

Cognitive Transmit Beamforming From Binary CSIT

Balasubramanian Gopalakrishnan and Nicholas D. Sidiropoulos, *Fellow, IEEE*

Abstract—Transmit beamforming is used to steer radiated power towards a receiver of interest and to limit interference to unintended receivers, thereby facilitating coexistence. Transmit beamforming requires accurate channel state information (CSI) at the transmitter, which is often difficult to acquire, particularly in cognitive underlay settings, where the primary receiver cannot be expected to cooperate with the secondary system to enable it to learn the secondary to primary crosstalk channel. This paper considers cases where it is not realistic to assume channel reciprocity, or that the receivers are capable of accurate CSI estimation and feedback—because they are legacy systems, or have limited computation/energy resources. Transmit beamforming from binary and infrequent CSI is first considered for an isolated link. An online beamforming and learning algorithm is developed using the analytic center cutting plane method and is shown to asymptotically attain optimal performance. A robust maximum-likelihood formulation is next developed to handle feedback errors and correlation drift. The setup is then generalized to a cognitive underlay setting, also exploiting the standard acknowledgement/negative-acknowledgement feedback on the reverse primary link. This is the first solution to jointly tackle secondary signal-to-noise ratio maximization and primary interference mitigation from only rudimentary CSI, without assuming channel reciprocity.

Index Terms—Transmit beamforming, spatial channel correlation, online learning, cutting plane method, maximum likelihood, cognitive radio network underlay.

I. INTRODUCTION

TRANSMIT beamforming uses multiple antennas and channel state information at the transmitter (CSIT) to steer radiated power towards directions of interest, limiting leakage in other directions [1]. The direction(s) of interest correspond to line-of-sight and specular multipath components of the propagation channel to the desired receiver(s), while limiting leakage controls interference to nearby co-channel systems. Transmit beamforming therefore facilitates coexistence, which is crucial for dynamic spectrum access in cognitive radio networks. The price paid is the need for accurate channel estimation at the receiver (Rx), and channel state feedback to the transmitter (Tx). In order to mitigate the communication overhead involved in feeding back instantaneous channels, an alternative is to work

Manuscript received February 27, 2014; revised July 2, 2014 and September 24, 2014; accepted September 30, 2014. Date of publication October 9, 2014; date of current version February 6, 2015. This work was supported by the National Science Foundation under Grant EARS/AST-1247885. This paper was presented in part at the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing Florence, Italy May 2014. The associate editor coordinating the review of this paper and approving it for publication was Y. Jing.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: gopa0022@umn.edu; nikos@umn.edu).

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

Digital Object Identifier 10.1109/TWC.2014.2362119

with the channel correlation matrix, which enables the Tx to optimize the *average* received signal to noise ratio (SNR). This is known as *long-term* transmit beamforming.

Designing long-term transmit beamforming vectors for the secondary system in cognitive radio networks presents the secondary Tx with the additional challenge of learning the correlation matrix of the channel to the non-cooperative primary Rx, in addition to the correlation matrix of the channel to the intended secondary Rx. For a cognitive radio underlay network, where the secondary and the primary system share the same frequency band for communication, this knowledge is essential for optimizing the average received SNR at the secondary Rx, while limiting interference to the primary network, thereby maintaining the Quality of Service (QoS) at the primary Rx. It is much more difficult for the secondary Tx to learn the correlation matrix of the channel to the primary Rx versus the secondary Rx, due to lack of cooperation from the primary Rx [2], [3]. This is especially true when the primary is a legacy system, that is not only unwilling, but also unable to cooperate.

A. Prior Work

The transmit beamforming vector that maximizes the average received SNR for a multiple-input single-output (MISO) link is the principal eigenvector of the channel correlation matrix [4]. In a cognitive underlay scenario [5] comprising a primary and a secondary link, the optimal transmit beamforming vector that maximizes the average SNR at the secondary Rx while limiting the interference caused to the primary Rx can be obtained by solving a convex optimization problem [6], [7], provided the secondary Tx knows the correlation matrices of the channel to the secondary Rx and the crosstalk channel to the primary Rx. Information about these correlation matrices can only be obtained through the respective receivers, except in cases where channel reciprocity can be assumed. Channel reciprocity can only be assumed in time-division duplex systems, but even there reciprocity can be a very coarse approximation, e.g., due to differences in local scattering, or when nodes use different transmit and receive beam patterns.

For most types of systems, CSI is acquired at the Rx-side and fed back to the corresponding Tx. In our context, this means that each Rx is responsible for estimating the channel correlation matrix for its own link. Instead of feeding back the correlation matrix, each Rx may compute the optimal beamforming vector (i.e., its principal eigenvector, e.g., via the power method) and send this back to the Tx. Either way, scalar or vector quantization is needed to limit the feedback rate. The beamforming vector can be quantized using a custom codebook, see [8] where bounds on the codebook size needed for a given SNR loss are given. The approach in [8] was developed for instantaneous

feedback, but it can be extended [8] to long-term feedback. A rate-distortion analysis of vector quantization performance in this context can be found in [9]. In [8] and [9], the codebook is assumed to be designed off-line and shared between the Tx and Rx beforehand; this is not suitable when the Tx and Rx are opportunistically paired.

For designing secondary transmit beamforming vectors, almost all methods in the literature assume crosstalk channel reciprocity and suggest that the secondary Tx can learn the channel to the primary Rx by overhearing the transmissions of the primary Rx to the primary Tx [2], [3], [10], [11]; and either secondary channel reciprocity or that the secondary Rx estimates and feeds back accurate link CSI to the secondary Tx.

B. Motivation

What if the receivers cannot perform correlation matrix estimation, summarization, and feedback? This can happen if they have limited computation and communication capabilities (low cost, small size and battery), when they are hard-wired for legacy protocols, or when they are opportunistically paired with a transmitter, with limited negotiation before payload transmission. Is it possible to acquire accurate CSI at the transmitter with only rudimentary feedback from the receivers, e.g., acknowledgement/negative-acknowledgement (ACK/NACK)-type? Is it possible to design effective transmit beamforming solutions this way for cognitive underlay and more sophisticated (e.g., broadcast/multicast) beamforming scenarios?

