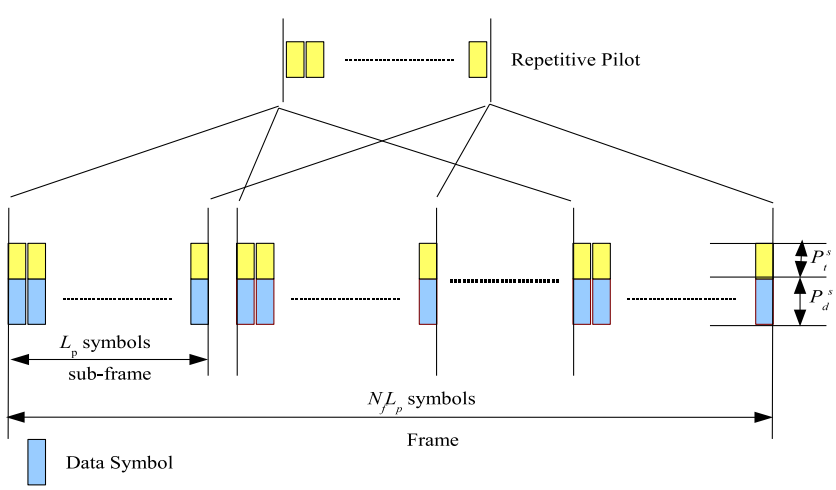

Figure 2. Schematic diagram of the superimposed pilot frame structure.

DOI: 10.1002/dac

The SP system model has the form

$$\mathbf{y}^s(k) = \mathbf{H}\underbrace{\left(\mathbf{x}_d^s(k) + \mathbf{x}_p\left(\mathrm{mod}(k-1, L_p)+1\right)\right)}_{\mathbf{x}^s(k)} + \boldsymbol{\eta}(k), \quad 1 \leqslant k \leqslant N_b, \tag{3}$$

where $\mathbf{y}^s(k), \mathbf{x}^s(k)$ are the kth received symbol vector and transmitted symbol, respectively.

2.1. Superimposed pilot mean estimate

A first-order statistic based SPME scheme for estimation of the MIMO channel $\mathbf{H}$ can be described as follows. This scheme is similar to the ones suggested in [17, 19, 33] to estimate the channel $\mathbf{H}$, which exploits the periodicity in the SP symbol transmission. Let $\bar{\mathbf{y}}^s(k) \in \mathbb{C}^{r\times 1}, 1 \leqslant k \leqslant L_p$ be defined as

$$\bar{\mathbf{y}}^s(k) \triangleq \frac{1}{N_f}\sum_{j=0}^{N_f-1}\mathbf{y}^s\left(k + jL_p\right), \quad 1 \leqslant k \leqslant L_p. \tag{4}$$

Let $\bar{\mathbf{Y}}^s \in \mathbb{C}^{r\times L_p} \triangleq \left[\bar{\mathbf{y}}^s(1), \bar{\mathbf{y}}^s(2), \ldots, \bar{\mathbf{y}}^s(L_p)\right]$ be a stacking of the processed received symbol vectors and $\mathrm{E}\left\{\bar{\mathbf{Y}}^s\right\} = \mathbf{H}\mathbf{X}_p$. The channel estimate $\hat{\mathbf{H}}_s$ is now computed by the standard least-squares procedure [34, 35] as

$$\hat{\mathbf{H}}_s = \bar{\mathbf{Y}}^s\mathbf{X}_p^\dagger = \bar{\mathbf{Y}}^s\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1} = \mathbf{Y}^s\left(\mathbf{X}_p^s\right)^\dagger, \tag{5}$$

where $\mathbf{X}_p^s \triangleq \left[\mathbf{X}_p, \mathbf{X}_p, \ldots, \mathbf{X}_p\right] \in \mathbb{C}^{t\times N_f L_p}$ is the SP signal. Refer to the preceding estimate as the SPME as it employs the mean of the received signal $\mathbf{Y}^s$. Because it is based only on the first-order statistics of $\mathbf{Y}^s$, it converges faster (compared with methods based on the second-order and higher-order statistics) while having a low complexity of implementation. The estimate $\hat{\mathbf{H}}_s$ is then used for detection of the transmitted data $\mathbf{x}_d^s(k)$ after removing the SP symbol $\mathbf{x}_p\left(\mathrm{mod}(k-1, L_p)+1\right)$.

2.2. Mean-squared error of estimation for superimposed pilot mean estimate

In this section, the MSE of estimation for the SPME given in (5) is computed. From Equation (4), the quantity $\bar{Y}^s$ is given as $\bar{\mathbf{Y}}^s = \mathbf{H}\mathbf{X}_p + \mathbf{H}\bar{\mathbf{X}}_d^s + \bar{\mathbf{N}}$, where $\bar{\mathbf{X}}_d^s$ and $\bar{\mathbf{N}}$ are defined analogously for $\mathbf{x}_d^s(k), \boldsymbol{\eta}(k), 1 \leqslant k \leqslant N_b$ as in (4). Simplifying the expression for the SP estimate given in (5), the quantity $\hat{\mathbf{H}}_s$ can be seen to be given as

$$\hat{\mathbf{H}}_s = \mathbf{H} + \mathbf{H}\bar{\mathbf{X}}_d^s\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1} + \bar{\mathbf{N}}\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}. \tag{6}$$

The MSE of the mean estimate for SP, defined as $\mathrm{MSE}_s \triangleq \mathrm{E}\left\{\left\|\hat{\mathbf{H}}_s - \mathbf{H}\right\|_F^2\right\}$, can be simplified as demonstrated in Appendix Section A to yield

$$\mathrm{MSE}_s = \mathrm{E}\left\{\mathrm{tr}\left(\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)^H\right)\right\} = \frac{1}{N_f}\left(\mathrm{tr}\left(\mathbf{H}\mathbf{H}^H\right)P_d^s + r\sigma_n^2\right)\mathrm{tr}\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}. \tag{7}$$

The optimal pilot symbol matrix for SP estimation that minimizes the MSE of estimation can be obtained as $\mathbf{X}_p^\star = \arg\min \mathrm{MSE}_s = \arg\min \mathrm{tr}\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}$. The following result gives the structure of the optimal pilot matrix $\mathbf{X}_p$.

Lemma 1
For a fixed total pilot power $\mathrm{tr}\left(\mathbf{X}_p\mathbf{X}_p^H\right) = tL_pP_t^s$, the optimal pilot symbol matrix $\mathbf{X}_p^\star \in \mathbb{C}^{t\times L_p}$ that minimizes the quantity MSE_s, the MSE of estimation of the MIMO channel $\mathbf{H}$ using SPs, is given by $\mathbf{X}_p^\star$ such that $\mathbf{X}_p^\star\left(\mathbf{X}_p^\star\right)^H = P_t^sL_p\mathbf{I}_t$, that is, the pilot symbol matrix $\mathbf{X}_p$ is orthogonal.

DOI: 10.1002/dac

Proof
Similar to [36, 37]. □

In the remainder of the paper, it is assumed that $\mathbf{X}_p = \mathbf{X}_p^\star$, the optimal orthogonal pilot sequence. Thus, the MSE for SP based estimation is given as

$$\mathrm{MSE}_s = \frac{tP_d^s}{N_b P_t^s}\mathrm{tr}\left(\mathbf{H}\mathbf{H}^H\right) + \frac{rt\sigma_n^2}{N_b P_t^s} \tag{8}$$

The asymptotic MSE of the SPME at high SNR, $\mathrm{MSE}_s^\infty \in \mathrm{O}\left(P_d^s\right)$, can be defined as

$$\mathrm{MSE}_s^\infty \triangleq P_d^s\left(\lim_{P_d^s\to\infty}\frac{\mathrm{MSE}_s}{P_d^s}\right) = \frac{tP_d^s}{N_b P_t^s}\mathrm{tr}\left(\mathbf{H}\mathbf{H}^H\right). \tag{9}$$

Hence, MSE_s can be expressed as $\mathrm{MSE}_s = \mathrm{MSE}_s^\infty + \mathrm{o}\left(\mathrm{P_d^s}\right)$. Ideally, it is desired $\mathrm{MSE}_s^\infty = 0$, to ensure that the MSE does not progressively increase without bound as the data power P_d^s increases. However, as is seen from preceding text, this is not true of the SPME, which is adversely affected as P_d^s increases. Thus, increasing P_d^s might potentially result in worsening the estimate $\hat{\mathbf{H}}_s$ and in turn results in poor detection performance. This aspect of SP estimation is of critical significance in pilot-data power allocation and is addressed in detail in Section 3. In the next section, the CRB for estimation is presented to provide a benchmark for optimal MSE performance of SP estimation.

2.3. Cramer–Rao bound for superimposed pilot estimation

In this section, the complex CRB for the SP estimation of $\mathbf{H}$ is presented. This can be seen to yield a lower asymptotic MSE than that obtained for the estimator in (5) as the SPME ignores the channel information available in the second-order statistics(source covariance). It is demonstrated for a SIMO channel that this asymptotic MSE bound for SP is 3 dB lower than that achieved by the SPME. This motivates us to develop the SPSB MIMO estimation scheme that achieves this bound at high SNR (ρ_d^s). To make the analysis tractable and demonstrate insights into SP estimation, a Gaussian symbol source is considered, that is, $\mathbf{x}_d(k) \sim \mathcal{N}\left(0, P_d^s\mathbf{I}_t\right)$. It is worth mentioning that the results derived employing this simplification are in close agreement with the performance of a system employing a discrete signal constellation such as quadrature phase-shift keying (QPSK). As suggested in [38] for the construction of CRBs of complex parameters, let the complex parameter vector $\bar{\theta} \in \mathbb{C}^{2rt\times 1}$ be constructed by stacking the parameter vector $\mathbf{H}$ and its conjugate as $\bar{\theta} \triangleq [\mathrm{vec}\,(\mathbf{H}), \mathrm{vec}\,(\mathbf{H}^*)]^T$. From the SP system model for pilot symbol outputs given in (3), the parameter-dependent log-likelihood (log-likelihood ignoring additive constants) for the estimation of the parameter vector $\bar{\theta}$ is given as

$$\mathcal{L}\left(\mathbf{Y}^s|\mathbf{X}_p^s;\bar{\theta}\right) = -N_b\ln|\mathbf{R}_e| - \sum_{i=1}^{N_b}\left(\mathbf{y}^s(i) - \mathbf{H}\mathbf{x}_p^s(i)\right)^H\mathbf{R}_e^{-1}\left(\mathbf{y}^s(i) - \mathbf{H}\mathbf{x}_p^s(i)\right) \tag{10}$$

where $\mathbf{Y}^s \triangleq [\mathbf{y}^s(1), \mathbf{y}^s(2), \ldots, \mathbf{y}^s(N_b)]$, $\mathbf{x}_p^s(i) \triangleq \mathbf{x}_p\left(\mathrm{mod}\left(i-1, L_p\right)+1\right)$ and $\mathbf{R}_e$, the covariance of this effective noise is given as $\mathbf{R}_e \triangleq \sigma_s^{2s}\mathbf{H}\mathbf{H}^H + \sigma_n^2\mathbf{I}_r$. The CRB for the estimation of $\bar{\theta}$ is given by the matrix $\mathbf{J}_{\bar{\theta}}^{-1}$, where $\mathbf{J}_{\bar{\theta}} \in \mathbb{C}^{2r\times 2r}$ is the complex Fisher information matrix (FIM) for the parameter vector $\bar{\theta} \in \mathbb{C}^{2r\times 1}$ and is given as

$$\mathbf{J}_{\bar{\theta}} = -\mathrm{E}\left\{\frac{\partial^2\mathcal{L}\left(\mathbf{Y}^s\left|\mathbf{X}_p\,;\bar{\theta}\right.\right)}{\partial\bar{\theta}\,\partial\bar{\theta}^H}\right\}$$

Therefore, the MSE lower bound for SP based estimation denoted by MSE_b is given as $\mathrm{MSE}_b = \mathrm{tr}\left(\mathbf{J}_{\bar{\theta}}^{-1}\right)$, which is also the asymptotic MSE of a ML estimator that maximizes the likelihood in (10). The SPME suggested in Zhou *et al*. [19] and described in Equation (5) is the ML estimate

