On Combined Beamforming and OSTBC over the Cognitive Radio S-Channel with Partial CSI

Efthymios Stathakis, Mikael Skoglund, and Lars K. Rasmussen
School of Electrical Engineering and ACCESS Linnaeus Center
KTH - Royal Institute of Technology, Stockholm, Sweden
emails:{eftsta,skoglund,lkra}@kth.se

***Abstract*—A pair of secondary (cognitive, unlicensed) users is communicating in the presence of multiple primary (licensed) user pairs. The cognitive transceiver is implementing beamforming and orthogonal space-time block coding in the presence of external interference, induced by the primary system transmission, that has to be properly handled. Moreover, the cognitive link nodes, both the transmitter and the receiver, are supplied with different levels of network side information (NSI), i.e. primary messages and partial channel side information (CSI). We investigate how this side information can be taken into account in the cognitive system design and how interference affects the behaviour of the beamforming solution. Through numerical simulations, we illustrate the impact of partial CSI on system performance and discuss its implications on the feasibility of cognitive systems.**

I. Introduction

The demand for higher data rates and ubiquitous coverage has been following an increasing trend over the last decade. Provision of high Quality-of-Service (QoS) for emerging applications requires, among others, bandwidth efficient wireless technology. However, rigid regulation of spectrum allocation policies has hindered efforts in this direction. A shift came with the introduction of cognitive radios (CRs) [1]; systems which sense their surrounding environment and dynamically adapt to changes in the wireless channel, promising to bridge this gap in spectrum underutilization.

Since the introduction of the CR paradigm, signal processing frameworks for cognitive operation have been defined, see [2], and fundamental information-theoretic limits have been explored [3]. A more elaborate definition of possible architectures for CR networks (CRNs), namely the interweave, the underlay and the overlay CRNs, was provided in [4]. Interweave CRNs support cognitive operation exclusively over spectrum holes (idle frequency bands) whereas underlay and overlay CRNs realize the spectrum sharing (SS) principle, by allowing simultaneous transmission of both licensed and unlicensed systems. The latter two architectures are more spectrally efficient and, in terms of CR feasibility, overlay systems are even more favorable since the underlying set of assumptions allows for advanced signal processing on the cognitive link side. Therefore, studying overlay CRNs provides an indication of the best achievable performance in a cognitive network setup.

This work has been partially funded through the EU FP7 project INFSO-ICT-248303 QUASAR. This work has further been supported in parts by the [illegible] 4666.

CROWNCOM 2012, June 18-20, Stockholm, Sweden

DOI 10.4108/icst.crowncom.2012.248344

A promising technology, in the aforementioned direction, are multiple-input multiple-output (MIMO) systems [5], a paradigm which led to the introduction of new research directions; beamforming [6], for transmission calibration, and orthogonal space-time block coding (OSTBC), for diversity gain [7]. The advantages stemming from the combination of beamforming and OSTBC, termed as BOSTBC, were demonstrated in the pioneering work of [8]. There it was shown that BOSTBC achieves very good performance in the presence of partial channel state information (CSI) and always outperforms each individual strategy.

In the context of overlay CRNs, the investigation of BOSTBC requires the integration of external incoming interference in the design. Important work extending the design of [8] in other directions appears in [9]-[11]. However, in all those considered cases, the system was not subject to interference and the only source of disturbance has been white noise. A different treatise, similar to our objective from a mathematical viewpoint, is found in [12], where space-time block coding design was considered in the case of correlated noise, yet the system was not implementing beamforming.

In this work, we investigate the BOSTBC design when the operating primary systems are causing interference on the receiver side of the cognitive link. The target is to characterize and evaluate how the system behaves in the presense of the interfering cross-links, for which only partial CSI is available. Under a statistical channel uncertainty model, we describe how primary interference affects the behavior of the optimal beamforming solutions. Moreover, we numerically evaluate the induced performance back-off, from single-user transmission, due to imperfect side information and draw conclusions regarding the feasibility of BOSTBC in an overlay CRN.

This paper is organized as follows. In Section II we introduce the system model under consideration whereas the design criterion and the associate mathematical analysis of our system are provided in Section III. Numerical evaluation of our system is provided in Section IV and finally Section V summarizes and highlights the most important parts of this work.

Notation: We use bold lower case and bold capital case letters for vectors and matrices. The operators $(\cdot)^\star, (\cdot)^T$ and $(\cdot)^H$ denote conjugate, transpose and conjugate-transpose, respectively. The N-length all-zero, all-one vectors and the $N \times N$ zero, identity matrices are denoted by $\mathbb{0}_N, \mathbb{1}_N$ and $\mathbf{0}_N, \mathbf{I}_N$. Finally, the operators $\mathrm{vec}(\cdot), \mathrm{tr}(\cdot), \|\cdot\|_2$ denote vectorized form, trace and Frobenius norm of a matrix.

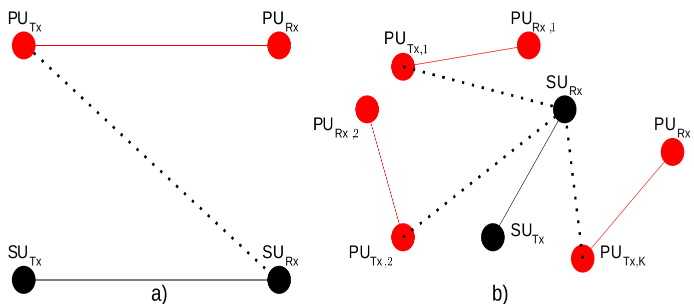

Fig. 1. a) The basic CSC, and b) the general network setup.

II. System Model

In this section we describe the system under consideration, along with the operational assumptions. In particular, we provide in what follows, the network model, its associate channel model and the model that captures CSI quality.