For a point-to-point MISO channel, Mudumbai *et al.* [12] and Banister *et al.* [13] have proposed transmit weight vector adaptation algorithms using 1-bit feedback from the Rx. In both the cases, the Tx periodically updates and stores the beamforming vector $\mathbf{w}_{\text{best}}(t)$, that resulted in the maximum SNR at the Rx until the current time slot t and the Rx updates and stores the maximum received SNR till t denoted by $SNR_{\text{best}}(t)$. At the start of time slot $(t + 1)$, the Tx perturbs $\mathbf{w}_{\text{best}}(t)$ to obtain the beamforming vector $\mathbf{w}(t + 1)$ and uses it to send data to the Rx. The Rx measures the corresponding received SNR, compares it with $SNR_{\text{best}}(t)$ and reports a “1” or a “0” to the Tx depending on whether the SNR at current time slot is \geq or $<$ SNR_{best} . Upon receiving a “1,” the Tx updates $\mathbf{w}_{\text{best}}(t + 1)$ with $\mathbf{w}(t + 1)$ or with $\mathbf{w}_{\text{best}}(t)$ otherwise. The difference in [12] and [13] is that in the former all entries of the beamforming vector have fixed magnitude equal to $\sqrt{\frac{1}{N}}$, (N = number of Tx antennas) and the phase of every entry is randomly perturbed each time; while in the latter the beamforming vector is perturbed with a coarse estimate of the gradient of the received signal power, and the resulting vector is normalized to unit norm. It is assumed that the channel remains fixed or changes very slowly during this adaptation process. Asymptotic convergence to the optimal beamforming vector has been shown for both algorithms when the channel remains fixed. Simulations show that these random exploration algorithms result in a slow convergence rate.

For the cognitive radio underlay scenario, notable exceptions to the pervasive use of the crosstalk channel reciprocity assumption are recent works by Zhang [14] and Noam and

Goldsmith [15], [16]. A primary and a secondary single antenna Tx–Rx link were considered in [14], where it was assumed that the communication protocol in the primary network and the transmission rate and power adaptations by the primary Tx are known to the secondary Tx. The communication is split into two phases: a) active learning, and b) supervised transmission. During the active learning phase, the secondary Tx probes the primary Rx with interfering signals, observes the corresponding transmit rate/power adaptations in the primary network and uses this as indirect feedback from the primary network to estimate the interference channel gain from the secondary Tx to the primary Rx. During the supervised transmission phase, the secondary Tx uses the interference channel estimate obtained by active learning and transmits data in such a way that the secondary Rx receives it with high SNR and the interference to the primary Rx is below its interference threshold, which is also assumed known at the secondary Tx. Overall, [14] requires inside knowledge and tight monitoring of the primary system, which may not be possible in ad-hoc deployments.

Noam and Goldsmith [15], [16] proposed an algorithm that enables the secondary Tx to learn the *fixed* interference channel to the primary Rx, by measuring a monotonic function of the interference to the primary Rx [15] or overhearing the ACK/NACK feedback [16] in the reverse primary link. They proposed to vary the secondary transmit precoding matrix to probe the primary Rx, gradually collecting information on what it can tolerate. They proposed using a cyclic Jacobi subspace estimation algorithm, and proved that it converges to the interference channel between the secondary Tx and primary Rx. The secondary Tx can indirectly learn the best signaling subspace this way, without assuming channel reciprocity, and without altering the primary communication protocol—an exciting development. Ideally though, such primary Rx “probing” should be done in parallel with secondary Rx channel exploration (and possibly also payload transmission); and acquisition speed is of essence. We explore these and other issues next.

C. Contributions

We begin by considering the long-term transmit beamforming problem for a MISO link in the case when the Rx has limited computational capabilities, and/or is paired up opportunistically with the Tx. We explore how the Tx can learn to beamform on-the-fly from very low-rate channel quality indicator bits fed back from the Rx (average received SNR \gtrsim pre-determined threshold), while transmitting payload simultaneously. The beamforming vectors are designed such that they not only exploit the acquired information gathered in the past to maximize a Tx-side estimate of the average received SNR, but are also diverse enough to explore the channel correlation space efficiently and learn the channel correlation matrix accurately over time. Towards this end, the analytic center cutting plane method (ACCPM) from optimization is leveraged to develop an online channel correlation matrix learning algorithm based on one-bit SNR feedback. In the absence of binary measurement or feedback communication errors, the proposed algorithm restricts the channel correlation estimate to a ball of radius r centered around the true value, within $\mathcal{O}\left(\frac{N^2}{r^2}\right)$ iterations

(where N is the number of transmit antennas) and the average received SNR converges asymptotically to the maximum achievable SNR value (obtained with perfect knowledge of the channel correlation matrix at the Tx) [17].

For the practically important case where there are occasional SNR measurement errors at the receiver, or binary feedback errors in the reverse link from the receiver to the transmitter, a robust maximum likelihood formulation is proposed and shown to be effective in dealing with such errors. A discounted maximum likelihood formulation is also proposed to enable CSI tracking and adaptation of the transmit beamforming vector in cases where the channel correlation matrix itself changes slowly, as time goes by.

The techniques proposed for an isolated MISO link are subsequently used as a foundation for designing secondary transmit beamforming vectors for cognitive radio networks (CRNs). Learning to beamform at the secondary transmitter in a cognitive underlay setting is far more challenging, as the primary receiver cannot be assumed to cooperate in “teaching” the secondary transmitter how to avoid causing interference. Based on novel formulations that exploit the binary feedback in the secondary system *and* the possibility of “overhearing” the usual ACK/NACK feedback on the reverse primary link, joint cognitive beamforming and primary interference avoidance algorithms are developed. Two distinct scenarios are considered, depending on whether or not the secondary transmitter knows the primary interference threshold. When it does, convergence of the beamforming vector to the optimal one (obtained with perfect CSIT) is proven; otherwise a power back-off mechanism is proposed to enable the secondary transmitter to learn the unknown primary interference threshold. Interestingly, simulations show that it is possible to learn the primary interference threshold and approach optimal secondary link performance this way, as if perfect knowledge of the interference threshold and CSIT were available—albeit we do not have proof of convergence in this case.