DOI: 10.1002/dac

ignoring the dependance of the covariance $\mathbf{R}_e$ on $\mathbf{H}$ and employs a straightforward LS estimator, that is,

$$\hat{\mathbf{H}}_s = \arg\min\ \mathcal{L}\left(\mathbf{Y}^s|\mathbf{X}_p^s, \mathbf{R}_e; \mathbf{H}\right) = \arg\min \left\{ \sum_{i=1}^{N_b} \left(\mathbf{y}^s(i) - \mathbf{H}\mathbf{x}_p^s(i)\right)^H \mathbf{R}_e^{-1} \left(\mathbf{y}^s(i) - \mathbf{H}\mathbf{x}_p^s(i)\right) \right\},$$

where $\mathbf{R}_e$ is assumed known. This procedure, although sub-optimal, results in a simple estimation algorithm when compared with minimizing the true cost function involving $\mathbf{R}_e$ $(\mathbf{H})$. The FIM corresponding to such an estimation procedure, which exclusively employs the information in the pilots while ignoring the information in the covariance $\mathbf{R}_e$, is given by the pilot FIM component $\mathbf{J}_{\bar{\theta}}^p$ of the total FIM $\mathbf{J}_{\bar{\theta}}$ as

$$\mathbf{J}_{\bar{\theta}}^p = -\mathrm{E}\left\{ \frac{\partial^2 \mathcal{L}\left(\mathbf{Y}^s|\mathbf{X}_p^s, \mathbf{R}_e; \mathbf{H}\right)}{\partial\bar{\theta}\,\partial\bar{\theta}^H} \right\} = \begin{bmatrix} \left(\mathbf{X}_p^s\left(\mathbf{X}_p^s\right)^H\right) \bigotimes \left(\mathbf{R}_e^{-1}\right)^T & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{X}_p^s\left(\mathbf{X}_p^s\right)^H\right)^T \bigotimes \mathbf{R}_e^{-1} \end{bmatrix}. \tag{11}$$

where $\bigotimes$ denotes the matrix Kronecker product. Hence, the MSE for the exclusive pilot based SP estimation of $\mathbf{H}$ is given as

$$\mathrm{MSE}_b^p = \frac{1}{2}\mathrm{tr}\left(\mathbf{J}_{\bar{\theta}}^{p^{-1}}\right) = \mathrm{tr}\left(\left(\mathbf{X}_p^s\left(\mathbf{X}_p^s\right)^H\right)^{-1}\right)\mathrm{tr}\left(\mathbf{R}_e\right) = \mathrm{tr}\left(\mathbf{H}\mathbf{H}^H\right)\frac{P_d^s}{N_b P_t^s}t + \frac{\sigma_n^2}{N_b P_t^s}rt,$$

which is equal to the MSE for the SP estimate given in Section 2.2. The factor $\frac{1}{2}$ in the preceding expression is to account for the fact that the parameter vector $\bar{\theta}$ represents the MSE of $\mathbf{H}$ and $\mathbf{H}^*$. Thus, the SPME is sub-optimal in the sense that it considers the pilot information $\mathbf{J}_{\bar{\theta}}^p$ while ignoring the information in the second-order statistics (covariance $\mathbf{R}_e$).

The true FIM corresponding to information in both $\mathbf{X}_p$ and $\mathbf{R}_e$ can be obtained as $\mathbf{J}_{\bar{\theta}} = \mathbf{J}_{\bar{\theta}}^p + \mathbf{J}_{\bar{\theta}}^r$, where the FIM component $\mathbf{J}_{\bar{\theta}}^r$ corresponds to the information in the covariance matrix $\mathbf{R}_e$. Let the block Toeplitz parameter derivative matrix $\mathcal{E}(k) \in \mathbb{C}^{r\times t}$ be defined employing complex derivatives as $\mathcal{E}(i) \triangleq \frac{\partial\mathbf{H}}{\partial\bar{\theta}_i}$. The component $\mathbf{J}_{\bar{\theta}}^r$ can be seen to be given as [39,40],

$$\mathbf{J}_{i,j}^r = \mathbf{J}_{rt+j,rt+i}^r = N_b\left(P_d^s\right)^2 \mathrm{tr}\left(\mathcal{E}(i)\mathbf{H}^H\mathbf{R}_e^{-1}\mathbf{H}\mathcal{E}(j)^H\mathbf{R}_e^{-1}\right)$$
$$\mathbf{J}_{i,rt+j}^r = \left(\mathbf{J}_{rt+j,i}^r\right)^* = N_b\left(P_d^s\right)^2 \mathrm{tr}\left(\mathcal{E}(i)\mathbf{H}^H\mathbf{R}_e^{-1}\mathcal{E}(j)\mathbf{H}^H\mathbf{R}_e^{-1}\right).$$

The covariance FIM $\mathbf{J}_{\bar{\theta}}^r$ can be block partitioned as

$$\mathbf{J}^r \triangleq N_b\left(P_d^s\right)^2 \begin{bmatrix} \mathbf{J}_{11}^r & \mathbf{J}_{12}^r \\ \mathbf{J}_{21}^r & \mathbf{J}_{22}^r \end{bmatrix}.$$

It can be verified that $\mathbf{J}_{21}^r = \mathbf{J}_{12}^{r\,H}$ and $\mathbf{J}_{22}^r = \mathbf{J}_{11}^{r\,T}$. The block components of the FIM are given as $\mathbf{J}_{11}^r = \left(\mathbf{H}^H\mathbf{R}_e^{-1}\mathbf{H}\right) \bigotimes \left(\mathbf{R}_e^{-1}\right)^T$ and

$$\mathbf{J}_{12}^r = \left(\mathbf{e}^H \bigotimes \mathbf{H}^H\mathbf{R}_e^{-1} \bigotimes \mathbf{e}\right) \odot \left(\mathbf{e}^H \bigotimes \mathbf{H}^H\mathbf{R}_e^{-1} \bigotimes \mathbf{e}\right)^T,$$

where $\mathbf{e} = [1, 1, \ldots, 1]^T \in \mathbb{C}^{r\times 1}$ and $\odot$ denotes the matrix Hadamard product. The expressions for the FIM components $\mathbf{J}_{\bar{\theta}}^p, \mathbf{J}_{\bar{\theta}}^r$ from can be employed to obtain the true FIM $\mathbf{J}_{\bar{\theta}}$. Thus, the CRB for SP based estimation of $\mathbf{H}$ is obtained as

$$\mathrm{E}\left\{\left(\bar{\theta} - \hat{\bar{\theta}}\right)\left(\bar{\theta} - \hat{\bar{\theta}}\right)^H\right\} \geqslant \mathbf{J}_{\bar{\theta}}^{-1}. \tag{12}$$

Also, $\mathbf{J}_{\bar{\theta}} > \mathbf{J}_{\bar{\theta}}^p$ in the positive semi-definite matrix sense [41], and hence, $\mathrm{MSE}_b < \mathrm{MSE}_s$. A more insightful result can be obtained in the context of a SIMO channel $\mathbf{h} \in \mathbb{C}^{r\times 1}$, that is, $t = 1$.

DOI: 10.1002/dac

The high SNR approximation to the CRB matrix given by the succeeding result yields a critical insight into the relation between this MSE bound MSE_b and the quantity MSE_s.

Theorem 1
In the context of a SIMO wireless channel $\mathbf{h} \in \mathbb{C}^{r\times 1}$, the MSE bound for SP based estimation is given as $\mathrm{MSE}_b = \mathrm{MSE}_b^{\infty} + \mathrm{o}\left(P_d^s\right)$, where MSE_b^{∞} the high SNR asymptote is

$$\mathrm{MSE}_b^{\infty} \triangleq P_d^s \left(\lim_{P_d^s \to \infty} \frac{\mathrm{MSE}_b}{P_d^s} \right) = \frac{1}{2} \left(\frac{P_d^s}{N_b P_t^s} \right) \|\mathbf{h}\|^2. \tag{13}$$

Proof
Given in Appendix B. □

Employing the preceding result, it is straightforward to derive the asymptotic performance gain of the SPSB estimator, which makes complete use of the pilot and covariance information, over the SPME.

Lemma 2
The asymptotic MSE measures MSE_b^{∞} and MSE_s^{∞}, the asymptotic MSE bound and the asymptotic MSE of the SPME, respectively, are related as

$$\frac{\mathrm{MSE}_b^{\infty}}{\mathrm{MSE}_m^{\infty}} = \frac{1}{2}. \tag{14}$$

Proof
Follows from (9) and (13). □

The preceding result implies that at reasonably high SNRs, the MSE of estimating the channel by employing the complete information in the likelihood function in (10) is 3 dB lower than that of the SPME. Neglecting the covariance information in $\mathbf{R}_e$ results in a 3 dB loss of estimation performance in the SIMO context. We now describe a SPSB estimation algorithm in the succeeding text, which asymptotically achieves the preceding MSE bound for SP based MIMO estimation and thus has a lower MSE of estimation compared with the SPME of Section 2.

2.4. Superimposed pilot semi-blind estimation

In this section, a semi-blind SP estimator is derived, that asymptotically achieves the CRB for SP estimation by employing the information in the output covariance $\mathbf{R}_e$. Observe that the output covariance $\mathbf{R}_y$ is given as

$$\mathbf{R}_y = \mathrm{E}\left\{\mathbf{y}^s(i)\mathbf{y}^s(i)^H\right\} = \left(P_d^s + P_t^s\right)\mathbf{H}\mathbf{H}^H + \sigma_n^2\mathbf{I}_r = P_t^s\mathbf{H}\mathbf{H}^H + \mathbf{R}_e.$$

Hence, let the output covariance $\hat{\mathbf{R}}_y$ be estimated from the received data symbols as $\hat{\mathbf{R}}_y \triangleq \frac{1}{N_b}\left(\sum_{i=1}^{N_b}\mathbf{y}^s(i)\mathbf{y}^s(i)^H\right)$. Employing a Cholesky matrix factorization, one can compute the matrix estimate $\hat{\mathbf{W}}$, such that

$$\hat{\mathbf{W}}\hat{\mathbf{W}}^H = \frac{1}{\left(P_d^s + P_t^s\right)}\left(\hat{\mathbf{R}}_y - \sigma_n^2\mathbf{I}_r\right) \tag{15}$$

The matrix $\hat{\mathbf{W}}$ is also known as the whitening matrix [26] and differs from the estimate of the channel $\hat{\mathbf{H}}_b$ by a unitary matrix $\hat{\mathbf{Q}}_b$, that is, $\hat{\mathbf{H}}_b = \hat{\mathbf{W}}\hat{\mathbf{Q}}_b^H$. The unitary matrix $\hat{\mathbf{Q}}_b$ can be estimated from $\mathbf{X}_p^s$ by minimizing the likelihood

$$\hat{\mathbf{Q}}_b = \arg\min \mathrm{tr}\left(\left(\mathbf{Y}^s - \hat{\mathbf{W}}\hat{\mathbf{Q}}^H\mathbf{X}_p^s\right)^H \hat{\mathbf{R}}_e^{-1}\left(\mathbf{Y}^s - \hat{\mathbf{W}}\hat{\mathbf{Q}}^H\mathbf{X}_p^s\right)\right), \tag{16}$$

DOI: 10.1002/dac

subject to the constraint $\hat{\mathbf{Q}}_b\hat{\mathbf{Q}}_b^H = \hat{\mathbf{Q}}_b^H\hat{\mathbf{Q}}_b = \mathbf{I}_t$. The quantity $\hat{\mathbf{R}}_e$ is given as $\hat{\mathbf{R}}_e \triangleq P_d^s\hat{\mathbf{W}}\hat{\mathbf{W}}^H + \sigma_n^2\mathbf{I}_r$. It can then be demonstrated that the optimal unitary matrix $\hat{\mathbf{Q}}_b$ that minimizes the preceding likelihood is given as

$$\hat{\mathbf{Q}}_b = \hat{\mathbf{U}}_b\hat{\mathbf{V}}_b^H, \quad \text{where,} \quad \hat{\mathbf{U}}_b\hat{\mathbf{\Sigma}}_b\hat{\mathbf{V}}_b^H = \text{SVD}(\mathbf{X}_p^s\,(\mathbf{Y}^s)^H\,\hat{\mathbf{R}}_e^{-1}\hat{\mathbf{W}}), \tag{17}$$

under the condition that the pilot symbol matrix $\mathbf{X}_p$ is orthogonal, as has been assumed for optimal MSE performance. The semi-blind channel estimate is obtained as $\hat{\mathbf{H}}_b = \hat{\mathbf{W}}\hat{\mathbf{Q}}_b^H$. This SPSB estimation scheme is akin to the whitening-rotation SB procedure elaborated in [26]. The SPSB estimator yields a biased estimate at low SNR, owing to the constrained ML estimator in (16). However, the bias progressively decreases as the SNR increases. Hence, theoretically, the SB estimator asymptotically achieves the MSE lower bound in (13) at high SNR [26]. Simulation studies demonstrated in Section 5 suggest that the performance of the proposed SB scheme is close to the bound even for moderate SNR.

It can also be noted that the quantity MSE_b^∞, or the dominant term in the SP MSE bound, increases linearly with P_d^s, similar to the MSE of the simplistic mean-estimator in Section 2. This means that even if one were to use the complete available statistical information (pilot and covariance) for estimation, the least achievable MSE of estimation still increases with SNR, similar to the SPME in Section 2.2. Hence, the problem of source-pilot power allocation assumes a critical significance in the context of SP and is addressed in Section 3.