A. Network Model

The cognitive S-channel (CSC), depicted in Figure 1a, is a subset of the general cognitive interference channel (CIC), with a cognitive and a primary transmit-receive pair denoted as $\mathrm{SU_{Tx}}-\mathrm{SU_{Rx}}$ and $\mathrm{PU_{Tx}}-\mathrm{PU_{Rx}}$, respectively. In the CSC setup the fundamental assumption is that the cognitive transmission is not disturbing the licensed system. This could be attributed to a number of factors, e.g. low secondary transmit power, large $\mathrm{SU_{Tx}} - \mathrm{PU_{Rx}}$ distance, obstacles, etc. Therefore, the goal, in this context, is to robustify the $\mathrm{SU_{Tx}} - \mathrm{SU_{Rx}}$ transmission strategy against external interference or simply render communication over the cognitive link feasible and reliable.

In order to properly design our system against primary interference we need some level of side information. Detection at the receiver and transmit precoding require the availability of non-causal primary message knowledge and partial CSI of the primary interfering cross-links, at both nodes of the cognitive link. The $\mathrm{SU_{Tx}}$ node has partial CSI of the $\mathrm{SU_{Tx}} - \mathrm{SU_{Rx}}$ link whereas the receiver $\mathrm{SU_{Rx}}$ is performing ML-detection assuming perfect knowledge of the cognitive link channel.

B. Channel Model

Assume K primary user pairs and a single secondary link, as in Figure 1b. The secondary transmitter and the j-th primary transmitter are supplied with M, M_j antennas whereas their associated receivers are supplied with N, N_j antennas, respectively. We use, for consistency with [8], $\mathbf{H} \in \mathbb{C}^{M\times N}$ to denote the secondary link channel, so that the element $(\cdot)^H_{ij}$ represents the channel between transmit antenna i and receive antenna j. The statistics of $\mathbf{H}$ are captured by $\mathbf{h} = \mathrm{vec}(\mathbf{H}^H)$ which is distributed as $\mathbf{h} \sim \mathbb{C}\mathcal{N}(\mathbf{m}, \mathbf{K})$. In the same spirit we define the $\mathrm{PU_{Tx,\mathit{j}}} - \mathrm{SU_{Rx}}$ cross-channel as $\mathbf{F}_j \in \mathbb{C}^{M_j\times N}$ and $\mathbf{f}_j = \mathrm{vec}(\mathbf{F}_j^H)$, with $\mathbf{f}_j \sim \mathbb{C}\mathcal{N}(\mathbf{m}_j, \mathbf{K}_{jj})$. The covariance matrices $\mathbf{K}, \mathbf{K}_{jj}$ have a Kronecker product structure [13], where transmit-receive covariance is separable, i.e. $\mathbf{K} = \mathbf{K}_T \otimes \mathbf{K}_R$, $\mathbf{K}_{jj} = \mathbf{K}_{T,jj} \otimes \mathbf{K}_{R,jj}$, with $\mathbf{K}_T \in \mathbb{C}^{M\times M}$, $\mathbf{K}_{T,jj} \in \mathbb{C}^{M_j\times M_j}$ and $\mathbf{K}_R, \mathbf{K}_{R,jj} \in \mathbb{C}^{N\times N}$. Moreover, we assume that all covariance matrices are non-singular.

Our study focuses on the design of the cognitive link transmission strategy, which takes place in blocks of L time-slots. During each block the $\mathrm{SU_{Tx}}$ terminal is transmitting a matrix $\mathbf{C} \in \mathbb{C}^{M\times L}$, of the form $\mathbf{C} = \mathbf{W}\bar{\mathbf{C}}$, over the channel. The matrix $\bar{\mathbf{C}} \in \mathbb{C}^{M\times L}$ is a mapping $\mathbb{D}$ of the actual information vector $\mathbf{c}$ onto an OSTBC which is weighted, prior to transmission, by some beamforming matrix $\mathbf{W} \in \mathbb{C}^{M\times M}$, whose choice relies on some criterion. Under these premises the signal at the cognitive receiver can be written as

$$\mathbf{Y} = \mathbf{H}^H\mathbf{C} + \sum_{j=1}^{K} \mathbf{F}_j^H \widetilde{\mathbf{C}}_j + \mathbf{N}, \tag{1}$$

where $\mathbf{N}$ denotes complex i.i.d Gaussian noise with variance σ_n^2 and whose statistics are described by $\mathbf{n} = \mathrm{vec}(\mathbf{N})$ as $\mathbf{n} \sim \mathbb{C}\mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I}_{NL})$. The summation term, in (1), corresponds to the interference, generated by the primary transmitters. Note here that, for consistency with the rest of our setup, the primary messages $\widetilde{\mathbf{C}}_j$ are also assumed to be OSTBCs.

C. Statistical Uncertainty CSI Model

We employ a statistical model in order to characterize the uncertainty around the actual channel. We assume that the actual channel $\mathbf{H}$ is jointly Gaussian distributed with the available channel estimate $\hat{\mathbf{H}}$, the latter statistically described by its vectorized counterpart $\hat{\mathbf{h}} = \mathrm{vec}(\hat{\mathbf{H}}^H)$ as $\hat{\mathbf{h}} \sim \mathbb{C}\mathcal{N}(\hat{\mathbf{m}}, \hat{\mathbf{K}})$. The joint distribution of $\mathbf{h}, \hat{\mathbf{h}}$ can be completetely characterized by the statistics of the vector $\mathbf{h}^{\mathrm{c}} = [\mathbf{h}^T \hat{\mathbf{h}}^T]^T$, with $\mathbf{h}^{\mathrm{c}} \sim \mathbb{C}\mathcal{N}(\mathbf{m}^{\mathrm{c}}, \mathbf{K}^{\mathrm{c}})$. The quantity which captures the uncertainty is $\tilde{\mathbf{h}} = \mathbf{h}|\hat{\mathbf{h}}$, i.e. the conditional distribution of $\mathbf{h}$ given $\hat{\mathbf{h}}$, which is Gaussian with mean $\tilde{\mathbf{m}}$ and covariance $\tilde{\mathbf{K}}$, that can be easily calculated from $\mathbf{m}^{\mathrm{c}}, \mathbf{K}^{\mathrm{c}}$ and the realization of $\hat{\mathbf{h}}$ [14].