The main novelty is the ability to gradually acquire CSI and design optimal transmit beampatterns from rudimentary CSI feedback. This is the first solution to jointly tackle secondary SNR maximization *and* primary interference avoidance, without assuming reciprocity or altering the primary’s signaling protocol, while enabling asymptotically optimal performance from only binary CSI. Some of the proposed approaches are very simple, making them ideal for practical implementation. As cognitive radio research and development inches closer to deployment, algorithms that can work under realistic channel and feedback conditions are likely to have a big impact in terms of practical transceiver and network engineering.

Relative to its conference precursor [17], this journal version brings in the robust discounted maximum likelihood formulation, generalization to the cognitive radio underlay setting, including proof of convergence of a pair of jointly driven cutting plane iterations for the case when the primary interference threshold is known to the secondary transmitter, and extensive simulation results.

As we were putting the finishing touches on this manuscript, we became aware, through the reviews of our conference paper [17], of a parallel submission to the same conference by Xu

and Zhang [18], containing an idea that is very similar to [17]. Namely, [18] considers transmit beamforming for wireless energy transfer (vs. communication), and a cutting plane method to learn the channel correlation matrix from one-bit feedback. While several design choices are naturally different, these being independent pieces of work, the core idea is the same in both papers. Summarizing the differences, [18] assumes separate learning and “bulk transfer” phases, uses higher-rank precoding instead of beamforming during the learning stage, and does not communicate thresholds to the receiver—at the cost of not controlling the transmission power during the learning phase. Using a higher-rank precoder may enable faster exploration and learning, but requires separate up-conversion chains that are not needed during show time (the payload phase) where beamforming is used. These different design choices are complementary in many ways, and it is interesting to see that the same basic idea was independently discovered but fleshed out in different ways by the two groups.

II. SYSTEM MODEL AND BASIC PROBLEM FORMULATION

Consider a point-to-point MISO link comprising a Tx with N antennas and a Rx with a single antenna. Time is divided into transmission rounds or *slots* of length T seconds, with each slot comprising enough symbols for the Rx to perform relatively accurate power estimation. Initially, the Tx starts transmitting data using an arbitrary beamforming vector \mathbf{w}_0 . At time $tT + \tau$, where t is a “slow time” (slot index) and τ is “fast time,” the channel from the Tx to the Rx is modeled as a complex random $N \times 1$ vector $\mathbf{h}(tT + \tau)$, with $E[\mathbf{h}(tT + \tau)\mathbf{h}^H(tT + \tau)] = \mathbf{R}_h, \forall t, \forall \tau$. At the same time, the Tx sends the complex zero-mean unit-variance symbol $x(tT + \tau)$ times a complex beamforming vector \mathbf{w}_t , and the Rx measures

$$y(tT + \tau) = \mathbf{w}_t^H \mathbf{h}(tT + \tau)x(tT + \tau) + z(tT + \tau), \quad (1)$$

where the additive noise $z(\cdot)$ has zero mean, variance σ^2 , and is independent of $x(\cdot)$ and $\mathbf{h}(\cdot)$. In order to decode the data, the Rx should at least estimate $\mathbf{w}_t^H \mathbf{h}(tT + \tau)$. This can be accomplished using a few pilot symbols per slot (or differential modulation/demodulation), and it is far simpler than estimating the vector $\mathbf{h}(tT + \tau)$. The average received SNR for slot t is given by $E\left(\frac{|\mathbf{w}_t^H \mathbf{h}|^2}{\sigma^2}\right) = \frac{\mathbf{w}_t^H \mathbf{R}_h \mathbf{w}_t}{\sigma^2}$. The beamforming vector that maximizes the average received SNR is the principal eigenvector of \mathbf{R}_h scaled according to the available transmit power. The Tx does not have any initial CSI and its objective is to learn \mathbf{R}_h and maximize the average received SNR based on binary CSIT—that is, binary slot-average SNR feedback. More specifically, in each time slot t , the Rx estimates the average SNR and compares it with a threshold γ_t . A “1” is fed back to the Tx if the average SNR is $\geq \gamma_t$ and a “0” is fed back otherwise. It is initially assumed that there are no measurement or feedback communication errors. Based on the single-bit feedback at time t , the Tx learns that

$$\begin{cases} \mathbf{w}_t^H \mathbf{R}_h \mathbf{w}_t \geq \gamma_t, & \text{when } s_t = 1; \text{ or} \\ \mathbf{w}_t^H \mathbf{R}_h \mathbf{w}_t < \gamma_t, & \text{when } s_t = 0, \end{cases} \quad (2)$$

where s_t is the 1-bit feedback at time t . For every feedback bit, the Tx learns an additional inequality, which can reduce the uncertainty about \mathbf{R}_h . This naturally raises the question whether appropriate choice of $\{\mathbf{w}_t, \gamma_t\}$ can quickly shrink the feasible region for \mathbf{R}_h , and even yield a sequence of estimates $\hat{\mathbf{R}}_h(t) \rightarrow \mathbf{R}_h$, as $t \rightarrow \infty$. More importantly, is it possible to approach the SNR attained with full CSIT, i.e., using the principal eigenvector of \mathbf{R}_h ? While this may seem an ambitious goal with only rudimentary CSIT, we will see that the answer is on the affirmative, and in fact relatively few feedback bits suffice to approach optimum performance, for practical purposes.