2.5. *Superimposed pilot robust estimation*

Recently, robust second-order statistical side information based channel estimation and detection schemes [42, 43] have gained significant attention in the presence of uncertainty. In the context of SP based MIMO channel estimation, the robust estimate $\hat{\mathbf{H}}_r$ of the MIMO wireless channel $\mathbf{H}$ can naturally be obtained as a solution of the optimization problem,

$$\hat{\mathbf{H}}_r = \arg\min_{\mathbf{H}_r} \text{tr}\left(\mathbf{H}_r^H\mathbf{R}_y^{-1}\mathbf{H}_r\right) \text{ s.t. } \left\|\mathbf{H}_r - \hat{\mathbf{H}}_s\right\|^2 \leqslant \epsilon,$$

where the output covariance matrix $\mathbf{R}_y$ constitutes the available second-order statistical side information. It is readily seen from the preceding formulation that the superimposed pilot robust (SPR) estimator computes the optimal estimate $\mathbf{H}_r$ belonging to the spherical uncertainty set of radius ϵ centered around the SP mean estimate $\hat{\mathbf{H}}_s$. More generally, as demonstrated in several works such as [44], the spherical uncertainty region can be relaxed to an ellipsoidal uncertainty set. However, considering the isotropic nature of the orthogonal SP sequence, the spherical uncertainty set considered previously is optimal. Further, it can be readily demonstrated that the optimal estimate $\hat{\mathbf{H}}_r$ is given as $\hat{\mathbf{H}}_r = \left(\mathbf{I} - \left(\mathbf{I} + \lambda\mathbf{R}_y\right)^{-1}\right)\hat{\mathbf{H}}_s$, where the Lagrange dual variable λ is derived as the solution of $\left\|\left(\mathbf{I} + \lambda\mathbf{R}_y\right)^{-1}\hat{\mathbf{H}}_s\right\|^2 = \epsilon$. However, the preceding SPR estimator is inefficient compared with the semi-blind estimator as it does not minimize the true likelihood cost function given in (10). As a result, it is outperformed by the SP semi-blind estimator as shown in the simulation results in Section 5.1.

3. OPTIMAL POWER ALLOCATION IN SUPERIMPOSED PILOT

It can be seen from (6) that $\hat{\mathbf{H}}_s$, the estimate of the channel is corrupted by the data symbols $\mathbf{x}_d^s(k)$, which enhance the noise during the estimation of the channel. This scenario presents an interesting tradeoff in SP systems. Although on one hand, higher data power improves the detection performance, it also results in a poor channel estimate and loss in detection performance. In fact, for a given number of frames N_f, if the source power P_d is too high, the detection performance tends to be very poor. Motivated by this observation, we derive expressions for the optimal data SNR $\rho_d^s\left(\triangleq P_d^s/\sigma_n^2\right)$ in the SIMO context to maximize the PSNR for Capon beamforming. Consider

DOI: 10.1002/dac

the analogous SIMO SP system model, obtained by setting the number of receive antennas $r = 1$ in (3), with channel vector denoted as $\mathbf{h}$. After estimation of $\mathbf{h}_s$ and subtracting the pilot symbol $x_p\left(\mathrm{mod}(k-1, L_p)+1\right)$, the model for the detection of the symbol $x_d^s(k)$ employing the computed estimate $\hat{\mathbf{h}}_s$ is given as demonstrated in (28) as

$$\tilde{\mathbf{y}}^s(k) = \hat{\mathbf{h}}_s x_d^s(k) + \underbrace{\Delta\mathbf{h}_s\left(x_p\left(\mathrm{mod}(k-1, L_p)+1\right) + x_d^s(k)\right) + \boldsymbol{\eta}(k)}_{\mathbf{v}^s(k)}, \tag{18}$$

where $\mathbf{v}(k)$ represents the effective detection noise and $\Delta\mathbf{h}_s$, the error in the estimate of $\mathbf{h}$ is defined as $\Delta\mathbf{h}_s \triangleq \mathbf{h} - \hat{\mathbf{h}}_s$. The expression for the covariance of the effective noise $\mathbf{R}_v^s$ is given in Table I. In the succeeding discussion, the expression for the optimum SNR–PNR allocation for PSNR maximization at the receiver is presented.

3.1. Minimum variance distortionless response beamformer

The minimum variance distortionless response (MVDR) beamformer [45] $\mathbf{w}_m$ is given as a solution to the detection SNR maximization criterion described as

$$\mathbf{w}_m = \arg\min \mathbf{w}^H \mathbf{R}_v^s \mathbf{w} \quad \text{subject to,} \quad \mathbf{w}^H \hat{\mathbf{h}}_s = 1.$$

From the result in [45], $\mathbf{w}_m$ is given as $\mathbf{w}_m^H = \left(\hat{\mathbf{h}}_s^H \left(\mathbf{R}_v^s\right)^{-1} \hat{\mathbf{h}}_s\right)^{-1} \hat{\mathbf{h}}^H \left(\mathbf{R}_v^s\right)^{-1}$. Substituting this in (18), the expression for the estimation of $\hat{x}_d(k)$ can be obtained as

$$\mathbf{w}_m^H \tilde{\mathbf{y}}^s(k) = \mathbf{w}_m^H \hat{\mathbf{h}}_s x_d^s(k) + \mathbf{w}_m^H \mathbf{v}^s(k) = x_d^s(k) + \mathbf{w}_m^H \mathbf{v}^s(k).$$

Hence, the PSNR for the MVDR beamformer can be seen to be given as

$$\kappa_m = \frac{P_d^s}{\mathrm{E}\left\{\left|\mathbf{w}_m^H \mathbf{v}^s(k)\right|^2\right\}} = P_d \hat{\mathbf{h}}_s^H \left(\mathbf{R}_v^s\right)^{-1} \hat{\mathbf{h}}_s. \tag{19}$$

As demonstrated in Appendix D, the preceding expression can be simplified by substituting the expression for $\mathbf{R}_v^s$ in Table I to yield

$$\kappa_m \approx \frac{\rho_t^s \rho_d^s N_b \|\mathbf{h}\|^2}{\left(\rho_d^s + \rho_t^s\right)\rho_d^s \|\mathbf{h}\|^2 + \rho_t^s\left(N_b + 1\right) + \rho_d^s}, \tag{20}$$

where ρ_d^s, ρ_t^s are the data and pilot SNR, respectively, as defined previously. Let the total symbol transmit power be constrained as

$$\rho_t^s + \rho_d^s = \alpha^s. \tag{21}$$

Table I. Covariance matrices for SP and CP systems with channel estimation error.

	SPME	**CP**
$\mathbf{R}_{vs}$	$\mathbf{R}_{vs}^s = -\frac{t\left(P_d^s\right)^2}{N_b P_t^s}\left(1+\frac{t}{N_b}\right)\mathbf{H}\mathbf{H}^H - \frac{P_d^s t\sigma_n^2}{N_b P_t^s}\mathbf{I}_r$	$\mathbf{R}_{vs}^c = -\frac{t\sigma_n^2 P_d^c}{N_f P_t^c}\mathbf{I}_r$
$\mathbf{R}_s$	$\mathbf{R}_s^s = P_d^s\left(1+\frac{tP_d^s}{N_b P_t^s}\right)\mathbf{H}\mathbf{H}^H + \frac{t\sigma_n^2 P_d^s}{N_b P_t^s}\mathbf{I}_r$	$\mathbf{R}_s^c = P_d^c \mathbf{H}\mathbf{H}^H + \frac{tP_d^c\sigma_n^2}{N_f P_t^c}\mathbf{I}_r$
$\mathbf{R}_v$	$\mathbf{R}_v^s = \sigma_n^2\mathbf{I}_r + \left(P_d^s + P_t^s\right)\left(\frac{tP_d^s}{N_b P_t^s}\mathbf{H}\mathbf{H}^H + \frac{t\sigma_n^2}{N_b P_t^s}\mathbf{I}_r\right)$	$\mathbf{R}_v^c = \sigma_n^2\mathbf{I}_r + \frac{t\sigma_n^2 P_d^c}{N_f P_t^c}\mathbf{I}_r$
	SPSB	Perfect CSI $\left(\hat{\mathbf{H}} = \mathbf{H}\right)$
$\mathbf{R}_{vs}$	$\mathbf{R}_{vs}^b = -P_d^s \mathcal{J}^b$	$\mathbf{R}_{vs}^p = \mathbf{0}_r$
$\mathbf{R}_s$	$\mathbf{R}_s^b = P_d^s \mathbf{H}\mathbf{H}^H + P_d^s \mathcal{J}^b$	$\mathbf{R}_s^p = \sigma_s^2 \mathbf{H}\mathbf{H}^H$
$\mathbf{R}_v$	$\mathbf{R}_v^b = \left(P_d^s + P_t^s\right)\mathcal{J}^b + \sigma_n^2\mathbf{I}_r$	$\mathbf{R}_v^p = \sigma_n^2\mathbf{I}_r$

DOI: 10.1002/dac

The optimum power allocation $\rho_t^\star, \rho_d^\star$ that maximizes the preceding expression for the PSNR κ_m is given by the following result.

Lemma 3
The optimum PNR $\rho_t^\star$ that maximizes the PSNR κ_m for the MVDR beamformer with the transmit power constraint in (21), is given as

$$\rho_t^\star = \frac{1}{\gamma}\left(\sqrt{\delta^2 + \delta\alpha^s\gamma} - \delta\right), \quad \rho_d^\star = \alpha^s - \rho_t^\star \tag{22}$$

where $\delta \triangleq (\alpha^s)^2 \|\mathbf{h}\|^2 + \alpha^s$ and $\gamma \triangleq N_b - \alpha^s \|\mathbf{h}\|^2$.

Proof
The preceding result can be readily obtained by differentiating the expression in (20) and noting that $\rho_t^\star > 0$. □

The expression in (22) gives the optimum pilot and data power allocation that maximizes the PSNR for MVDR reception.

4. THROUGHPUT PERFORMANCE

One of the promising aspects of SP estimation compared with CP is the potential savings in bandwidth due to the transmission of superimposed data and pilot signal, thereby eliminating an exclusive slot for the transmission of pilot symbols. In this section, the throughput performance of SP is quantified and contrasted with that of CP. The result in [30] provides an expression to characterize the worst-case capacity of a communication channel in the presence of channel estimation errors. The framework therein relies on the central assumption that the channel estimate $\hat{\mathbf{H}}$ and the estimation error $\mathbf{H} - \hat{\mathbf{H}}$ satisfy the decorrelation property, that is, $\mathrm{E}\left\{\hat{\mathbf{H}}\left(\mathbf{H} - \hat{\mathbf{H}}\right)^H\right\} = \mathbf{0}_{r\times r}$, which is satisfied by the MMSE estimate. However, this result cannot be used in the context of SP based estimation for the following reasons.

A.1 The SP channel estimate $\hat{\mathbf{H}}_s$ is correlated with the transmitted data symbols $\bar{\mathbf{X}}_d^s$ as $\mathrm{E}\left\{\hat{\mathbf{H}}_s\left[\mathbf{x}_d^s(k)\right]_i\right\} = \left(\frac{P_d^s}{N_b P_t^s}\right)\mathbf{h}_j\left(\mathbf{x}_p^s(k)\right)^H$. This can be seen from (6) and is unlike the scenario in [30] where the channel estimate is uncorrelated with the data, that is, $\mathrm{E}\left\{\hat{\mathbf{H}}\left[\mathbf{x}_d^s(k)\right]_i\right\} = \mathbf{0}_{r\times r}$.

A.2 Further, it can also be observed that the decorrelation property mentioned previously is not satisfied by many estimators including the least-squares (LS) estimator. For instance, it can be observed from Section 2.2 that $\mathrm{tr}\left(\mathrm{E}\left\{\hat{\mathbf{H}}_c\left(\mathbf{H} - \hat{\mathbf{H}}_c\right)^H\right\}\right) = -\left(\frac{t\sigma_n^2}{N_b P_t^s}\right)\mathrm{tr}\left(\mathbf{I}_r\right) \neq 0$. This is a disadvantage because the LS estimate is robust and has a low computational complexity, which makes it especially suited for implementation in wireless systems. Therefore, it is of interest to develop a framework that takes into account such estimators.

The following discussion presents a result for the worst-case capacity C_w of a channel with non-zero signal–noise correlation. This framework can be employed to quantify the throughput performance of the SP system and for comparison with a system employing CPs for channel estimation.

4.1. A throughput lower bound for channels with correlated noise

In this section, similar to the result in [30], an expression for the throughput lower bound of a communication system with correlated noise is derived. Consider the communication channel,

$$\mathbf{y}(k) = \hat{\mathbf{H}}\mathbf{x}(k) + \mathbf{v}(k) = \mathbf{s}(k) + \mathbf{v}(k), \quad \mathbf{s}(k), \mathbf{y}(k) \in \mathbb{C}^{r\times 1} \tag{23}$$

DOI: 10.1002/dac

where $\mathbf{v}(k) \in \mathbb{C}^{r\times 1}$ is additive noise and $\mathbf{s}(k) \triangleq \hat{\mathbf{H}}\mathbf{x}(k)$. The worst-case capacity for the preceding channel is

$$C_w = \min_{p_{\mathbf{v}}(\cdot),\, \mathrm{tr}(\mathbf{R}_v)=r\sigma_n^2} \; \max_{p_{\mathbf{x}}(\cdot),\, \mathrm{tr}(\mathbf{R}_x)=tP_d} \mathbf{I}(\mathbf{y};\mathbf{x}) \tag{24}$$

The important difference between the preceding model and the one in [30] is that the signal and noise components $\mathbf{s}(k), \mathbf{v}(k)$ are not necessarily uncorrelated, that is, $\mathrm{E}\left\{\mathbf{v}(k)\mathbf{s}(l)^H\right\} = \delta(k-l)\mathbf{R}_{vs} \neq \mathbf{0}$. The succeeding result gives the expression for the worst-case capacity of the preceding channel.