The conditional mean $\tilde{\mathbf{m}}$ is a function of the current channel estimate $\hat{\mathbf{h}}$ whereas the conditional covariance matrix $\tilde{\mathbf{K}}$ is a statistical measure of the distance between the estimate and the actual channel. Therefore if we relate the statistics of $\mathbf{h}$ and $\hat{\mathbf{h}}$ through a single scalar parameter δ then we can model the two asymptotic CSI cases as: a) No CSI: $\delta \to 0, \tilde{\mathbf{m}} \to \mathbb{0}_{MN}, \tilde{\mathbf{K}} \to \mathbf{K}$ and b) Full CSI: $\delta \to 1, \tilde{\mathbf{m}} \to \mathbf{h}, \tilde{\mathbf{K}} \to \mathbf{0}_{MN}$. Similarly, as above, we define $\delta_j, \tilde{\mathbf{m}}_j, \tilde{\mathbf{K}}_{jj}$ in order to describe the statistics of $\tilde{\mathbf{f}}_j = \mathbf{f}_j|\hat{\mathbf{f}}_j$, which is modelling the uncertainty of the j-th cross-channel estimate $\hat{\mathbf{f}}_j \sim \mathbb{C}\mathcal{N}(\hat{\mathbf{m}}_j, \hat{\mathbf{K}}_{jj})$, around $\mathbf{f}_j$. Note here that, our no CSI definition is implicitly assuming knowledge of the primary system geometry, which is a prerequisite in a CRN.

III. BOSTC for the Cognitive S-Channel

In this section, we describe the extension of [8] to the system model described in Section II. We briefly go through the performance criterion that dictates the design and further analyze the system behavior under some special cases of the CSI quality and covariance matrix model.

A. Performance Criterion

In order to analyze our system it will be convenient to start by rewriting (1) in its equivalent vectorized form as

$$\mathbf{y} = \left(\mathbf{C}^T \otimes \mathbf{I}_N\right)\mathbf{h} + \sum_{j=1}^{K} \left(\widetilde{\mathbf{C}}_j^T \otimes \mathbf{I}_N\right)\mathbf{f}_j + \mathbf{n}. \tag{2}$$

For ease of exposition we define the interference-plus-noise vector $\mathbf{e} = \sum_{j=1}^{K}\left(\widetilde{\mathbf{C}}_j^T \otimes \mathbf{I}_N\right)\mathbf{f}_j + \mathbf{n}$ but we are, in fact, interested in finding the mean $\tilde{\mathbf{m}}_{\mathbf{e}}$ and the covariance $\tilde{\mathbf{K}}_{\mathbf{ee}}$ of the conditional distribution $\tilde{\mathbf{e}} = \mathbf{e}|\{(\hat{\mathbf{f}}_j, \tilde{\mathbf{C}}_j)\}_{j=1}^K$. By Lemma 1 (Appendix A) it follows that its pdf $\mathrm{p}_{\tilde{\mathbf{e}}}$ can be expanded as

$$\mathrm{p}_{\mathbf{e}|\{(\hat{\mathbf{f}}_j,\tilde{\mathbf{C}}_j)\}_{j=1}^K} = \mathrm{p}_{\mathbf{n}} \prod_{j=1}^K \mathrm{p}_{\mathbf{C}'_j \mathbf{f}_j|\hat{\mathbf{f}}_j} = \mathrm{p}_{\mathbf{n}} \prod_{j=1}^K \mathrm{p}_{\mathbf{C}'_j \tilde{\mathbf{f}}_j},$$

where $\mathbf{C}'_j = \tilde{\mathbf{C}}_j^T \otimes \mathbf{I}_N$, rendering the calculation of the desired quantities straightforward. Given the statistics of $\tilde{\mathbf{e}}$ we can calculate the pairwise error probability $\mathrm{P}_{kl} = \mathrm{P}(\mathbf{C}_k \to \mathbf{C}_l, |\mathbf{C}_k, \mathbf{h}, \mathbf{\Phi})$, i.e. the probability of erroneously deciding in favor of $\mathbf{C}_l$, given that $\mathbf{C}_k$ was transmitted and side information, described by the set $\mathbf{\Phi} = \{\hat{\mathbf{h}}, (\hat{\mathbf{f}}_1, \tilde{\mathbf{C}}_1), \ldots, (\hat{\mathbf{f}}_K, \tilde{\mathbf{C}}_K)\}$, is available at the transmitter. It is then easy to show that [15]

$$\begin{aligned}\mathrm{P}_{k,l} = Q\Big\{\Big[\tfrac{1}{2}\mathbf{h}^H\Big((\Delta\mathbf{C}_{kl}^\star \otimes \mathbf{I}_N)\mathbf{K}_{ee}^{-1} \\ (\Delta\mathbf{C}_{kl}^T \otimes \mathbf{I}_N)\Big)\mathbf{h}\Big]^{\frac{1}{2}}\Big\} \\ \le e^{-\mathbf{h}^H\mathbf{A}(\mathbf{Q},\mathbf{W})\mathbf{h}}\end{aligned} \quad (3)$$

In order to simplify the expressions in (3), we have defined the quantities $\Delta\mathbf{C}_{kl} = \mathbf{W}(\bar{\mathbf{C}}_k - \bar{\mathbf{C}}_l) = \mathbf{W}\Delta\bar{\mathbf{C}}_{kl}$ and $\mathbf{Q} = (\Delta\bar{\mathbf{C}}_{kl} \otimes \mathbf{I}_N)\tilde{\mathbf{K}}_{ee}^{-\star}(\Delta\bar{\mathbf{C}}_{kl}^H \otimes \mathbf{I}_N)$. The last inequality follows from the well-known upper bound on the Q-function and upon defining $\mathbf{A}(\mathbf{Q},\mathbf{W}) = \frac{1}{4}[(\mathbf{W}\otimes\mathbf{I}_N)\mathbf{Q}(\mathbf{W}^H \otimes \mathbf{I}_N)]^\star$ in order to express the exponential term in a compact form.