Simultaneous Exploration–Exploitation: For every slot t , the Tx can choose \mathbf{w}_t in such a way that it not only gathers a significant amount of information about \mathbf{R}_h (from the 1-bit feedback), but also tries to deliver a high average received SNR, thus enabling channel learning in parallel with payload transmission. To accomplish the former objective, the beamforming vectors chosen at each instant should be as diverse as possible relative to the previously chosen weight vectors, so that, over time, the Tx will learn about \mathbf{R}_h from as many different directions as possible. For the latter, the best that the Tx can do to deliver a high average received SNR is to assume that $\hat{\mathbf{R}}_h$ is close to \mathbf{R}_h and choose the beamforming weight vector along the direction of the principal eigenvector of $\hat{\mathbf{R}}_h$. Since the Tx does not have any CSI to start with, initially it has to give preference towards choosing weight vectors that aggressively explore the channel correlation space to improve the accuracy of $\hat{\mathbf{R}}_h$; and as time progresses, slowly shift emphasis towards beamforming vectors in the direction of the principal eigenvector of $\hat{\mathbf{R}}_h$. This ensures that as $\hat{\mathbf{R}}_h$ approaches \mathbf{R}_h (as will be shown later), \mathbf{w}_t will approach the direction of the principal eigenvector of \mathbf{R}_h , thus attaining the maximum average received SNR that is achievable with perfect knowledge of \mathbf{R}_h at the Tx.

At the end of slot t , the Tx has learned the following inequalities about \mathbf{R}_h from the t received feedback bits from the Rx.

$$\mathbf{w}_i^H \mathbf{R}_h \mathbf{w}_i \geq \gamma_i, \quad \forall i \in \mathcal{G}_1 \quad (3)$$

$$\mathbf{w}_i^H \mathbf{R}_h \mathbf{w}_i < \gamma_i, \quad \forall i \in \mathcal{G}_2 \quad (4)$$

where $\mathcal{G}_1 = \{i : 1 \leq i \leq t, s_i = 1\}$, $\mathcal{G}_2 = \{i : 1 \leq i \leq t, s_i = 0\}$, $\mathcal{G}_1 \cup \mathcal{G}_2 = \{1, 2, \dots, t\}$ and t is the number of elapsed time slots.

We propose to update $\hat{\mathbf{R}}_h(t)$ (the Tx-side estimate of \mathbf{R}_h at time t) as follows.

$$\begin{aligned} \Pi_1 \hat{\mathbf{R}}_h(t) = \arg \max_{\mathbf{R}_h} & \sum_{i \in \mathcal{G}_1} \log(\text{Tr}(\mathbf{W}_i \mathbf{R}_h) - \gamma_i) \\ & + \sum_{j \in \mathcal{G}_2} \log(\gamma_j - \text{Tr}(\mathbf{W}_j \mathbf{R}_h)) + \log \det \mathbf{R}_h \end{aligned}$$

where $\mathbf{W}_i = \mathbf{w}_i \mathbf{w}_i^H$ and the term $\mathbf{w}_i^H \mathbf{R}_h \mathbf{w}_i$ has been rewritten as $\text{Tr}(\mathbf{W}_i \mathbf{R}_h)$. Π_1 is a convex optimization problem which obtains the analytic center of the feasible region at time slot t formed by the linear inequalities (3), (4) and the positive semi-definite cone [19], [20]. It can be solved efficiently using interior point methods with worst case complexity $\mathcal{O}(N^7)$.

A) *Design of Beamforming Vector \mathbf{w}_{t+1} and Threshold γ_{t+1} :*

- *Design of beamforming vector \mathbf{w}_{t+1} :* After updating $\hat{\mathbf{R}}_h(t)$, we propose to select \mathbf{w}_{t+1} as follows.

$$\Pi_2 : \quad \mathbf{w}_{t+1} = \arg \max_{\|\mathbf{w}\|=1} \mathbf{w}^H \hat{\mathbf{R}}_h(t) \mathbf{w} - \lambda_t \mathbf{w}^H \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H \mathbf{w}$$

where $\mathbf{V}_{w,t} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t]$, and λ_t is a non-increasing function of t , e.g., $\lambda_t = \frac{\lambda_1}{[0.1t]}$, with $\lambda_1 \gg 1$. The solution of Π_2 can be obtained in closed form, i.e., \mathbf{w}_{t+1} is the unit vector along the principal eigenvector of the matrix $\hat{\mathbf{R}}_h(t) - \lambda_t \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H$. The objective function in Π_2 consists of two terms, the first one is proportional to the Tx-side estimate of the average received SNR (which is close to the actual average received SNR if the Tx has estimated $\hat{\mathbf{R}}_h(t)$ close to \mathbf{R}_h), and the second one is the squared norm of the vector $\mathbf{V}_{w,t}^H \mathbf{w}$. Maximization of this objective function gives a weight vector that strikes a balance between maximizing the estimated average received SNR and minimizing similarity to the weight vectors chosen in previous time slots. For small t , the choice of weight vector is dictated by $\mathbf{V}_{w,t}^H \mathbf{w}$, yielding diverse weight vectors that explore different directions, gathering information about \mathbf{R}_h ; for large t , $\lambda_t \ll 1$ and preference shifts to the first term $\mathbf{w}^H \hat{\mathbf{R}}_h(t) \mathbf{w}$, resulting in weight vectors aligned with the principal eigenvector of $\hat{\mathbf{R}}_h(t)$. Therefore, if $\hat{\mathbf{R}}_h(t) \rightarrow \mathbf{R}_h$ as $t \rightarrow \infty$, the beamforming vector chosen by the Tx will asymptotically align itself with the principal eigenvector of \mathbf{R}_h , thus attaining the maximum average received SNR.