Theorem 2
Worst-Case Correlated Capacity: Let the system input–output model of a matrix-valued noisy communication channel be given as

$$\mathbf{y}(k) = \hat{\mathbf{H}}\mathbf{x}(k) + \mathbf{v}(k) = \mathbf{s}(k) + \mathbf{v}(k), \tag{25}$$

where $\mathbf{x}(k) \in \mathbb{C}^{t\times 1}, \mathbf{v}(k) \in \mathbb{C}^{r\times 1}$ represent the signal and the unknown noise components, respectively. Let $\mathbf{v}(k), \mathbf{x}(k)$ satisfy the covariance constraints,

$$\mathrm{E}\left\{\mathbf{x}(k)^H\mathbf{x}(l)\right\} = \delta(k-l)\mathrm{tr}\left(\mathbf{R}_x\right) = tP_d, \quad \mathrm{E}\left\{\mathbf{v}(k)^H\mathbf{v}(l)\right\} = \delta(k-l)\mathrm{tr}\left(\mathbf{R}_v\right) = r\sigma_n^2,$$

and $\delta(k) = 1$ if and only if $k = 0$ and $\delta(k) = 0$ otherwise. Further, let $\mathbf{R}_s \triangleq \mathrm{E}\left\{\mathbf{s}(k)\mathbf{s}(k)^H\right\} = \hat{\mathbf{H}}\mathbf{R}_x\hat{\mathbf{H}}^H$ and the correlation between the signal and noise components be given as

$$\mathrm{E}\left\{\mathbf{v}(k)\mathbf{s}(l)^H\right\} = \delta(k-l)\mathbf{R}_{vs} = \delta(k-l)\mathbf{R}_{sv}^H,$$

where $\mathbf{R}_{vs}$ is not necessarily $\mathbf{0}_{r\times r}$. For the preceding communication system, the worst-case capacity C_w as defined in (24) is given by the expression,

$$C_w\left(\mathbf{R}_s, \mathbf{R}_v, \mathbf{R}_{vs}\right) = \min_{\mathrm{tr}(\mathbf{R}_v)=r\sigma_n^2} \; \max_{\mathrm{tr}(\mathbf{R}_x)=tP_d} \log\left|\mathbf{I} + \mathbf{R}_{v|s}^{-1}\left(\mathbf{R}_s + \mathbf{R}_{vs}\right)\mathbf{R}_s^{-1}\left(\mathbf{R}_s + \mathbf{R}_{vs}\right)^H\right|, \tag{26}$$

where the conditional covariance $\mathbf{R}_{v|s} \in \mathbb{C}^{r\times r}$ is given as $\mathbf{R}_{v|s} \triangleq \mathbf{R}_v - \mathbf{R}_{vs}\mathbf{R}_s^{-1}\mathbf{R}_{sv}$.

Proof
See Appendix C. □

It can be seen from (26) that for the case of uncorrelated noise, that is, $\mathbf{R}_{sv} = \mathbf{R}_{vs} = \mathbf{0}_{r\times r}$, the preceding expression reduces to the result in [30],

$$C_w = \min_{\mathrm{tr}(\mathbf{R}_v)=r\sigma_n^2} \; \max_{\mathrm{tr}(\mathbf{R}_x)=tP_d} \log\left|\mathbf{I} + \mathbf{R}_v^{-1}\mathbf{R}_s\right| = \min_{\mathrm{tr}(\mathbf{R}_v)=r\sigma_n^2} \; \max_{\mathrm{tr}(\mathbf{R}_x)=tP_d} \log\left|\mathbf{I} + \mathbf{R}_v^{-1}\hat{\mathbf{H}}\mathbf{R}_x\hat{\mathbf{H}}^H\right|. \tag{27}$$

4.2. Throughput comparison of superimposed and conventional pilots

The result for the worst-case capacity derived previously is now applied to the scenarios of SP and CP based channel estimation. Let $\tilde{\mathbf{y}}^s(k)$ denote the output of the SP system after removal of the pilot symbol $\mathbf{x}_p\left(\mathrm{mod}(k-1, L_p)+1\right)$ employing the estimate $\hat{\mathbf{H}}_s$ described in (6). From (3), the input–output relation for the SP channel after pilot removal is given as

$$\tilde{\mathbf{y}}^s(k) = \hat{\mathbf{H}}_s x_d^s(k) + \underbrace{\left(\mathbf{H} - \hat{\mathbf{H}}_s\right)\left(\mathbf{x}_p\left(\mathrm{mod}(k-1, L_p)+1\right) + \mathbf{x}_d^s(k)\right) + \boldsymbol{\eta}(k)}_{\mathbf{v}^s(k)}, \tag{28}$$

where $\mathbf{s}^s(k) \triangleq \hat{\mathbf{H}}x_d^s(k)$ and $\mathbf{s}^s(k), \mathbf{v}^s(k) \in \mathbb{C}^{r\times 1}$ denote the effective SP channel output (after pilot removal), noise, respectively, at the kth time instant. The optimal $\mathbf{R}_x$ that maximizes the worst-case capacity in (26) depends on the channel matrix $\mathbf{H}$ and can be fairly challenging to compute. However, in simplistic communication scenarios where the channel information is not fed back to the

DOI: 10.1002/dac

transmitter, a reasonable choice for the transmit covariance matrix is $\mathbf{R}_x = P_t \mathbf{I}_t$, where power is loaded uniformly on all the transmit antennas. Further, because our study focuses on a comparison between the SP and CP scenarios, the preceding choice of $\mathbf{R}_x$ can be used as a benchmark. Hence, the throughput lower bounds for the SPME and SPSB estimation schemes in bits per channel use are given as

$$C_w^s = C_w\left(\mathbf{R}_s^s, \mathbf{R}_v^s, \mathbf{R}_{vs}^s\right), \quad C_w^b = C_w\left(\mathbf{R}_s^b, \mathbf{R}_v^b, \mathbf{R}_{vs}^b\right), \tag{29}$$

where the expressions for the covariance matrices $\mathbf{R}_s, \mathbf{R}_v, \mathbf{R}_{vs}$ for the different estimation schemes are listed in Table I.

4.3. *Computation of* $\mathbf{R}_s, \mathbf{R}_v, \mathbf{R}_{vs}$ *for superimposed pilot estimation*

The covariance matrices $\mathbf{R}_s^s, \mathbf{R}_v^s, \mathbf{R}_{vs}^s$ for the SPME and $\mathbf{R}_s^b, \mathbf{R}_v^b, \mathbf{R}_{vs}^b$ for the SPSB can be derived as follows. Employing the expression for SPME in (6), the quantity $\mathrm{E}\left\{\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)^H\right\}$ can be simplified as

$$\begin{aligned}
\mathrm{E}\left\{\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)\left(\hat{\mathbf{H}}_s - \mathbf{H}\right)^H\right\} &= \left(\frac{1}{N_b P_t^s}\right)^2 \left(\mathrm{E}\left\{\mathbf{H}\bar{\mathbf{X}}_d^s \mathbf{X}_p^H \mathbf{X}_p \left(\bar{\mathbf{X}}_d^s\right)^H \mathbf{H}^H\right\} + \mathrm{E}\left\{\bar{\mathbf{N}}\mathbf{X}_p^H \mathbf{X}_p \bar{\mathbf{N}}^H\right\}\right) \\
&= \left(\frac{1}{N_b P_t^s}\right)^2 \left(\mathrm{E}\left\{\mathbf{H}\bar{\mathbf{X}}_d^s \mathbf{G}\left(\bar{\mathbf{X}}_d^s\right)^H \mathbf{H}^H\right\} + \mathrm{E}\left\{\bar{\mathbf{N}}_s \mathbf{G} \bar{\mathbf{N}}_s^H\right\}\right) \\
&= \left(\frac{t P_d^s}{N_b P_t^s}\right)\mathbf{H}\mathbf{H}^H + \left(\frac{t\sigma_n^2}{N_b P_t^s}\right)\mathbf{I}_r, \quad \mathbf{G} \triangleq \mathbf{X}_p^H \mathbf{X}_p,
\end{aligned}$$

where the last equality follows from the simplification $\mathrm{E}\left\{X_d \mathbf{G} X_d^H\right\} = \mathrm{tr}\left(\mathbf{G}\right) P_d^s \mathbf{I}_t = t N_b P_t^s P_d^s \mathbf{I}_t$, as demonstrated in Appendix Section A. Hence, the expression for $\mathbf{R}_s^s = \mathrm{E}\left\{\hat{\mathbf{H}}_s x_d^s(k)\left(\hat{\mathbf{H}}_s x_d^s(k)\right)^H\right\}$ can be obtained as

$$\begin{aligned}
\mathbf{R}_s^s &= \mathrm{E}\left\{\hat{\mathbf{H}}_s x_d^s(k)\left(\hat{\mathbf{H}}_s x_d^s(k)\right)^H\right\} = P_d^s \mathrm{E}\left\{\hat{\mathbf{H}}_s \hat{\mathbf{H}}_s^H\right\} \\
&= P_d^s \mathrm{E}\left\{\left(\hat{\mathbf{H}}_s - \mathbf{H} + \mathbf{H}\right)\left(\hat{\mathbf{H}}_s - \mathbf{H} + \mathbf{H}\right)^H\right\} \\
&= P_d^s\left(\left(\frac{t P_d^s}{N_b P_t^s}\right)\mathbf{H}\mathbf{H}^H + \left(\frac{t\sigma_n^2}{N_b P_t^s}\right)\mathbf{I}_r\right) + P_d^s \mathbf{H}\mathbf{H}^H \\
&= P_d^s\left(1 + \frac{t P_d^s}{N_b P_t^s}\right)\mathbf{H}\mathbf{H}^H + \frac{t\sigma_n^2 P_d^s}{N_b P_t^s}\mathbf{I}_r.
\end{aligned}$$

Because of the non-linear nature of the SPSB outlined in Section 2.4, it is fairly complex to derive the expressions for the covariance matrices $\mathbf{R}_s^b, \mathbf{R}_v^b, \mathbf{R}_{vs}^b$ in a straightforward fashion similar to that of the preceding SPME. However, the asymptotic covariance matrices can be derived by employing the expression for the semi-blind CRB in (12) as follows. Let $\mathcal{J}^b$, the error matrix for SPSB estimation, be given as

$$\mathcal{J}_{ij}^b \triangleq \sum_{k=0}^{t-1}\left[\mathbf{J}_{\bar{\theta}}^{-1}\right]_{i+kr, j+kr}. \tag{30}$$

It can then be demonstrated that the asymptotic error covariance $\mathrm{E}\left\{\left(\mathbf{H} - \hat{\mathbf{H}}_b\right)\left(\mathbf{H} - \hat{\mathbf{H}}_b\right)^H\right\}$ can be derived as

$$\lim_{N_b \to \infty} \mathrm{E}\left\{\left(\hat{\mathbf{H}}_b - \mathbf{H}\right)\left(\hat{\mathbf{H}}_b - \mathbf{H}\right)^H\right\} = \mathcal{J}^b. \tag{31}$$

DOI: 10.1002/dac

The preceding expression provides a good approximation for the error covariance and is tight even at moderate SNR. Now, in a fashion similar to that for the SPME illustrated previously, one can proceed to derive the error covariance matrices associated with SP semi-blind estimation. The expressions for these covariance matrices are given in Table I.