The next step is to average the pairwise error probability P_{kl} over all channel realizations. Therefore, we need to integrate both sides of (3), using the conditional marginal distribution $\mathrm{p}_{\mathbf{h}|\hat{\mathbf{h}}}(\cdot|\cdot)$, as

$$\mathrm{P}(\mathbf{C}_k \to \mathbf{C}_l, |\mathbf{C}_k, \mathbf{\Phi}) \le \int e^{-\mathbf{h}^H\mathbf{A}(\mathbf{Q},\mathbf{W})\mathbf{h}}\mathrm{p}_{\mathbf{h}|\hat{\mathbf{h}}}\mathrm{d}\mathbf{h}. \quad (4)$$

Following the lines of [8] it turns out that minimizing the right hand side (RHS) of (4) is equivalent to maximizing the following function $l(\mathbf{W}, \Delta\bar{\mathbf{C}}_{kl}) = l_{kl}(\mathbf{W})$

$$\begin{aligned}l_{kl}(\mathbf{W}) = -\tilde{\mathbf{m}}^H\tilde{\mathbf{K}}^{-1}(\mathbf{A}(\mathbf{Q},\mathbf{W}) + \tilde{\mathbf{K}}^{-1})^{-1}\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{m}} \\ + \log\det(\mathbf{A}(\mathbf{Q},\mathbf{W}) + \tilde{\mathbf{K}}^{-1}).\end{aligned} \quad (5)$$

If we compare (5) with the cost function derived in [8], we can see that it generalizes the beamforming design, studied in [8], to the case where the noise is correlated. The optimal beamformer that minimizes the RHS of (4) is found as the maximizer of $l_{kl}(\mathbf{W})$ under the constraint that $\mathbf{W}$ does not boost or weaken the signal, i.e. $\|\mathbf{W}\|_2^2 = 1$.

The analytical form of (5) is rather complicated rendering the analysis of the beamforming design a very challenging problem. An interesting case, however, arises when the covariance matrix $\tilde{\mathbf{K}}_{ee}$ can be expressed as a Kronecker product, i.e. $\tilde{\mathbf{K}}_{ee} = \tilde{\mathbf{K}}_{T,ee} \otimes \tilde{\mathbf{K}}_{R,ee}$. This assumption can be satisfied in many practical cases where our system model, described in Section II, holds. In this case we can simplify the expression of $\mathbf{A}(\mathbf{W},\mathbf{Q})$ as follows

$$\begin{aligned}\mathbf{A}(\mathbf{W},\mathbf{Q}) &= \tfrac{1}{4}(\mathbf{W}\Delta\bar{\mathbf{C}}_{kl}\tilde{\mathbf{K}}_{T,ee}^{-\star}\Delta\bar{\mathbf{C}}_{kl}^H\mathbf{W}^H)^\star \otimes \tilde{\mathbf{K}}_{R,ee}^{-1} \\ &= \tfrac{1}{4}(\mathbf{W}\mathbf{\Omega}\mathbf{W}^H)^\star \otimes \tilde{\mathbf{K}}_{R,ee}^{-1},\end{aligned} \quad (6)$$

where $\mathbf{\Omega} = \Delta\bar{\mathbf{C}}_{kl}\mathbf{K}_{T,ee}^{-\star}\Delta\bar{\mathbf{C}}_{kl}^H$. Plugging (6) into $l(\mathbf{W}, \Delta\bar{\mathbf{C}}_{kl})$ will lead to a significant simplification of the cost function (5). The resulting problem is still difficult but the most important outcome is that it admits asymptotic analysis with respect to the CSI quality of the $\mathrm{SU}_{\mathrm{Tx}} - \mathrm{SU}_{\mathrm{Rx}}$ link. Under this premise, the problem to be solved onwards is the following

$$\begin{aligned}&\text{maximize} \quad l(\mathbf{W}, \Delta\bar{\mathbf{C}}_{kl}) \\ &\text{subject to} \quad \mathrm{tr}(\mathbf{W}\mathbf{W}^H) = 1.\end{aligned} \quad \text{(P0)}$$

In the next two subsections we provide some analytical results characterizing the extreme cases of no and full CSI, assuming that $\mathbf{A}(\mathbf{W},\mathbf{Q})$ is given by (6).

B. Limiting Case-No CSI for Secondary Link

When the transmitter $\mathrm{SU}_{\mathrm{Tx}}$ is completely oblivious of the cognitive link channel, i.e. when $\delta \to 0, \tilde{\mathbf{m}} \to \mathbb{0}_{MN}$ and $\tilde{\mathbf{K}} \to \mathbf{K}$, then the initial cost function $l_{kl}(\mathbf{W})$, tends to $l_{kl,1}(\mathbf{W}) = \log\det\left[\frac{1}{4}\left(\mathbf{W}\mathbf{\Omega}\mathbf{W}^H\right)^\star \otimes \tilde{\mathbf{K}}_{R,ee}^{-1} + \mathbf{K}^{-1}\right]$ and the problem becomes

$$\begin{aligned}&\text{maximize} \quad \log\det\left[\tfrac{1}{4}\left(\mathbf{W}\mathbf{\Omega}\mathbf{W}^H\right)^\star \otimes \tilde{\mathbf{K}}_{R,ee}^{-1} + \mathbf{K}^{-1}\right] \\ &\text{subject to} \quad \mathrm{tr}(\mathbf{W}\mathbf{W}^H) = 1.\end{aligned} \quad \text{(P1)}$$

Analysis of (P1) can be carried out by decomposing $\mathbf{W}, \mathbf{\Omega}$ as $\mathbf{W} = \mathbf{U}_W\mathbf{\Sigma}_W\mathbf{V}_W^H$ and $\mathbf{\Omega} = \mathbf{U}_\Omega\mathbf{\Lambda}_\Omega\mathbf{U}_\Omega^H$. Further, recalling that $\mathbf{K} = \mathbf{K}_T \otimes \mathbf{K}_R$ we set $\mathbf{P} = \mathbf{K}_R^{1/2}\mathbf{K}_{R,ee}^{-1}\mathbf{K}_R^{1/2}$ and define $\mathbf{P} = \mathbf{U}_P\mathbf{\Lambda}_P\mathbf{U}_P^H, \mathbf{K}_T = \mathbf{U}_{K_T}\mathbf{\Lambda}_{K_T}\mathbf{U}_{K_T}^H$. Characterization of the optimal beamformer follows from the next proposition.