- *Design of threshold γ_{t+1} : Analytic Center Cutting Plane Method (ACCPM).* After choosing \mathbf{w}_{t+1} , the Tx selects an appropriate SNR threshold γ_{t+1} such that the subsequent inequality constraint for \mathbf{R}_h obtained from the 1-bit feedback at time slot $t+1$ considerably reduces the feasible region at time t denoted by \mathcal{P}_t , where $\mathcal{P}_t = \{\mathbf{R} : \mathbf{R} \succeq 0, \mathbf{w}_i^H \mathbf{R} \mathbf{w}_i \geq \gamma_i, \forall i \in \mathcal{G}_1, \mathbf{w}_i^H \mathbf{R} \mathbf{w}_i < \gamma_i, \forall i \in \mathcal{G}_2, \mathcal{G}_1 \cup \mathcal{G}_2 = \{1, 2, \dots, t\}\}$. This is crucial for the convergence of $\hat{\mathbf{R}}_h(t)$ to \mathbf{R}_h . Since the Tx already communicates payload information to the Rx in parallel to learning to beamform, the new threshold can “piggyback” on the payload transmission at limited overhead—unlike the Rx feedback on the reverse link, which is more severely limited in terms of rate. (The basic method still works without having the Tx dictate thresholds to the Rx, albeit convergence to the true channel correlation matrix cannot be guaranteed in this case.)

One way to ensure that the feasible region is reduced at each time step is to choose \mathbf{w}_{t+1} and γ_{t+1} , such that the resulting hyperplane $\mathbf{w}_{t+1}^H \mathbf{R} \mathbf{w}_{t+1} = \gamma_{t+1}$ passes through an interior point of \mathcal{P}_t . Here, we propose to design the beamforming vector \mathbf{w}_{t+1} and the threshold γ_{t+1} such that the resulting hyperplane passes through the analytic center of \mathcal{P}_t (Analytic Center Cutting Plane Method—ACCPM), which is $\hat{\mathbf{R}}_h$. Since the analytic center is the point that maximizes the product of distances to the defining hyperplanes and the positive

semi-definite cone, it gives the deepest interior point of \mathcal{P}_t . Hence for a given \mathbf{w}_{t+1} , we choose $\gamma_{t+1} = \mathbf{w}_{t+1}^H \hat{\mathbf{R}}_{\mathbf{h}}(t) \times \mathbf{w}_{t+1}$. This ensures that the resulting cutting plane $\mathbf{w}_{t+1}^H \times \mathbf{R}_{\mathbf{h}} \mathbf{w}_{t+1} = \gamma_{t+1}$ will pass through $\hat{\mathbf{R}}_{\mathbf{h}}(t)$ and cut off a significant part of the current feasible region \mathcal{P}_t .

From known convergence results for ACCPM [20], [21], it then follows that $\hat{\mathbf{R}}_{\mathbf{h}}(t)$ (updated as the analytic center of the current feasible region) is restricted to a ball of radius r around $\mathbf{R}_{\mathbf{h}}$ within $\mathcal{O}\left(\frac{N^2}{r^2}\right)$ iterations. Therefore, if λ_t is designed so that it becomes negligible by $\left\lceil \frac{N^2}{r^2} \right\rceil$ iterations, then the objective function of $\mathbf{\Pi}_2$ can be approximated as $\mathbf{w}^H \hat{\mathbf{R}}_{\mathbf{h}}(t) \mathbf{w}$. Hence asymptotically, as $\hat{\mathbf{R}}_{\mathbf{h}}(t) \rightarrow \mathbf{R}_{\mathbf{h}}$, the beamforming weight vector will converge to the principal eigenvector of $\mathbf{R}_{\mathbf{h}}$ and the average received SNR will approach the maximum achievable average SNR (obtained with perfect *a priori* knowledge of $\mathbf{R}_{\mathbf{h}}$).

Remark 1: One can avoid the threshold communication overhead by fixing the threshold at the receiver, and scaling the transmit beamforming vector instead. This relinquishes control of transmission power, however, and is not as good a fit to our simultaneous exploration–exploitation setup, where power control is important and threshold information can piggyback on the payload. On the other hand, one can dispense with threshold communication when the transmission power can be allowed to vary freely, and this is in line with separate learning and payload phases, as in [18]. Threshold communication appears appealing from a practical point of view, taking into account the limited dynamic range of the receiver front-end, and power amplifier nonlinearities.

III. MAXIMUM LIKELIHOOD FORMULATION

In practice, bits may be flipped due to inaccurate SNR estimation at the Rx or communication errors in the reverse link. Assuming a memoryless feedback link, these errors are independent from slot to slot. We propose modeling both types of errors using an additive measurement noise model that is equivalent *from the point of view of the Tx*. Including this noise, the inequalities become

$$\begin{cases} \mathbf{w}_t^H \mathbf{R}_{\mathbf{h}} \mathbf{w}_t + n_t \geq \gamma_t, & \text{when } s_t = 1, \text{ or} \\ \mathbf{w}_t^H \mathbf{R}_{\mathbf{h}} \mathbf{w}_t + n_t < \gamma_t, & \text{when } s_t = 0, \end{cases} \quad (5)$$

where $n_t \sim \mathcal{N}(0, \sigma_n^2)$ is the equivalent noise at time t . The measurements received at the Tx are the bits $s_t = (\text{sign}(\mathbf{w}_t^H \mathbf{R}_{\mathbf{h}} \mathbf{w}_t + n_t - \gamma_t) + 1)/2$. The conditional likelihood of these bits s_1, s_2, \dots, s_t conditioned on the unknown parameter $\mathbf{R}_{\mathbf{h}}$ can be written as

$$\begin{aligned} f(\mathbf{s}_t | \mathbf{R}_{\mathbf{h}}) &= \prod_{i \in \mathcal{G}_1} \Pr[\text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}}) + n_i \geq \gamma_i] \\ &\quad \times \prod_{i \in \mathcal{G}_2} \Pr[\text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}}) + n_i < \gamma_i] \\ &= \prod_{i \in \mathcal{G}_1} \Phi\left(\frac{\text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}}) - \gamma_i}{\sigma_n}\right) \\ &\quad \times \prod_{i \in \mathcal{G}_2} \Phi\left(\frac{\gamma_i - \text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}})}{\sigma_n}\right), \end{aligned}$$

where $\mathbf{s}_t = [s_1, s_2, \dots, s_t]^T$, \mathcal{G}_1 and \mathcal{G}_2 are defined as before, and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz$ is the standard Gaussian c.d.f. At time slot $t+1$, $\hat{\mathbf{R}}_{\mathbf{h}}(t)$ is updated as the maximum likelihood estimate (MLE) $\hat{\mathbf{R}}_{\mathbf{h}}^{MLE}(t)$ (as compared to the analytic center in the error-free case) obtained from $\mathbf{\Pi}_3$ which maximizes the log-likelihood function $\log(f(\mathbf{s}_t | \mathbf{R}_{\mathbf{h}}))$ with a positive semi-definite constraint.