4.4. Conventional pilot based estimation

In contrast to SP, CP based channel estimation involves the exclusive transmission of pilot symbols, which results in a bandwidth overhead. The CP system frame can be modeled as a transmission of L_p pilot symbols followed by $\left(N_f - 1\right) L_p$ information bearing data symbols. Let P_d^c, P_t^c denote the transmit power of the data and pilot symbols, respectively, for the CP system. A schematic diagram for the frame structure of a CP based system is given in Figure 3. The input–output model for the CP system is given as

$$\mathbf{y}^c(k) = \mathbf{H}\mathbf{x}^c(k) + \boldsymbol{\eta}(k), \quad \text{where} \quad \mathbf{x}^c(k) = \begin{cases} \mathbf{x}_p^c(k), 1 \leqslant k \leqslant L_p \\ \mathbf{x}_d^c(k), L_p + 1 \leqslant k \leqslant N_b \end{cases} \tag{32}$$

Defining a stacking of the received pilot symbol outputs as $\mathbf{Y}^c \triangleq \left[\mathbf{y}^c(1), \mathbf{y}^c(2), \ldots, \mathbf{y}^c(L_p)\right]$, the conventional estimate $\hat{\mathbf{H}}_c$ is then given by the well-known LS estimate as

$$\hat{\mathbf{H}}_c = \mathbf{Y}^c \left(\mathbf{X}_p^c\right)^{\dagger} = \mathbf{Y}^c \left(\mathbf{X}_p^c\right)^H \left(\mathbf{X}_p^c \left(\mathbf{X}_p^c\right)^H\right)^{-1}, \tag{33}$$

where $\mathbf{X}_p^c = \left[\mathbf{x}_p^c(1), \ldots, \mathbf{x}_p^c\left(L_p\right)\right]$. The worst-case throughput performance of CP is given as

$$C_w^c = \left(1 - \frac{1}{N_f}\right) C_w\left(\mathbf{R}_s^c, \mathbf{R}_v^c, \mathbf{R}_{vs}^c\right), \tag{34}$$

where the factor $\left(\frac{N_b - L_p}{N_b}\right) = \left(1 - \frac{1}{N_f}\right)$ arises because of a loss of one sub-frame per frame owing to exclusive transmission of the pilot symbols. This results in a loss in throughput in CP systems, especially for a low number of sub-frames N_f. As illustrated by the simulation results, for reasonable values of SNR $\left(= P_d/\sigma_n^2\right)$, PNR$\left(= P_t/\sigma_n^2\right)$ and number of sub-frames$(= N_f)$, an SP scheme has a throughput of approximately 0.5 bits per channel use greater than that of CP. This is predominantly because the CP is disadvantaged by the loss of one sub-frame of bandwidth due to the transmission of pilot symbols exclusively, whereas the estimation errors are comparable at low SNRs. Hence, for reasonable SNRs and short data frame sizes SP has a higher throughput than CP. This makes SP especially suitable for employment in scenarios such as ad hoc and sensor networks, where the information transmitted is typically bursty and of short duration and the pilot overhead in CP would be comparable with the total transmitted data.

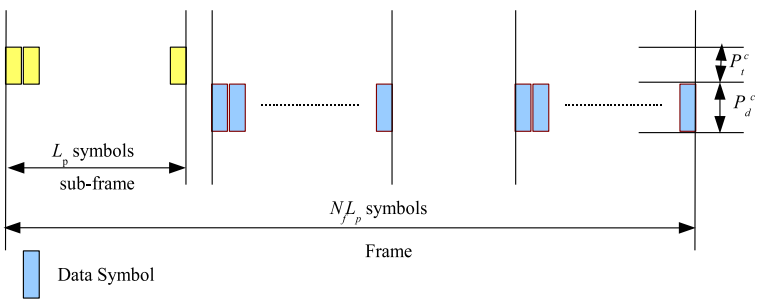

Figure 3. Schematic of conventional (time-multiplexed) pilots frame (block) structure.

DOI: 10.1002/dac

5. SIMULATION RESULTS

5.1. Mean-squared error of estimation

In our simulations a MIMO/SIMO wireless channel is employed with $r = 4$ receive antennas, $t \in \{4, 1\}$ transmit antennas and QPSK symbol modulation. We consider a Rayleigh fading channel with coefficients $\mathbf{H}_{ij}$, $1 \leqslant i, j \leqslant 4$, generated as independent zero-mean circularly symmetric complex Gaussian random variables of unit variance, that is, $\mathrm{E}\left\{\|\mathbf{H}_{ij}\|^2\right\} = 1$. In the first example, the estimation performance of the semi-blind SP scheme with an orthogonal pilot sequence $\mathbf{X}_p$ of length $L_p = 12$ symbols and $N_f = 10$ sub-frames per frame, or a total of $N_b = 120$ symbols per frame is considered. It is assumed that the channel does not vary significantly over the block, or in other words, the channel coherence time is larger than the block duration. Figure 4 shows computed MSE averaged over 2000 independent realizations of the wireless channel $\mathbf{H}$. It is seen that the MSE of the SP estimate $\hat{\mathbf{H}}_s$ given by (5) is in close agreement with theory from Section 2.2. The semi-blind estimate in (17) has a lower MSE than the SPME and achieves the CRB in (12). The asymptotic semi-blind estimator, which has the least MSE, is the semi-blind estimate as $N_b \to \infty$, implying that the estimate of the whitening matrix $\hat{\mathbf{W}} = \mathbf{W}$. It can also be seen that even though the CRB results are derived assuming Gaussian signaling, they are in close agreement with the performance of a system employing a discrete constellation, QPSK in the preceding case. Figure 5 shows the MSE of estimation of a SIMO channel $\mathbf{h}$ with $r = 4$ receive antennas. In this scenario, the semi-blind estimate in (17) involves the constrained estimation of a scalar phase. As illustrated in lemma 2, it is seen that at high SNR the SP MSE bound is 3 dB lower than the MSE of the SP SPME. The MSE of the estimate $\hat{\mathbf{h}}_f$, which is obtained by employing a numerical optimization routine **fminunc**$(\cdot)$ in `MATLAB` to optimize the likelihood in (10), is also plotted. The SPME $\hat{\mathbf{h}}_s$ is employed to initialize the procedure. This estimate can also be seen to achieve the asymptotic MSE bound for SP estimation. Thus, both these estimators asymptotically outperform the SP mean-estimator. Finally, the MSE of the CP estimate from (33) is plotted for comparison and can be seen to outperform SP based estimation. This is expected as the performance of CP is not limited by the data SNR. However, CP has a net throughput loss compared with SP as is seen next. Further, the MSE performance of the robust SP estimator presented in Section 2.5 is compared with that of the SP mean and semi-blind

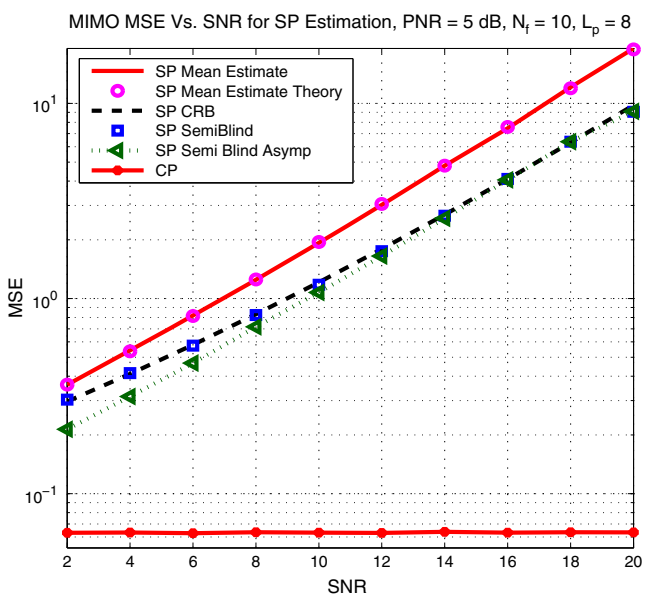

Figure 4. Mean-squared error of estimation of MIMO wireless channel with $r = t = 4$, PNR $= 5$ dB, $N_f = 10$, and $L_p = 8$ symbols.

DOI: 10.1002/dac

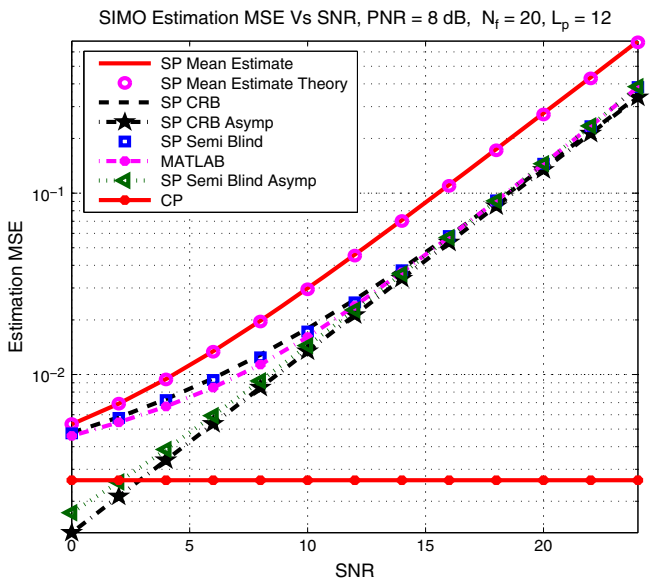

Figure 5. Mean-squared error of estimation of SIMO Rayleigh wireless channel with $r = 4$ antennas, $N_f = 20$, $L_p = 8$, and $\mathrm{PNR} = 5$ dB.

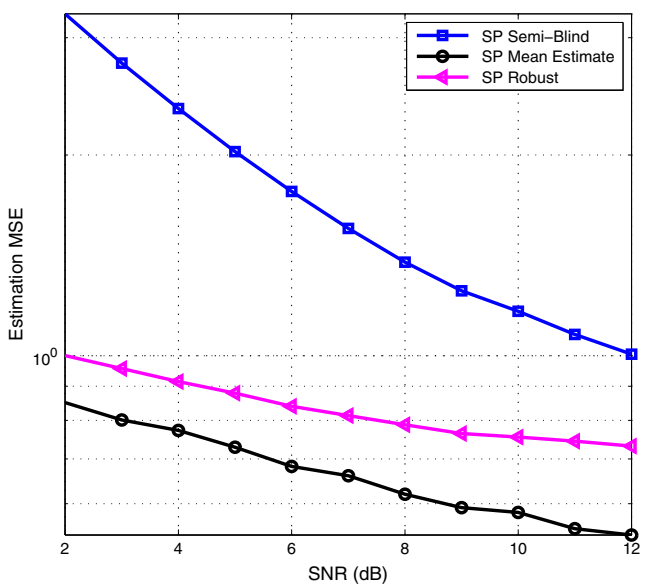

Figure 6. Mean-squared error comparison of SP robust and semi-blind estimation for MIMO wireless channel with $r = 4, t = 2$, and $L_p = 8$ symbols.

estimators in Figure 6. It can be clearly seen that the SP semi-blind estimator has a significant performance gain over the SP robust estimator, thus making it an ideal choice for SP channel estimation in the presence of second-order statistics based covariance side information.

5.2. *Optimal power allocation*

Next, the problem of optimal data/pilot transmit power allocation for receive PSNR maximization is considered, which was examined in Section 3. Figure 7 demonstrates the symbol error rate (SER) performance as a function of SNR versus transmitted SNR for a QPSK transmit constellation, when MVDR beamforming is employed at the receiver. The receiver SER corresponding to several

DOI: 10.1002/dac

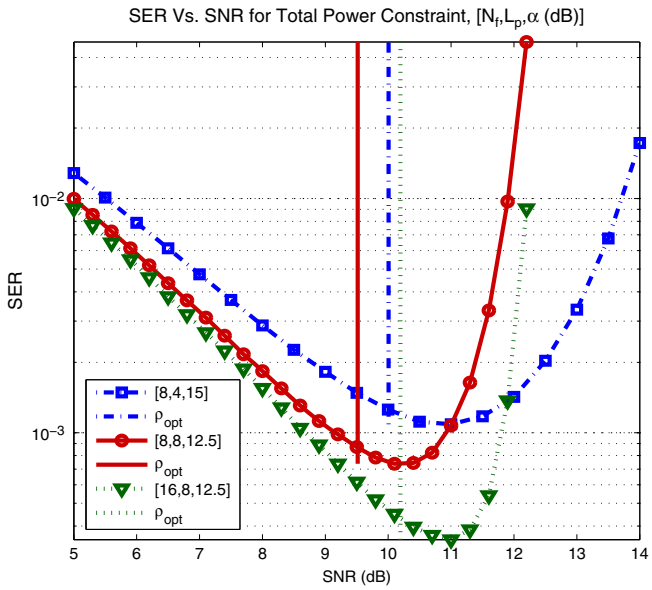

Figure 7. Detection performance versus SNR for SP based estimation. SER versus SNR $\left(\sigma_s^2/\sigma_n^2\right)$ for QPSK signaling, $r = 4$ SIMO channel, and different $\left[N_f, L_p, \alpha^s \text{ (dB)}\right]$.

choices of sub-frame number, pilot sequence length and total transmit power is considered, given in the figure by the legend entry $\left[N_f, L_p, \alpha^s \text{ (dB)}\right]$. For instance, the legend entry $[16, 8, 12.5]$ denotes the SIMO SER performance curve for $N_f = 16, L_p = 8, \alpha^s = 12.5$ dB. The SER performance reaches a minimum for a unique SNR (in this scenario for $\rho_d \approx 11$ dB) and increases for higher data power allocation. The corresponding vertical line represents the analytically computed optimal power allocation from the expression in (22) and is seen to be approximately 0.5 dB away from optimal performance. Thus, it yields a reliable benchmark for optimal power allocation. Figure 8 illustrates the optimal power allocation ratio $10\log_{10}\left(\frac{\rho_d^\star}{\rho_t^\star}\right)$ versus total transmit power α^s dB for different numbers of pilot length L_p and sub-frames N_f. It can be seen that as the

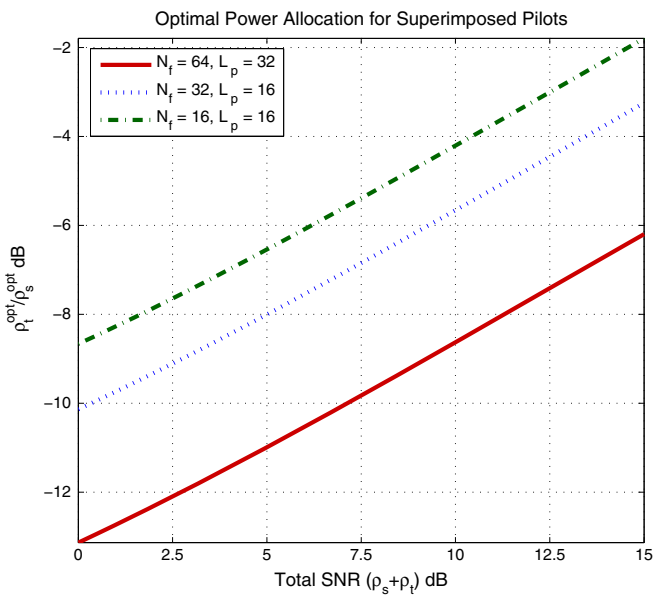

Figure 8. Optimal power allocation ratio $10\log_{10}\left(\frac{\rho_d^\star}{\rho_t^\star}\right)$ of a $r = 4$ antenna SIMO channel versus total power (α^s dB) for various N_f, L_p.