Proposition 1. *When no CSI is available for the direct-link, then the optimal beamformer is given as* $\mathbf{W}_1^{\mathrm{o}} = \mathbf{U}_{K_T}^\star\mathbf{\Sigma}_W\mathbf{U}_\Omega^H$, *where* $\mathbf{\Sigma}_W$ *is a properly chosen power loading matrix.*

Proof: Invoking the above definitions it is possible to manipulate the objective $l_{kl,1}(\mathbf{W})$ as follows

$$\begin{aligned}l_{kl,1}(\mathbf{W}) &= \log\det\left[\tfrac{1}{4}\left(\mathbf{W}\mathbf{\Omega}\mathbf{W}^H\right)^\star \otimes \mathbf{P} + \mathbf{K}_T^{-1} \otimes \mathbf{I}_N\right] \\ &\le \log\det\left[\tfrac{1}{4}\left(\mathbf{\Sigma}_W\mathbf{\Lambda}_\Omega\mathbf{\Sigma}_W^H\right) \otimes \mathbf{\Lambda}_P + \mathbf{\Lambda}_{K_T}^{-1} \otimes \mathbf{I}_N\right],\end{aligned}$$

where the inequality is basically the Hadamard inequality, satisfied with equality when $\mathbf{U}_W, \mathbf{V}_W$ are chosen as in Proposition 1. The new objective is separable in the diagonal elements of the matrix $\mathbf{\Sigma}_W$ and thus it can be maximized very efficiently, using Newton-based or gradient methods, with complexity $\mathbb{O}(M)$ operations per iteration. ∎

Since our design should address the worst case scenario, problem (P1) has to be solved over all possible codeword pairs $(\mathbf{C}_k, \mathbf{C}_l)$, i.e. a total of $\binom{|\mathcal{U}|^M}{2}$ combinations, where $|\cdot|$ denotes the cardinality of the constellation set $\mathcal{U}$, a task which is rather demanding. However this complexity can be affordable in practice, if the constellation cardinality $|\mathcal{U}|$ and the antenna array size M are small and we exploit potential symmetries that might arise for different pairs (k,l) of indices.

C. Limiting Case-Full CSI for Secondary Link

When the cognitive transmitter $\mathrm{SU}_{\mathrm{Tx}}$ has perfect knowledge of the cognitive link channel, i.e. $\delta \to 1, \tilde{\mathbf{m}} \to \mathbf{h}, \tilde{\mathbf{K}} \to \mathbf{0}_{MN}$, then extrapolating the relevant analysis, performed in [8], for the same asymptotic scenario, it turns out that $l_{kl}(\mathbf{W})$ tends to the following limiting function

$$\begin{aligned} l_{kl,2}(\mathbf{W}) &= \mathbf{h}^H\big[\big(\mathbf{W}\boldsymbol{\Omega}\mathbf{W}^H\big)^\star \otimes \tilde{\mathbf{K}}_{R,ee}^{-1}\big]\mathbf{h} \\ &= \operatorname{tr}\Big(\mathbf{W}\boldsymbol{\Omega}\mathbf{W}^H\mathbf{H}\tilde{\mathbf{K}}_{R,ee}^{-1}\mathbf{H}^H\Big) \\ &= \operatorname{tr}(\mathbf{W}\boldsymbol{\Omega}\mathbf{W}^H\boldsymbol{\Psi}), \end{aligned} \tag{7}$$

where we have defined the matrix $\boldsymbol{\Psi} = \mathbf{H}\tilde{\mathbf{K}}_{R,ee}^{-1}\mathbf{H}^H$ with decomposition $\boldsymbol{\Psi} = \mathbf{U}_\Psi\boldsymbol{\Lambda}_\Psi\mathbf{U}_\Psi^H$ and $\boldsymbol{\Omega}$ is previously defined. The problem that has to be solved onwards is

$$\begin{aligned} &\text{maximize} \quad \operatorname{tr}(\mathbf{W}\boldsymbol{\Omega}\mathbf{W}^H\boldsymbol{\Psi}) \\ &\text{subject to} \quad \operatorname{tr}(\mathbf{W}\mathbf{W}^H) = 1. \end{aligned} \tag{P2}$$

The optimal beamformer for problem (P2) has a very simple form, described in the following Proposition.

Proposition 2. *When full CSI is available for the direct-link, then the optimal beamformer is given as* $\mathbf{W}_2^\circ = \mathbf{u}_{\Psi,1}\mathbf{u}_{\Omega,1}^H$, *where* $\mathbf{u}_{\Psi,1}$ *and* $\mathbf{u}_{\Omega,1}$ *denote the principal eigenvectors of matrices* $\mathbf{U}_\Psi$ *and* $\mathbf{U}_\Omega$, *respectively.*