$$\mathbf{\Pi}_3 \quad \hat{\mathbf{R}}_{\mathbf{h}}^{MLE}(t) = \arg \max_{\mathbf{R}_{\mathbf{h}} \succeq 0} \sum_{i \in \mathcal{G}_1} \log \Phi\left(\frac{\text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}}) - \gamma_i}{\sigma_n}\right) + \sum_{i \in \mathcal{G}_2} \log \Phi\left(\frac{\gamma_i - \text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}})}{\sigma_n}\right)$$

$\mathbf{\Pi}_3$ is a convex optimization problem since it involves the maximization of the logarithm of the c.d.f. of a Gaussian distribution which is concave, with a positive semi-definite constraint which is convex. Once the channel correlation matrix estimate is updated, \mathbf{w}_{t+1} is chosen as the principal eigenvector of $\hat{\mathbf{R}}_{\mathbf{h}}^{MLE}(t) - \lambda_t \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H$ and $\gamma_{t+1} = \text{Tr}(\mathbf{W}_{t+1} \hat{\mathbf{R}}_{\mathbf{h}}^{MLE}(t))$, where $\mathbf{V}_{w,t} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t]$ and λ_t is a non-increasing function of t .

Our 1-bit/slot measurement model is a special case of what is known as a *probit model*. Statistical identifiability and MLE consistency conditions and proof for the probit model can be found in [22], and a more compact proof for a generalized model can be found in [23]. The basic idea is that, by the law of large numbers, the normalized log-likelihood function will converge to its expectation, and by the information inequality this will have a unique maximum at the true parameter when this is identifiable. However, the proof assumes that the regressors $\text{vec}(\mathbf{W}_t^T)$ are independently drawn from a distribution with nonsingular $E[\text{vec}(\mathbf{W}_t^T) \text{vec}(\mathbf{W}_t^T)^H]$. A random model for $\text{vec}(\mathbf{W}_t^T)$ is needed to invoke the information inequality. In our context, however, the $\text{vec}(\mathbf{W}_t^T)$'s are iteratively generated—in fact, judiciously designed—based on interim ML estimates of the sought channel correlation matrix, according to the proposed exploration–exploitation trade-off schedule. In certain cases, one can prove consistency of the MLE designed for independent and identically distributed (i.i.d.) data, but operating on non-i.i.d. data [24], [25]. We do not have proof of convergence and consistency of the MLE in our context, but our experiments indicate that the MLE approaches the true $\mathbf{R}_{\mathbf{h}}$ as the number of feedback bits increases.

A. Tracking Changes in $\mathbf{R}_{\mathbf{h}}$

When the channel correlation matrix $\mathbf{R}_{\mathbf{h}}$ changes over time due to mobility, the Tx should be capable of tracking these changes and adapting its beamforming vector to maintain high average SNR at the Rx. Assuming that $\mathbf{R}_{\mathbf{h}}$ changes slowly with time, it is natural to consider the following “discounted” modification of the MLE in $\mathbf{\Pi}_3$.

$$\mathbf{\Pi}_3' \quad \hat{\mathbf{R}}_{\mathbf{h}}^{dMLE}(t) = \arg \max_{\mathbf{R}_{\mathbf{h}} \succeq 0} \sum_{i \in \mathcal{G}_1} \beta^{t-i} \log \Phi\left(\frac{\text{Tr}(\mathbf{W}_i \mathbf{R}_{\mathbf{h}}) - \gamma_i}{\sigma_n}\right) + \sum_{j \in \mathcal{G}_2} \beta^{t-j} \log \Phi\left(\frac{\gamma_j - \text{Tr}(\mathbf{W}_j \mathbf{R}_{\mathbf{h}})}{\sigma_n}\right)$$

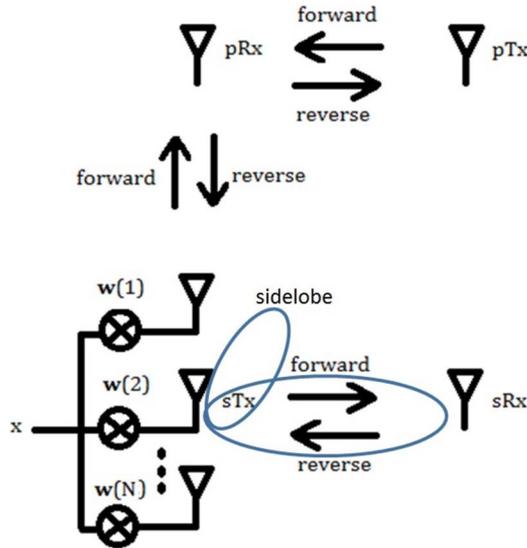


Fig. 1. Secondary beamforming schematic: primary system (#1, top) and secondary system (#2, bottom).

where $0 < \beta < 1$ and $\mathcal{G}_1 \cup \mathcal{G}_2 = \{1, 2, \dots, t\}$. Each term inside the summation of the objective function is weighted by a forgetting factor that decays exponentially with time. As a result, the terms corresponding to inequalities obtained from the recent past are given higher weight.