DOI: 10.1002/dac

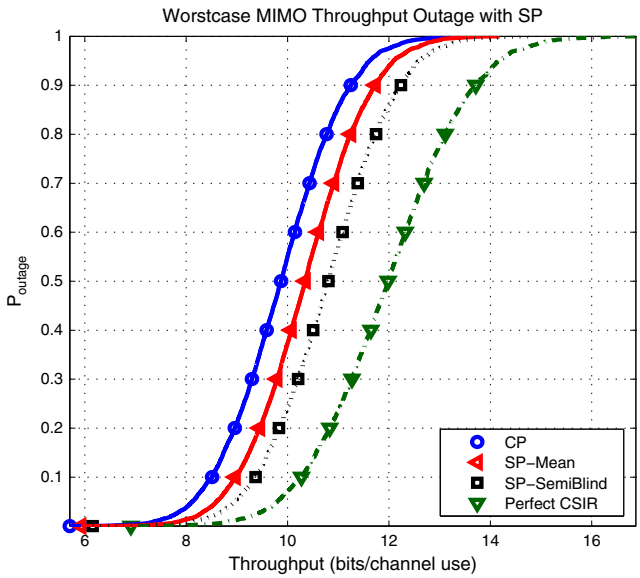

Figure 9. Throughput performance of SP and CP versus N_f, $\mathrm{SNR} = \mathrm{PNR} = 5$ dB, $L_p = 64$.

block length ($N_f L_p$) increases, the fraction of pilot power decreases from -8 to -13 dB. Further, increasing total transmit power results in increasing pilot power allocation to offset the increase in estimation error from data.

5.3. *Throughput performance*

Utilizing the framework of worst-case correlated capacity developed in Section 4.2, the net throughput performance of a SP system is computed and contrasted with the performance of a system employing conventional or time-multiplexed pilots. The results shown in figure 9 consider a system with $L_p = 64$ pilot symbols, $N_f = 4$ sub-frames and PNR, SNR fixed at 5 dB. For fairness, we scale the pilot power P_t^c and data power P_d^c of the CP by $\sqrt{N_f}$ and $\frac{1}{\sqrt{1-\frac{1}{N_f}}}$, respectively, so that the SP and CP systems are allocated equal total pilot-data power. Employing the expressions in (29) and (34), the probability of outage of the throughput lower bounds for SP and CP is plotted. In Fig. (9) it can be seen that the throughput of SP semi-blind estimation and SP mean estimation is approximately 1.0 to 0.5 bits per channel use, respectively, higher than that of CP based estimation. This throughput margin progressively decreases as the CP bandwidth loss relative to the block length, that is, L_p symbols per block of $N_b = N_f L_p$ symbols, decreases. Thus, SP estimation can potentially yield significant bandwidth gains, especially in scenarios that warrant communication in bursts of shorter block lengths, where employing CP results in significant pilot overheads.

6. CONCLUSION

In the preceding work, we presented a detailed analysis of semi-blind SP based MIMO wireless channel estimation. Employing a second-order statistics based analysis, the semi-blind CRB for MIMO SP estimation, which includes both the training and blind information, has been derived. It has been demonstrated that the semi-blind channel estimation scheme results in a 3 dB lower MSE of estimation, thus significantly improving the accuracy of the channel estimate compared with conventional channel estimation. A novel semi-blind SP MIMO channel estimation algorithm that asymptotically achieves the preceding semi-blind CRB has been presented. A closed form expression has also been derived for the superimposed pilot-data power allocation towards receive SINR maximization. The effective throughput of SP and CP systems has been analyzed employing a novel result for the worst-case capacity with correlated symbols and noise. It has been observed that SP has

DOI: 10.1002/dac

a higher effective throughput than CP based systems. Thus, SP based estimation can lead to a significant conservation of bandwidth in wireless MIMO communication systems. Further, semi-blind SP MIMO channel estimation can be employed to significantly enhance the accuracy of the channel estimate compared with conventional SP estimation that ignores the second-order statistical information. Simulations results have been presented, which illustrate the MSE, BER and throughput performance of the proposed SP channel estimation, power allocation schemes.

APPENDIX A: PROOF OF EXPRESSION FOR MSE_S IN SECTION 2.2

The expression for MSE_s can be simplified as

$$\mathrm{MSE}_s = \mathrm{tr}\left(\mathrm{E}\left\{\left(\mathbf{H}\bar{\mathbf{X}}_d^s\right)\mathbf{F}_p\left(\mathbf{H}\bar{\mathbf{X}}_d^s\right)^H\right\}\right) + \mathrm{tr}\left(\mathrm{E}\left\{\bar{\mathbf{N}}\mathbf{F}_p\bar{\mathbf{N}}^H\right\}\right),$$

where $\mathbf{F}_p \triangleq \left(\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\right)\left(\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\right)^H$. Let $\mathbf{U} = [\mathbf{u}(1), \mathbf{u}(2), \ldots, \mathbf{u}(L)] \in \mathbb{C}^{m\times L}$ be any matrix such that $\mathrm{E}\left\{\mathbf{u}(k)\mathbf{u}(l)^H\right\} = \sigma_u^2\delta(k-l)\mathbf{I}$. Then, $\mathrm{E}\left\{\mathbf{U}\mathbf{F}_p\mathbf{U}^H\right\}$ is given as

$$\begin{aligned}\mathrm{E}\left\{\mathbf{U}\mathbf{F}_p\mathbf{U}^H\right\} &= \sum_{j=1}^{L}\sum_{i=1}^{L}\mathrm{E}\left\{\mathbf{u}(i)\left[\mathbf{F}_p\right]_{ij}\mathbf{u}(j)^H\right\}\\ &= \sum_{j=1}^{L}\sum_{i=1}^{L}\sigma_u^2\delta(i-j)\left[\mathbf{F}_p\right]_{ij}\mathbf{I}_m\\ &= \sum_{j=1}^{L}\sigma_u^2\left[\mathbf{F}_p\right]_{jj}\mathbf{I}_m = \sigma_u^2\mathrm{tr}\left(\mathbf{F}_p\right)\mathbf{I}_m.\end{aligned}$$

Hence, it follows that $\mathrm{E}\left\{\bar{\mathbf{X}}_d^s\mathbf{F}_p\left(\bar{\mathbf{X}}_d^s\right)^H\right\} = \frac{P_d^s}{N_f}\mathrm{tr}\left(\mathbf{F}_p\right)\mathbf{I}_t$ and $\mathrm{E}\left\{\bar{\mathbf{N}}\mathbf{F}_p\bar{\mathbf{N}}^H\right\} = \frac{\sigma_n^2}{N_f}\mathrm{tr}\left(\mathbf{F}_p\right)\mathbf{I}_r$. Further,

$$\begin{aligned}\mathrm{tr}\left(\mathbf{F}_p\right) &= \mathrm{tr}\left(\left(\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\right)\left(\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\right)^H\right)\\ &= \mathrm{tr}\left(\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\mathbf{X}_p\mathbf{X}_p^H\right) = \mathrm{tr}\left(\left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1}\right)\end{aligned}$$

Hence, the expression in (7) follows.

APPENDIX B: PROOF OF THEOREM 1

It can be observed that

$$\begin{aligned}\mathrm{E}\left\{\mathrm{tr}\left(\frac{\partial\left(\mathbf{y}^s(i)-\mathbf{h}x_p^s(i)\right)^H}{\partial\bar{\theta}_i}\frac{\partial\mathbf{R}_e^{-1}}{\partial\bar{\theta}_j^*}\left(\mathbf{y}^s(i)-\mathbf{h}x_p^s(i)\right)\right)\right\} &= \mathrm{tr}\left(x_p^s(i)^*\frac{\partial\mathbf{h}^H}{\partial\bar{\theta}_i}\frac{\partial\mathbf{R}_e^{-1}}{\partial\bar{\theta}_j^*}\mathrm{E}\left\{\mathbf{y}^s(i)-\mathbf{h}x_p^s(i)\right\}\right)\\ &= 0\end{aligned}$$

as $\mathrm{E}\left\{\mathbf{y}^s(i)-\mathbf{h}x_p^s(i)\right\} = \mathrm{E}\left\{\mathbf{h}x_d^s(i)+\boldsymbol{\eta}(i)\right\} = 0$. Hence, the total FIM corresponding to information in both $\mathbf{X}_p$ and $\mathbf{R}_e$ can be obtained as $\mathbf{J}_{\bar{\theta}} = \mathbf{J}_{\bar{\theta}}^p + \mathbf{J}_{\bar{\theta}}^r$ where the FIM component $\mathbf{J}_{\bar{\theta}}^r$ corresponds to the information in the covariance matrix $\mathbf{R}_e$. From the results for the FIM of a complex Gaussian stochastic process [39,40], the covariance FIM component $\mathbf{J}_{\bar{\theta}}^r \in \mathbb{C}^{2r\times 2r}$ is given as

$$\mathbf{J}_{\bar{\theta}}^r(i,j) = \mathbf{J}_{\bar{\theta}}^r(r+j, r+i) = N_b\left(P_d^s\right)^2\mathrm{tr}\left(\frac{\partial\mathbf{h}}{\partial\mathbf{h}_i}\mathbf{h}^H\mathbf{R}_e^{-1}\mathbf{h}\frac{\partial\mathbf{h}}{\partial\mathbf{h}_j}^H\mathbf{R}_e^{-1}\right), \quad 1\leqslant i,j\leqslant r$$

$$\mathbf{J}_{\bar{\theta}}^r(i,r+j) = \left(\mathbf{J}_{\bar{\theta}}^r(r+j,i)\right)^* = N_b\left(P_d^s\right)^2\mathrm{tr}\left(\frac{\partial\mathbf{h}}{\partial\mathbf{h}_i}\mathbf{h}^H\mathbf{R}_e^{-1}\frac{\partial\mathbf{h}}{\partial\mathbf{h}_j}\mathbf{h}^H\mathbf{R}_e^{-1}\right), \quad 1\leqslant i,j\leqslant r.$$

DOI: 10.1002/dac

It can then be shown after simplification that the matrix $\mathbf{J}_{\bar{\theta}}^{r}$ is given as

$$\mathbf{J}_{\bar{\theta}}^{r} = N_b \left(P_d^s\right)^2 \begin{bmatrix} \left(\mathbf{h}^H \mathbf{R}_e^{-1}\mathbf{h}\right)\left(\mathbf{R}_e^{-1}\right)^T & \left(\mathbf{h}^H \mathbf{R}_e^{-1}\right)^T \mathbf{h}^H \mathbf{R}_e^{-1} \\ \left(\mathbf{h}^H \mathbf{R}_e^{-1}\right)^H \left(\mathbf{h}^H \mathbf{R}_e^{-1}\right)^* & \left(\mathbf{h}^H \mathbf{R}_e^{-1}\mathbf{h}\right)\mathbf{R}_e^{-1} \end{bmatrix}.$$

Using results on matrix inversion [41], the quantities $\mathbf{h}^H \mathbf{R}_e^{-1}$ and $\mathbf{h}^H \mathbf{R}_e^{-1}\mathbf{h}$ can be further simplified as $\mathbf{h}^H \mathbf{R}_e^{-1} = \frac{\mathbf{h}^H}{\sigma_n^2 + P_d^s \|\mathbf{h}\|^2}$ and $\mathbf{h}^H \mathbf{R}_e^{-1}\mathbf{h} = \frac{\|\mathbf{h}\|^2}{\sigma_n^2 + P_d^s \|\mathbf{h}\|^2}$. Substituting these expressions in the preceding FIM expression, the final expression for $\mathbf{J}_{\bar{\theta}}^{r}$ can be obtained. The SP FIM is given as

$$\mathbf{J}_{\bar{\theta}} = N_b P_t^s \begin{bmatrix} \left(\mathbf{R}_e^{-1}\right)^T & \mathbf{0}_{r\times r} \\ \mathbf{0}_{r\times r} & \mathbf{R}_e^{-1} \end{bmatrix} + N_b \left(P_d^s\right)^2 \begin{bmatrix} \frac{\|\mathbf{h}\|^2 \left(\mathbf{R}_e^{-1}\right)^T}{\sigma_n^2 + P_d^s \|\mathbf{h}\|^2} & \frac{\mathbf{h}^*\mathbf{h}^H}{\left(\sigma_n^2 + P_d^s \|\mathbf{h}\|^2\right)^2} \\ \frac{\mathbf{h}\mathbf{h}^T}{\left(\sigma_n^2 + P_d^s \|\mathbf{h}\|^2\right)^2} & \frac{\|\mathbf{h}\|^2 \mathbf{R}_e^{-1}}{\sigma_n^2 + P_d^s \|\mathbf{h}\|^2} \end{bmatrix}.$$