Proof: From Kristof's theorem [16], we have $l_{kl,2}(\mathbf{W}) \leq \operatorname{tr}(\boldsymbol{\Sigma}_W\boldsymbol{\Lambda}_\Omega\boldsymbol{\Sigma}_W\boldsymbol{\Lambda}_\Psi)$, with equality achieved for $\mathbf{V}_W = \mathbf{U}_\Omega$ and $\mathbf{U}_W = \mathbf{U}_\Psi$. If we define $\boldsymbol{\mu} = [\sigma_{W,1}^2 \ \dots \ \sigma_{W,M}^2]^T$ and $\boldsymbol{\nu} = [\lambda_{\Omega,1}\lambda_{\Psi,1} \ \dots \ \lambda_{\Omega,M}\lambda_{\Psi,M}]^T$, where $(\sigma_{W,k}, \lambda_{\Omega,k}, \lambda_{\Psi,k})$ denote the k-th diagonal elements of $(\boldsymbol{\Sigma}_W, \boldsymbol{\Lambda}_\Omega, \boldsymbol{\Lambda}_\Psi)$, then we can solve the following equivalent vector problem

$$\begin{aligned} &\text{maximize} \quad \boldsymbol{\mu}^T\boldsymbol{\nu} \\ &\text{subject to} \quad \boldsymbol{\mu}^T\mathbb{1}_M = 1 \end{aligned} \tag{P2e}$$

Since $\boldsymbol{\mu}, \boldsymbol{\nu} \succeq \mathbf{0}$, with their elements in descending order of magnitude, it is easy to see that the optimal value $p^\star$ of (P2e) is $p^\star = \nu_1$ attained by $\boldsymbol{\mu}^\circ = [1\ 0 \dots\ 0]$. Thus, for the original problem (P2) we get $\sigma_{W,1} = 1$ and $\sigma_{W,k} = 0, k \neq 1$ implying that the optimal beamforming vector would be $\mathbf{W}_2^\circ = \mathbf{u}_{\Psi,1}\mathbf{u}_{\Omega,1}^H$. Note that, the optimization problem (P2) should also be solved for all possible combinations of codeword pairs $(\mathbf{C}_k, \mathbf{C}_l)$. ■

D. Comments on System Design

Even thought the cases we have treated this far do not capture the most general model, they are still applicable to many scenarios and provide us with useful insights on the beamforming designs.

Both results of Propositions 1 and 2 are in line with intuition, regarding the beamforming behavior. It can be inferred that the optimal design serves a dual role; on one hand it compensates for the orthogonality loss of the OSTBC, due to interference, and on the other hand it properly steers transmission over the channel in order to minimize the upper bound on the error probability.

Prior to the presentation of the numerical results, let us have a look into a simple scenario, where we assume uncorrelated primary transmit antennas and $M_j \geq L\,\forall\, j$, with L defined in Section II-B as the code time span. Under these premises, it is possible to write $\tilde{\mathbf{K}}_{ee}$ as $\tilde{\mathbf{K}}_{ee} = \mathbf{I}_M \otimes \tilde{\mathbf{K}}_{R,ee}$ for some matrix $\tilde{\mathbf{K}}_{R,ee}$ and express $\mathbf{A}(\mathbf{W}, \Delta\bar{\mathbf{C}}_{k,l})$ in terms of $\mathbf{Z} = \mathbf{W}\mathbf{W}^H$ as $\mathbf{A}(\mathbf{Z}) = (\frac{\mu_{kl}^2}{4})\mathbf{Z} \otimes \tilde{\mathbf{K}}_{R,ee}^{-1}$, where $\mu_{kl} = \|\Delta\bar{\mathbf{C}}_{kl}\|_2$. Using the high SNR approximation we can argue that the worst case scenario, in terms of probability of error, is sufficiently captured by solving (P0) for $\mu_{\min} = \min\{\mu_{kl}\ \forall\ k \neq l\}$. Moreover the non-causal knowledge of $\widetilde{\mathbf{C}}_j\,\forall j$ can be removed, from the secondary transmitter $\mathrm{SU_{Tx}}$. Therefore under these moderate assumptions the problem simplifies greatly both in overhead and complexity requirements.

IV. Numerical Results

In this section we investigate the performance of our system, through numerical simulations for various system parameter setups. Since our main interest is to investigate the impact of cross-link CSI on system performance, we confine our study to the cases where the $\mathrm{SU_{Tx}} - \mathrm{SU_{Rx}}$ link CSI quality is better than that of the cross-links, and consider various combinations of the cross-link CSI quality and the antenna correlation.

Our evaluation is carried out for the simple scenario, described in Section III-D, for one primary link. We assume rate-one ($L = M$) OSTBCs and uncorrelated transmit antennas everywhere. The receiver correlation for the $\mathrm{SU_{Tx}} - \mathrm{SU_{Rx}}$ channel is $\mathbf{K}_R = (1-c)\sigma_h^2\mathbf{I}_N + c\sigma_h^2\mathbb{1}_N\mathbb{1}_N^T$, i.e. antenna correlation (normalized with channel variance σ_h^2) equals $c \in [0, 1)$. Similarly, we define the cross-link channel using $\mathbf{K}_{R,11}, \sigma_{f_1}^2, c$. Assuming Rayleigh fading with $\sigma_h^2 = 1, \sigma_{f_1}^2 = {}^1\!/_{M_1}$ we further define the received power, the interference power and the noise power as $P = \mathbf{E}\|\mathbf{H}^H\mathbf{W}\bar{\mathbf{C}}\|_2^2, \sigma_I^2 = \mathbf{E}\|\mathbf{F}_1^H\tilde{\mathbf{C}}_1\|_2^2$ and $\sigma_n^2 = \mathbf{E}\|\mathbf{N}\|_2^2$, respectively. Based on these, we obtain the ratios $\mathrm{INR} = \frac{\sigma_I^2}{\sigma_n^2}$ and $\mathrm{SINR} = \frac{P}{\sigma_I^2+\sigma_n^2}$. Moreover, the CSI qualities of the cognitive and the $\mathrm{PU_{Tx,1}} - \mathrm{SU_{Rx}}$ link are indicated by δ, δ_1, which characterize the cross-covariance matrices as $\mathbf{K}_{\mathbf{h}\hat{\mathbf{h}}} = \delta\mathbf{I}_M \otimes \mathbf{K}_R$ and $\mathbf{K}_{\mathbf{f}_1\hat{\mathbf{f}}_1} = \delta_1\mathbf{I}_{M_1} \otimes \mathbf{K}_{R,11}$.