IV. COGNITIVE BEAMFORMING AND INTERFERENCE AVOIDANCE FROM BINARY CSIT

Consider a cognitive radio underlay scenario where the secondary and the primary Tx share the same frequency spectrum for communicating with the corresponding receivers. The constraint for the secondary Tx in this scenario is that it should transmit in such a way that the interference caused due to its transmission at the primary Rx is restricted below a primary interference threshold. In our cognitive radio setup, the secondary system consists of a multi-antenna Tx serving a single antenna Rx, coexisting with a primary system comprising a Tx–Rx pair, see Fig. 1. Let \mathbf{h}_{ij} denote the random complex channel from the Tx of system i to the Rx of system j and \mathbf{R}_{ij} be the auto-correlation matrix of \mathbf{h}_{ij} , $\{i, j\} \in \{1, 2\}$, where index 1 refers to the primary system, and index 2 to the secondary system. Initially, the secondary Tx does not know \mathbf{R}_{22} and \mathbf{R}_{21} , the correlation matrices of the channels to the secondary Rx and to the primary Rx, respectively. Its goal is to design beamforming vectors for learning these matrices while sending payload data to the secondary Rx, with the ultimate objective of maximizing the average SNR at the secondary Rx without seriously degrading the QoS of the primary Rx. The challenge here for the secondary Tx is that the primary Rx cannot be assumed to cooperate in the learning process. Note that we do not assume channel reciprocity, as this does not hold with frequency-division duplex or when nodes employ different transmit and receive beam patterns. The secondary Rx (sRx) operates as in the case of an isolated MISO link, considered earlier. In addition to the feedback bits from the sRx, the secondary Tx (sTx) can overhear the regular ACK-

NACK feedback that the primary Rx (pRx) sends to the primary Tx (pTx) over its reverse link. We consider two scenarios in what follows, depending on whether or not the sTx knows the primary interference threshold τ_p . Note that, in the absence of third-party interference, τ_p primarily depends on the modulation and coding scheme employed over the primary link, which are often fixed in legacy systems. Assuming that τ_p is known to the secondary system is far less realistic, on the other hand, if the primary system uses adaptive modulation and coding.

Case 1—sTx Knows Primary Interference Threshold τ_p : At time slot t , the sTx employs beamforming vector \mathbf{w}_t . Let \mathcal{G}_{p1} and \mathcal{G}_{p0} be the set of time slots where the pRx sends an ACK and NACK feedback, respectively to the pTx. Listening in to the primary's reverse link, the sTx infers that

$$\begin{aligned} \mathbf{w}_i^H \mathbf{R}_{21} \mathbf{w}_i &< \tau_p \quad \forall i \in \mathcal{G}_{p1} \\ \mathbf{w}_j^H \mathbf{R}_{21} \mathbf{w}_j &\geq \tau_p \quad \forall j \in \mathcal{G}_{p0} \\ \iff \text{Tr}(\mathbf{W}_i \mathbf{R}_{21}) &< \tau_p \quad \forall i \in \mathcal{G}_{p1} \\ \text{Tr}(\mathbf{W}_j \mathbf{R}_{21}) &\geq \tau_p \quad \forall j \in \mathcal{G}_{p0}. \end{aligned} \quad (6)$$

On the other hand, the sTx also receives 1-bit feedback from the sRx, yielding the following inequalities.

$$\begin{aligned} \text{Tr}(\mathbf{W}_i \mathbf{R}_{22}) &\geq \gamma_i, \quad \forall i \in \mathcal{G}_{s1} \\ \text{Tr}(\mathbf{W}_i \mathbf{R}_{22}) &< \gamma_i, \quad \forall i \in \mathcal{G}_{s2} \end{aligned} \quad (7)$$

where $\mathcal{G}_{s1} = \{i : i \in \{1, 2, \dots, t\}, s_i = 1\}$ and $\mathcal{G}_{s2} = \{i : i \in \{1, 2, \dots, t\}, s_i = 0\}$. For every time slot t , the secondary Tx has to update its estimate of \mathbf{R}_{22} and \mathbf{R}_{21} and design the beamforming vector \mathbf{w}_t such that the average received signal power of the sRx is maximized without causing excessive interference to the pRx on average.

At time $t + 1$, the sTx updates $\hat{\mathbf{R}}_{22}(t + 1)$ and $\hat{\mathbf{R}}_{21}(t + 1)$ as follows

$$\begin{aligned} \hat{\mathbf{R}}_{22}(t + 1) &= \arg \max_{\mathbf{R}_{22}} \sum_{i \in \mathcal{G}_{s1}} \log(\text{Tr}(\mathbf{W}_i \mathbf{R}_{22}) - \gamma_i) \\ &\quad + \sum_{j \in \mathcal{G}_{s2}} \log(\gamma_j - \text{Tr}(\mathbf{W}_j \mathbf{R}_{22})) + \log \det \mathbf{R}_{22} \end{aligned} \quad (8)$$

$$\begin{aligned} \hat{\mathbf{R}}_{21}(t + 1) &= \arg \max_{\mathbf{R}_{21}} \sum_{i \in \mathcal{G}_{p0}} \log(\text{Tr}(\mathbf{W}_i \mathbf{R}_{21}) - \tau_p) \\ &\quad + \sum_{j \in \mathcal{G}_{p1}} \log(\tau_p - \text{Tr}(\mathbf{W}_j \mathbf{R}_{21})) + \log \det \mathbf{R}_{21} \end{aligned} \quad (9)$$

where $\hat{\mathbf{R}}_{22}(t + 1)$ and $\hat{\mathbf{R}}_{21}(t + 1)$ are the estimates of \mathbf{R}_{22} and \mathbf{R}_{21} at time $t + 1$. From (8) and (9), it can be seen that $\hat{\mathbf{R}}_{22}(t + 1)$ and $\hat{\mathbf{R}}_{21}(t + 1)$ are the analytic centers of the feasible region formed by the associated linear inequalities in (7) and (6) and the positive semi-definite cone [19], [20].