Let the constants $\alpha, \beta, \gamma, \theta$ be defined as $\alpha \triangleq N_b P_t^s$, $\gamma \triangleq \sigma_n^2 + P_d^s \|\mathbf{h}\|^2$, $\beta \triangleq \frac{N_b \left(P_d^s\right)^2}{\gamma}$ and $\theta \triangleq \frac{\beta}{\gamma}$. Substituting these in the preceding expression for the FIM, $\mathbf{J}_{\bar{\theta}}$ can be written as

$$\mathbf{J}_{\bar{\theta}} = \underbrace{\begin{bmatrix} \left(\alpha + \beta \|\mathbf{h}\|^2\right)\left(\mathbf{R}_e^{-1}\right)^T & \mathbf{0} \\ \mathbf{0} & \left(\alpha + \beta \|\mathbf{h}\|^2\right)\mathbf{R}_e^{-1} \end{bmatrix}}_{\mathbb{K}_{\bar{\theta}}^{-1}} + \theta \begin{bmatrix} \mathbf{0} & \mathbf{h}^*\mathbf{h}^H \\ \mathbf{h}\mathbf{h}^T & \mathbf{0} \end{bmatrix}.$$

where $\mathbb{K}_{\bar{\theta}}$ is defined as

$$\mathbb{K}_{\bar{\theta}} \triangleq \frac{1}{\alpha + \beta \|\mathbf{h}^2\|} \begin{bmatrix} \mathbf{R}_e^T & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_e \end{bmatrix}.$$

Employing the matrix inversion lemma [41], the CRB for the parameter vector $\bar{\theta}$ given by $\mathbf{J}_{\bar{\theta}}^{-1}$ can be expressed as

$$\mathbf{J}_{\bar{\theta}}^{-1} = \mathbb{K}_{\bar{\theta}} - \mathbb{K}_{\bar{\theta}} \begin{pmatrix} \mathbf{h}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{h} \end{pmatrix} \mathbb{R}_{\bar{\theta}}^{-1} \begin{pmatrix} \mathbf{0} & \mathbf{h}^H \\ \mathbf{h}^T & \mathbf{0} \end{pmatrix} \mathbb{K}_{\bar{\theta}},$$

where the matrix $\mathbb{R}_{\bar{\theta}}$ is defined as

$$\mathbb{R}_{\bar{\theta}} \triangleq \frac{1}{\theta}\mathbf{I}_{2r} + \begin{bmatrix} \mathbf{0} & \mathbf{h}^H \\ \mathbf{h}^T & \mathbf{0} \end{bmatrix} \mathbb{K}_{\bar{\theta}} \begin{bmatrix} \mathbf{h}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{h} \end{bmatrix} = \frac{1}{\theta}\mathbf{I}_{2r} + \frac{1}{\alpha + \beta \|\mathbf{h}\|}^2 \begin{bmatrix} \mathbf{0} & \mathbf{h}^H \mathbf{R}_e \mathbf{h} \\ \mathbf{h}^T \mathbf{R}_e^T \mathbf{h}^* & \mathbf{0} \end{bmatrix}.$$

The MSE bound for the estimation of the parameter vector $\bar{\theta}$ is given as

$$\mathrm{MSE}_b = \frac{1}{2}\mathrm{tr}\left(\mathbf{J}_{\bar{\theta}}^{-1}\right) = \frac{1}{2}\mathrm{tr}\left(\mathbb{K}_{\bar{\theta}}\right) - \frac{1}{2}\mathrm{tr}\left(\mathbb{R}_{\bar{\theta}}^{-1} \begin{pmatrix} \mathbf{0} & \mathbf{h}^H \\ \mathbf{h}^T & \mathbf{0} \end{pmatrix} \mathbb{K}_{\bar{\theta}} \mathbb{K}_{\bar{\theta}} \begin{pmatrix} \mathbf{h}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{h} \end{pmatrix}\right).$$

Simplifying the preceding expression, it can be demonstrated that the MSE lower bound for the estimation of $\mathbf{h}$ is given as

$$\mathrm{E}\left\{\left\|\hat{\mathbf{h}} - \mathbf{h}\right\|^2\right\} \geqslant \frac{\mathrm{tr}\left(\mathbf{R}_e\right)}{\alpha + \beta \|\mathbf{h}\|^2} + \frac{\left(\mathbf{h}^H \mathbf{R}_e \mathbf{h}\right)}{\left(\alpha + \beta \|\mathbf{h}\|^2\right)} \frac{\left(\mathbf{h}^H \mathbf{R}_e \mathbf{R}_e \mathbf{h}\right)}{\left(\alpha + \beta \|\mathbf{h}\|^2\right)^2 \left|\mathbb{R}_{\bar{\theta}}\right|}, \tag{35}$$

DOI: 10.1002/dac

where $\left|\mathbb{R}_{\bar{\theta}}\right|$ is the determinant of the matrix $\mathbb{R}_{\bar{\theta}}$ and is given as

$$\left|\mathbb{R}_{\bar{\theta}}\right| = \frac{1}{\theta^2} - \frac{\left(\mathbf{h}^H \mathbf{R}_e \mathbf{h}\right)\left(\mathbf{h}^T \mathbf{R}_e \mathbf{h}^*\right)}{\left(\alpha + \beta \left\|\mathbf{h}\right\|^2\right)^2} = \frac{\left(\sigma_n^2 + P_d^s \left\|\mathbf{h}\right\|^2\right)^4}{\left(N_b \left(P_d^s\right)^2\right)^2} - \frac{\left\|\mathbf{h}\right\|^4 \left(\sigma_n^2 + P_d^s \left\|\mathbf{h}\right\|^2\right)^2}{\left(\alpha + \beta \left\|\mathbf{h}\right\|^2\right)^2}$$

$$= \gamma^2 \left(\frac{1}{\beta^2} - \frac{\left\|\mathbf{h}\right\|^4}{\left(\alpha + \beta \left\|\mathbf{h}\right\|^2\right)^2}\right) = \frac{\alpha\gamma^2 \left(\alpha + 2\beta \left\|\mathbf{h}\right\|^2\right)}{\beta^2 \left(\alpha + \beta \left\|\mathbf{h}\right\|^2\right)^2}$$

At high SNR, that is, as $P_d^s \to \infty$, it can be seen that $\lim_{P_d^s \to \infty} \left|\mathbb{R}_{\bar{\theta}}\right| = \frac{2\alpha\gamma^2}{\beta^2 \|\mathbf{h}\|^2}$. It can be observed that $\mathbf{h}^H \mathbf{R}_e \mathbf{R}_e \mathbf{h} \to \left(P_d^s\right)^2 \|\mathbf{h}\|^6$ and $\frac{\left(\mathbf{h}^H \mathbf{R}_e \mathbf{h}\right)}{\left(\alpha + \beta \|\mathbf{h}\|^2\right)} \to \frac{P_d^s \|\mathbf{h}\|^2}{\beta}$ as $P_d^s \to \infty$. Substituting these and the expression for $\left|\mathbb{R}_{\bar{\theta}}\right|$ from preceding text in the MSE expression in (35), the high SNR CRB asymptote is obtained as

$$\mathrm{MSE}_b^{\infty} = P_d^s \left(\lim_{P_d^s \to \infty} \frac{\mathrm{MSE}_b}{P_d^s}\right) = P_d^s \left(\lim_{P_d^s \to \infty} \frac{1}{P_d^s} \left\{\frac{\|\mathbf{h}\|^2}{N_b} + \frac{\|\mathbf{h}\|^2 P_d^s}{2N_b \sigma_p^2}\right\}\right) = \frac{\|\mathbf{h}\|^2 P_d^s}{2N_b \sigma_p^2}.$$

APPENDIX C: PROOF OF THEOREM 2

The capacity of the communication channel of (23) with uncorrelated Gaussian noise is given by the well-known maximization of mutual information [5, 46]. When the nature of the noise process $\mathbf{v}(k)$ is unknown, the worst-case capacity [30] can be expressed as

$$C_w = \min_{p_{\mathbf{v}}(\cdot),\, \mathrm{tr}(\mathbf{R}_v) = r\sigma_n^2} \; \max_{p_{\mathbf{s}}(\cdot),\, \mathrm{tr}(\mathbf{R}_x) = t\sigma_s^2} \mathbf{I}\left(\mathbf{y}; \mathbf{s}\right)$$

The system in (25) can be equivalently written as

$$\mathbf{y}(k) \equiv \left(\mathbf{I} + \mathbf{R}_{vs}\mathbf{R}_s^{-1}\right)\mathbf{s}(k) + \tilde{\mathbf{v}}(k),$$

where $\mathbf{v} \triangleq \tilde{\mathbf{v}} + \mathbf{R}_{vs}\mathbf{R}_s^{-1}\mathbf{s}$, and the innovations noise $\tilde{\mathbf{v}}$ is uncorrelated with the source $\mathbf{s}$, that is, $\mathrm{E}\left\{\tilde{\mathbf{v}}\mathbf{s}^H\right\} = \mathbf{0}$. The covariance $\mathbf{R}_{\tilde{v}}$ is given as $\mathbf{R}_{\tilde{v}} = \mathbf{R}_{v|s} = \mathbf{R}_v - \mathbf{R}_{vs}\mathbf{R}_s^{-1}\mathbf{R}_{sv}$. It can be seen that the transformation,

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_r & \mathbf{R}_{vs}\mathbf{R}_s^{-1} \\ \mathbf{0}_{r\times r} & \mathbf{I}_r \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}} \\ \mathbf{s} \end{bmatrix} \tag{36}$$

is invertible (because the transform matrix is upper triangular). Therefore, given a distribution function $p_{\tilde{\mathbf{v}},\mathbf{s}}(\cdot)$, there exists a distribution $p_{\mathbf{v},\mathbf{s}}(\cdot)$ and vice versa. Hence, it can be seen that,

$$\min_{p_{\mathbf{v}}(\cdot),\, \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H\right\} = \mathbf{R}_v} \; \max_{p_{\mathbf{s}}(\cdot),\, \mathrm{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{R}_s} \mathbf{I}\left(\mathbf{y}; \mathbf{s}\right) = \min_{p_{\tilde{\mathbf{v}}}(\cdot),\, \mathrm{E}\left\{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H\right\} = \mathbf{R}_{\tilde{v}}} \; \max_{p_{\mathbf{s}}(\cdot),\, \mathrm{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{R}_s} \mathbf{I}\left(\mathbf{y}; \mathbf{s}\right),$$

Now, employing the result for worst-case capacity with uncorrelated noise from [30], the worst-case capacity for the preceding system can be seen to be given as

$$C_w = \min_{\mathrm{tr}(\mathbf{R}_v) = r\sigma_n^2} \; \max_{\mathrm{tr}(\mathbf{R}_x) = tP_d} \log\left|\mathbf{I} + \mathbf{R}_{\tilde{v}}^{-1}\left(\mathbf{I} + \mathbf{R}_{vs}\mathbf{R}_s^{-1}\right)\mathbf{R}_s\left(\mathbf{I} + \mathbf{R}_{vs}\mathbf{R}_s^{-1}\right)^H\right|,$$

$$= \min_{\mathrm{tr}(\mathbf{R}_v) = r\sigma_n^2} \; \max_{\mathrm{tr}(\mathbf{R}_x) = tP_d} \log\left|\mathbf{I} + \mathbf{R}_{v|s}^{-1}\left(\mathbf{R}_s + \mathbf{R}_{vs}\right)\mathbf{R}_s^{-1}\left(\mathbf{R}_s + \mathbf{R}_{vs}\right)^H\right|,$$

which is the expression for the worst-case capacity in (26).