The numerical results for various choices of parameters are depicted in Figures 2-4. In Figure 2 we see the bit-error-rate (BER) performance of the system, when the interference is of the same magnitude as the noise. We can see that the 2×4 secondary system performs quite well and even the 2×2 system has acceptable BER results. We note that performance loss, with diminishing CSI quality which is reflected in δ_1, for both setups is insensitive to the number of receive antennas. For both receive array sizes the induced degradation, when δ_1 decreases, is of the same magnitude.

The impact of correlation factor c is seen in Figure 3. As expected, our system experiences some loss, w.r.t the BER, when the correlation factor c increases because the MIMO direct-link channel matrix condition number is increasing. However, an interesting observation is that decreasing CSI quality incurs a smaller BER penalty for higher c. Finally, in Figure 4 we can see the connection between INR and δ_1. In line with intuition, as the INR increases, primary interference becomes more dominant, thus partial knowledge can significantly help the system. Conversely, having less side information, i.e. smaller δ_1, incurs higher loss for $\mathrm{INR} = 20\mathrm{dB}$ than $\mathrm{INR} = 0\mathrm{dB}$ due to the higher contribution, in the case of 20dB, of interference in the system disturbance.

It is good to keep in mind that these results serve the qualitative assessment of the impact of the cross-link CSI quality. In practical systems, where the direct-link CSI will likely be imperfect, we expect further degradation. However,

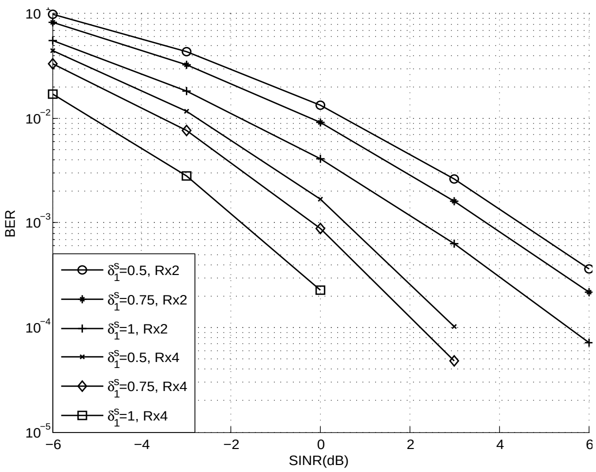

Fig. 2. BER performance for a system with $(M_1, M) = (2, 2)$ antennas, $c = 0.5$, INR $= 0$dB and BPSK input modulation.

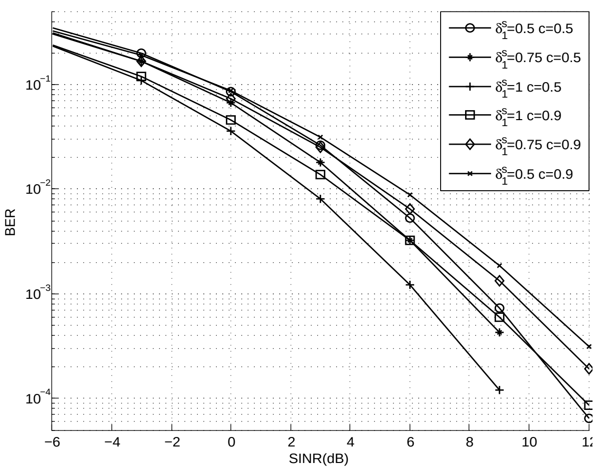

Fig. 3. BER performance for a system with $(M_1, M, N) = (2, 2, 2)$ antennas, INR $= 0$dB and QPSK input modulation.

these results are an indication that cognitive communication is feasible under moderate assumptions for the CRN components.

V. Conclusions and Future Work

In this work we investigated the design of a BOSTBC strategy in the context of overlay CNRs. In this framework, the system design has to take into consideration and sufficiently tackle the incoming interference, which is injected from the primary system. Assuming a statistical model for the description of the CSI quality, we incorporated the primary interference in our design optimization problem. Moreover we characterized, theoretically and numerically, the behaviour of the resulting beamforming solution in terms of the interference and for different CSI levels.

Numerical evaluation, performed for our simple scenarios, has shown that, albeit the incurred performance loss, the secondary system can still sustain reliable communication. The addition of extra primary links or the relaxation of the perfect CSI assumption for the direct-link will naturally lead to a larger system BER degradation. However, based on our

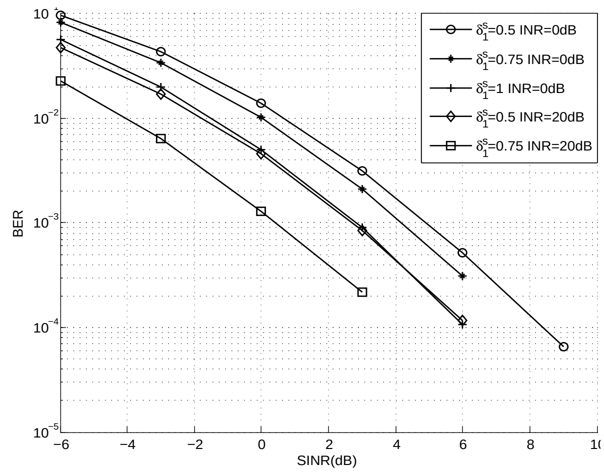

Fig. 4. BER performance for a system with $(M_1, M, N) = (2, 2, 2)$ antennas, $c = 0.7$ and BPSK input modulation.

numerical evalation, we can reason that the expected performance will still lie within the limits of acceptable operational standards for reliable communication. This is a positive indication, regarding the future realization of unlicensed systems.

Appendix

A. Statement of Lemma 1

Lemma 1. *Assume matrices* $\mathbf{A}_k \in \mathbb{C}^{N\times N}$*, Gaussian N-length vectors* $\mathbf{x}_k \sim \mathbb{C}\mathcal{N}(\mathbf{m}_{\mathbf{x}_k}, \mathbf{K}_{\mathbf{x}_k\mathbf{x}_k}), \mathbf{y}_k \sim \mathbb{C}\mathcal{N}(\mathbf{m}_{\mathbf{y}_k}, \mathbf{K}_{\mathbf{y}_k\mathbf{y}_k})$*, with* $k \in \{1\ldots, K\}$*, and* $\mathbf{w} \sim \mathbb{C}\mathcal{N}(\mathbf{m}_\mathbf{w}, \mathbf{K}_{\mathbf{ww}})$*. All vectors, but the pair* $(\mathbf{x}_k, \mathbf{y}_k)$ *with cross-covariance* $\mathbf{K}_{\mathbf{x}_k\mathbf{y}_k}$*, are pairwise independent. If we define* $\mathbf{z} = \sum_{k=1}^K \mathbf{A}_k\mathbf{x}_k + \mathbf{w}$ *then we can write the p.d.f* $p_{\mathbf{z}|\{\mathbf{y}_k\}_{k=1}^K}$ *as*

$$p_{\mathbf{z}|\{(\mathbf{A}_k,\mathbf{y}_k)\}_{k=1}^K} = p_\mathbf{w} \prod_{k=1}^K p_{\mathbf{A}_k\mathbf{x}_k|\mathbf{y}_k} \tag{A1}$$

Proof: For simplicity we define $\tilde{\mathbf{z}} = \mathbf{z}|\{\mathbf{y}_k\}_{k=1}^K$, such that $\tilde{\mathbf{z}} \sim \mathbb{C}\mathcal{N}(\mathbf{m}_{\tilde{\mathbf{z}}}, \mathbf{K}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}})$. In order to derive the mean and the covariance of $\tilde{\mathbf{z}}$ we need to calculate the quantities $\mathbf{K}_{\mathbf{zy}}, \mathbf{K}_{\mathbf{yy}}$, where $\mathbf{y} = [\mathbf{y}_1^T \ \ldots \ \mathbf{y}_K^T]^T$. It obviously holds that $\mathbf{K}_{\mathbf{yy}} = \mathrm{diag}(\mathbf{K}_{\mathbf{y}_1\mathbf{y}_1}, \ldots, \mathbf{K}_{\mathbf{y}_K\mathbf{y}_K})$ and invoking, further, the independence of the each summation term $\mathbf{A}_k\mathbf{x}_k$ with $\{\mathbf{y}_i\}_{i\neq k}$ it follows that $\mathbf{K}_{\mathbf{zy}} = [\mathbf{A}_1\mathbf{K}_{\mathbf{x}_1\mathbf{y}_1} \ldots \mathbf{A}_K\mathbf{K}_{\mathbf{x}_K\mathbf{y}_K}]$. We can now calculate $\mathbf{m}_{\tilde{\mathbf{z}}}$ and $\mathbf{K}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$, using known formulas [14], as

$$\begin{aligned} \mathbf{m}_{\tilde{\mathbf{z}}} &= \sum_{k=1}^K [\mathbf{m}_{\mathbf{x}_k} + \mathbf{A}_k(\mathbf{y}_k - \mathbf{m}_{y_k})] + \mathbf{m}_\mathbf{w} \\ \mathbf{K}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} &= \sum_{k=1}^K [\mathbf{A}_k(\mathbf{K}_{\mathbf{x}_k\mathbf{x}_k} - \mathbf{K}_{\mathbf{x}_k\mathbf{y}_k}\mathbf{K}_{\mathbf{y}_k\mathbf{y}_k}^{-1}\mathbf{K}_{\mathbf{y}_k\mathbf{x}_k})\mathbf{A}_k^H] \\ &\quad + \mathbf{K}_{\mathbf{ww}} \end{aligned} \tag{A2}$$

It is obvious, from the above equations, that $\mathbf{m}_{\tilde{\mathbf{z}}}, \mathbf{K}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$ are separable with respect to the statistics of each pair $(\mathbf{x}_k, \mathbf{y}_k)$. Since the mean and the covariance fully characterize a Gaussian distribution, this separability indicates that the p.d.f of $\tilde{\mathbf{z}}$ is given as in (A1). This characterization, not surprising in hindsight, facilitates easier calculations and allows for insightful interpretations. ∎

References

[1] J. Mitola, Cognitive radio: An Integrated Agent Architecture for Software. Ph.D Dissertation, KTH, Stockholm, Sweden, 2000.

[2] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[3] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.

[4] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894-914, May 2009.

[5] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585-596, Nov. 1999.

[6] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437-1450, Oct. 1998.

[7] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.

[8] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 611-627, Mar. 2002.

[9] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Proc.*, vol. 50, no. 10, pp. 2599-2613, Oct. 2002.

[10] M. Vu and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with non-zero mean in space-time coded systems," *IEEE Trans. on Signal Proc.*, vol. 54, no. 6, pp. 2318-2332, June 2006.

[11] A. Hjorungnes and D. Gesbert, "Precoding of orthogonal space-time block codes in arbitrarily correlated MIMO channels: Iterative and closed-form solutions," *IEEE Trans. on Wireless Commun.*, vol. 6, no. 3, pp. 1072-1082, Mar. 2007.

[12] A. Dogandzic, "Chernoff bounds on pairwise error probabilities of space-time codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1327-1336, May 2003.

[13] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wideband MIMO radio channels based on NLOS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 655-665, May 2004.

[14] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, Upper Sadle River, NJ, 1993.

[15] Van Trees, H. Detection, Estimation and Modulation Theory. John Wiley & Sons, Inc., New York, 1968.

[16] W. Kristof, "A theorem on the trace of certain matrix products and some applications." *Journal of Mathematical Psychology*, vol. 7, no. 3, pp. 515-530, 1970.