DOI: 10.1002/dac

APPENDIX D: MINIMUM VARIANCE DISTORTIONLESS RESPONSE-POST-PROCESSING SIGNAL-TO-NOISE POWER RATIO

In succeeding text, the expression in Equation (20) for the MVDR PSNR κ_m is derived. From Table I, the covariance of the effective noise $\mathbf{R}_v^s = \beta_h \mathbf{h}\mathbf{h}^H + \beta_n \mathbf{I}$, where the constants β_h, β_n are defined as $\beta_h \triangleq \frac{P_d^s}{N_b}\left(1 + \frac{P_d^s}{P_t^s}\right)$ and $\beta_n \triangleq \frac{\sigma_n^2}{N_b}\left(1 + \frac{P_d^s}{P_t^s}\right) + \sigma_n^2$. Using results on matrix inversion [41], the matrix $\left(\mathbf{R}_v^s\right)^{-1}$ can be expressed as

$$\left(\mathbf{R}_v^s\right)^{-1} = \frac{1}{\beta_h}\left(\frac{\beta_h}{\beta_n}\mathbf{I} - \frac{\beta_h}{\beta_n}\mathbf{I}\mathbf{h}\left(1 + \mathbf{h}^H \frac{\beta_h}{\beta_n}\mathbf{I}\mathbf{h}\right)^{-1}\frac{\beta_h}{\beta_n}\mathbf{I}\right) = \frac{1}{\beta_n}\mathbf{I} - \left(\frac{\beta_h}{\beta_n}\right)\frac{\mathbf{h}\mathbf{h}^H}{\beta_n + \beta_h \|\mathbf{h}\|^2}. \quad (37)$$

Substituting this expression for $\left(\mathbf{R}_v^s\right)^{-1}$ in (19), the expression for κ_m can be simplified as

$$\kappa_m = P_d^s \hat{\mathbf{h}}_s^H \left(\mathbf{R}_v^s\right)^{-1} \hat{\mathbf{h}}_s \approx P_d^s \mathbf{h}^H \left(\mathbf{R}_v^s\right)^{-1} \mathbf{h} = \frac{P_d^s}{\beta_n}\|\mathbf{h}\|^2 - \left(\frac{\beta_h}{\beta_n}\right)\frac{P_d^s \|\mathbf{h}\|^4}{\beta_n + \beta_h \|\mathbf{h}\|^2} = \frac{P_d^s \|\mathbf{h}\|^2}{\beta_n + \beta_h \|\mathbf{h}\|^2}.$$

Substituting the expressions for β_h, β_n defined previously, the final expression for κ_m in terms of the quantities P_d^s, P_t^s is obtained as

$$\kappa_m = \frac{N_b \|\mathbf{h}\|^2 P_d^s P_t^s}{\|\mathbf{h}\|^2 P_d^s \left(P_d^s + P_t^s\right) + \sigma_n^2 \left(P_t^s (1 + N_b) + P_d^s\right)},$$

which reduces to the expression in (20).

ACKNOWLEDGEMENTS

This work was supported by UC Discovery grant com04-10176 and com04-10173.

REFERENCES

1. Bhunia S, Misra S, Sanyal SK, Kundu A. Performance study of mobile WiMAX network with changing scenarios under different modulation and coding. *International Journal of Communication Systems* 2011; **24**:1087–1104.
2. Rim M. An irregular frequency reuse scheme for cellular OFDMA systems. *International Journal of Communication Systems* 2011; **24**:526–531.
3. Wang H, Jia W. An optimized scheduling scheme in OFDMA WiMax networks. *International Journal of Communication Systems* 2010; **23**(1):23–39.
4. Campos R, Ricardo M. An efficient mechanism for establishing IP connectivity in next-generation networks. *International Journal of Communication Systems* 2010; **23**:817–840.
5. Telatar E. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications* 1999; **10**(6):585–596.
6. Tse D, Viswanath P. *Fundamentals of Wireless Communication (Hardcover)*, 1st edition. Cambridge University Press: West Nyack, New York, 2005.
7. Biglieri E, Calderbank R, Constantinides A, Goldsmith A, Paulraj A, Poor HV. *MIMO Wireless Communications*. Cambridge University Press: West Nyack, New York, 2010.
8. Yu X, Chen X, Bi G. Multiuser receiver scheme for full-rate spacetime coded CDMA system with imperfect estimation. *International Journal of Communication Systems* 2011; **24**:295–309.
9. Zhang W, Guan Z, Zhou X. Analysis of SER of MIMO-MRC systems with imperfect channel estimation in the presence of non-rayleigh CCIs. *International Journal of Communication Systems* 2011; **24**:889–901.
10. Mousa A, Mahmoud H. Reducing ICI effect in OFDM system using low-complexity Kalman filter based on comb-type pilots arrangement. *International Journal of Communication Systems* 2011; **24**:53–61.
11. Tong L, Perreau S. Multichannel blind identification: from subspace to maximum likelihood methods. *Proceedings of the IEEE* October 1998; **86**(10):1951–1968. DOI: 10.1109/5.720247.
12. Hua Y. Fast maximum-likelihood for blind identification of multiple FIR channels. *IEEE transactions on Signal Processing* 1996; **44**(3):661–672.
13. Tu C, Champagne B. Subspace-based blind channel estimation for MIMO–OFDM systems with reduced time averaging. *IEEE Transactions on Vehicular Technology* 2010; **59**(3):1539–1543.
14. Carvalho E, Slock DTM. Blind and semi-blind FIR multichannel estimation: Global identifiability conditions. *IEEE Transactions on Signal Processing* 2004; **50**(8):1053–1064.

DOI: 10.1002/dac

15. Wang S, Hu J. Blind channel estimation for single-input multiple-output OFDM systems: zero padding based or cyclic prefix based? *Wireless Communications and Mobile Computing* 2011. DOI: 10.1002/wcm.1109.
16. Slock DTM. Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas. *Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers* 1994; **2**:1154–1158.
17. Mazzenga F. Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence. *IEEE Transactions on Broadcasting* 2000; **46**(2):170–176.
18. Farhang-Boroujeny B. Pilot-based channel identification: proposal for semi-blind identification of communication channels. *Electronics Letters* 1995; **31**(13):1044–1046.
19. Zhou GT, Viberg M, McKelvey T. Superimposed periodic pilots for blind channel estimation. *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, 2001*, Vol. 1, Asilomar Conference Grounds, Pacific Grove, CA, USA, November 2001; 653–657.
20. Tugnait JK, Luo W. On channel estimation using superimposed training and first-order statistics. *IEEE Communications Letters* 2003; **7**(9):413–415.
21. Wang G, Gao F, Tellambura C. Superimposed Pilots Aided Joint CFO and Channel Estimation for ZP-OFDM Modulated Two-Way Relay Networks. *2010 IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, Miami, Florida, USA, September 2010; 1–5.
22. Levanen T, Talvitie J, Renfors M. Improved performance analysis for superimposed pilot based short channel estimator. *2010 IEEE Eleventh International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Marrakech, Morocco, 2010; 1–6.
23. Cui T, Tellambura C. OFDM channel estimation and data detection with superimposed pilots. *European Transactions on Telecommunications* 2011; **22**(3):125–136.
24. Chen N, Zhou GT. A superimposed periodic pilot scheme for semi-blind channel estimation of OFDM systems. *Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop*, Pine Mountain, GA, October 2002; 362–365.
25. Pal D. Fractionally spaced semi-blind equalization of wireless channels. *The Twenty-Sixth Asilomar Conference* 1992; **2**:642–645.
26. Jagannatham AK, Rao BD. Whitening rotation based semi-blind MIMO channel estimation. *IEEE Transactions on Signal Processing* 2006; **54**(3):861–869.
27. Tugnait JK, Meng X. On superimposed training for channel estimation: performance analysis, training power allocation, and frame synchronization. *IEEE Transactions on Signal Processing* 2006; **54**(2):752–765.
28. Bohlin P, Tapio M. Performance evaluation of MIMO communication systems based on superimposed pilots. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004 (ICASSP '04)*, Montreal, Quebec, Canada, 2004; 425–428.
29. Coldrey M. *Ph.D. Thesis: Estimation and Performance Analysis for Wireless Multiple Antenna Communication Channels*. Chalmers University of Technology: Goteborg, Sweden, 2006.
30. Hassibi B, Hochwald B. How much training is needed in multiple-antenna wireless links. *IEEE Transactions on Information Theory* 2003; **49**(10):951–964.
31. Xiong N, Cao M, Vasilakos AV, Yang LT, Yang F. An energy-efficient scheme in next-generation sensor networks. *International Journal of Communication Systems* 2010; **23**(9):1189–1200.
32. Fang S, Berber SM, Swain AK. Energy distribution-aware clustering algorithm for dense wireless sensor networks. *International Journal of Communication Systems* 2010; **23**(9):1223–1251.
33. Orozco-Lugo AG, Lara MM, McLernon DC. Channel estimation using implicit training. *IEEE Transactions on Signal Processing* 2004; **52**(1):240–254.
34. Kay SM. *Fundamentals of Statistical Signal Processing, Vol I: Estimation Theory*, First edition. Prentice Hall, Inc.: Upper Saddle River, NJ, USA, 1993.
35. Gray RM, Davisson LD. *An Introduction to Statistical Signal Processing*, first edition. Cambridge University Press: West Nyack, New York, 2010.
36. Biguesh M, Gershman AB. Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals. *IEEE Transactions on Signal Processing* 2006; **54**(3):884–893.
37. Dong M, Tong L. Optimal design and placement of pilot symbols for channel estimation. *IEEE Transactions on Signal Processing* 2002; **50**(12).
38. Jagannatham AK, Rao BD. Cramer–Rao lower bound for constrained complex parameters. *IEEE Signal Processing Letters* 2004; **11**(11):875–878.
39. Jagannatham AK, Rao B. Fisher-information-matrix based analysis of semiblind MIMO frequency selective channel estimation. *ISRN Signal Processing* 2011; **2011**. DOI: 10.5402/2011/758918.
40. Miller KS. *Complex Stochastic Processes: An Introduction to Theory and Application*, first edition. Addison-Wesley, Reading: MA, USA, 1974.
41. Moon TK, Stirling WC. *Mathematical Methods and Algorithms for Signal Processing*, first edition. Prentice Hall: Upper Saddle River, NJ, USA, 1999.
42. Li J, Stoica P, Wang Z. Doubly constrained robust capon beamformer. *IEEE Transactions on Signal Processing* 2004; **52**(9):2407–2423.
43. Stoica P, Wang Z, Li J. Robust capon beamforming. *IEEE Signal Processing Letters* 2003; **10**(6):172–175.
44. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press: NY, USA, 2004.

45. Trees HLV. *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*. Wiley Interscience: New York, 2002.
46. Cover TM, Thomas JA. *Elements of Information Theory*. John Wiley & Sons, Inc.: N.Y., 1991.

AUTHORS' BIOGRAPHIES

Aditya K. Jagannatham received his bachelors degree from the Indian Institute of Technology, Bombay, and MS and PhD degrees from the University of California, San Diego, USA. From April 2007 to May 2009, he was employed as a senior wireless systems engineer at Qualcomm Inc., San Diego, California, where he worked on developing 3G UMTS/WCDMA/HSDPA mobile chipsets as part of the Qualcomm CDMA technologies division. His research interests are in the area of next-generation wireless communications and networking, sensor and ad hoc networks, digital video processing for wireless systems, wireless 3G/4G cellular standards, and CDMA/OFDM/MIMO wireless technologies. He has contributed to the 802.11n high-throughput wireless LAN standard and has published extensively in leading international journals and conferences. He was awarded the CAL(IT)2 fellowship for pursuing graduate studies at the University of California, San Diego, and in 2009 he received the Upendra Patel Achievement Award for his efforts towards developing HSDPA/HSUPA/HSPA+ WCDMA technologies at Qualcomm. Since 2009, he has been a faculty member in the Electrical Engineering Department at IIT Kanpur and is also associated with the BSNL-IITK Telecom Center of Excellence (BITCOE) at IIT Kanpur. At IIT Kanpur he has been awarded the P.K. Kelkar Young Faculty Research Fellowship (June 2012 to May 2015) for excellence in research.

Bhaskar D. Rao received his B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, in 1979, and his MS and PhD degrees from the University of Southern California, Los Angeles, in 1981 and 1983, respectively. Since 1983, he has been with the University of California at San Diego (UCSD), La Jolla, where he is currently a Professor with the Electrical and Computer Engineering Department. His interests are in the areas of digital signal processing, estimation theory, and optimization theory, with applications to digital communications, speech signal processing, and human-computer interactions.

Dr. Rao holds the Ericsson endowed chair in Wireless Access Networks and is the Director of the Center for Wireless Communications. His research group has received several paper awards. His paper received the Best Paper Award at the 2000 Speech Coding Workshop, and his students have received student paper awards at both the 2005 and 2006 International Conference on Acoustics, Speech and Signal Processing Conference, as well as the Best Student Paper Award at NIPS 2006. A paper he coauthored with B. Song and R. Cruz received the 2008 Stephen O. Rice Prize Paper Award in the Field of Communications Systems, and a paper he coauthored with S. Shivappa and M. Trivedi received the Best Paper Award at AVSS 2008. He also received the Graduate Teaching Award from the graduate students in the Electrical Engineering Department, UCSD, in 1998. He was elected an IEEE Fellow in 2000 for his contributions in high-resolution spectral estimation. He has been a member of the Statistical Signal and Array Processing Technical Committee, the Signal Processing Theory and Methods Technical Committee, and the Communications Technical Committee of the IEEE Signal Processing Society. He has also served on the editorial board of the EURASIP *Signal Processing Journal*.

DOI: 10.1002/dac