We propose the following steps to design the beamforming vector \mathbf{w}_{t+1} . First we solve

$$\begin{aligned} \Pi_6 \tilde{\mathbf{w}}_{t+1} &= \arg \max_{\mathbf{w}} \left[\mathbf{w}^H \hat{\mathbf{R}}_{22}(t + 1) \mathbf{w} - \lambda_t \mathbf{w}^H \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H \mathbf{w} \right] \\ \text{s.t.} \quad \mathbf{w}^H \hat{\mathbf{R}}_{21}(t + 1) \mathbf{w} &\leq \tau_p \end{aligned} \quad (10)$$

$$\|\mathbf{w}\|^2 \leq P_w \quad (11)$$

where $\mathbf{V}_{w,t} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t]$, P_w is the maximum available transmit power at the sTx and λ_t is a non-increasing function of t . By solving $\mathbf{\Pi}_6$, we get a $\tilde{\mathbf{w}}_{t+1}$ that not only maximizes the sTx-side estimate of average received signal power at sRx, but also is diverse enough to explore the channel correlation space, gain more information about \mathbf{R}_{22} and \mathbf{R}_{21} and eventually improve the accuracy of their estimates, while limiting the average interference to pRx below τ_p . $\mathbf{\Pi}_6$ involves the maximization of an indefinite quadratic objective function subject to two convex quadratic constraints. From the results in [7], the optimal solution to $\mathbf{\Pi}_6$ can be obtained by using semi-definite relaxation (SDR) and solving $\mathbf{\Pi}_{6a}$.

$$\mathbf{\Pi}_{6a} \tilde{\mathbf{W}}_{t+1} = \arg \max_{\mathbf{W} \succeq 0} \left[\text{Tr}(\mathbf{W}\hat{\mathbf{R}}_{22}) - \lambda_t \text{Tr}(\mathbf{W}\mathbf{V}_{w,t}\mathbf{V}_{w,t}^H) \right]$$

$$s.t. \quad \text{Tr}(\mathbf{W}\hat{\mathbf{R}}_{21}) \leq \tau_p \quad (12)$$

$$\text{Tr}(\mathbf{W}) \leq P_w \quad (13)$$

where $\mathbf{W} = \mathbf{w}\mathbf{w}^H$ and $\tilde{\mathbf{w}}_{t+1}$ can be obtained as the principal eigenvector of $\tilde{\mathbf{W}}_{t+1}$. Notice that semi-definite relaxation incurs no loss of optimality here, as shown in [7]. Once $\tilde{\mathbf{w}}_{t+1}$ is obtained, \mathbf{w}_{t+1} is designed as follows.

$$\mathbf{w}_{t+1} = \left(\sqrt{\frac{\tau_p}{\tilde{\mathbf{w}}_{t+1}^H \hat{\mathbf{R}}_{21}(t+1) \tilde{\mathbf{w}}_{t+1}}} \right) \tilde{\mathbf{w}}_{t+1}. \quad (14)$$

This step is necessary to ensure that the hyperplane $\mathbf{w}_{t+1}^H \mathbf{R}\mathbf{w}_{t+1} = \tau_p$ passes through $\mathbf{R} = \hat{\mathbf{R}}_{21}(t+1)$.

Remark 2: Here, the scaling step is adopted because the primary interference threshold τ_p is fixed and cannot be varied by the sTx, see also remark 1. The drawback is that transmission power cannot be directly controlled this way. In contrast to the case of an isolated MISO link that we considered earlier, this is necessary here because we wish to drive two parallel cutting plane iterations with a common beamforming vector and threshold, thus we must use all degrees of freedom we have.

After designing \mathbf{w}_{t+1} , the threshold for the sRx γ_{t+1} is designed as follows

$$\gamma_{t+1} = \mathbf{w}_{t+1}^H \hat{\mathbf{R}}_{22}(t+1) \mathbf{w}_{t+1}. \quad (15)$$

This ensures that the hyperplane $\mathbf{w}_{t+1}^H \mathbf{R}\mathbf{w}_{t+1} = \gamma_{t+1}$ passes through $\mathbf{R} = \hat{\mathbf{R}}_{22}(t+1)$.

Convergence of \mathbf{w}_{t+1} to the Optimal Beamforming Vector: The design of \mathbf{w}_{t+1} and γ_{t+1} ensures that the hyperplanes corresponding to the inequalities inferred by the sTx upon receiving the feedback bit and overhearing the primary ACK-NACK feedback for time $t+1$ pass through the analytic centers of the feasible regions of \mathbf{R}_{22} and \mathbf{R}_{21} , i.e., $\hat{\mathbf{R}}_{22}(t+1)$ and $\hat{\mathbf{R}}_{21}(t+1)$, respectively. Hence convergence of ACCPM [21] can be invoked. As a result, $\hat{\mathbf{R}}_{22}(t+1)$ and $\hat{\mathbf{R}}_{21}(t+1)$ are confined to a ball of radius r centered around \mathbf{R}_{22} and \mathbf{R}_{21} respectively, within $\mathcal{O}\left(\frac{N^2}{r^2}\right)$ iterations. Furthermore, λ_t decreases to zero as $t \rightarrow \infty$. Therefore, as $\hat{\mathbf{R}}_{22}(t+1) \rightarrow \mathbf{R}_{22}$, $\hat{\mathbf{R}}_{21}(t+1) \rightarrow \mathbf{R}_{21}$ and $\lambda_t \rightarrow 0$, $\mathbf{\Pi}_6$ becomes

$$\tilde{\mathbf{w}} = \arg \max_{\|\mathbf{w}\|^2 \leq P_w} \mathbf{w}^H \mathbf{R}_{22} \mathbf{w}$$

$$s.t. \quad \mathbf{w}^H \mathbf{R}_{21} \mathbf{w} \leq \tau_p$$

which is the optimal secondary beamforming vector design when sTx has perfect knowledge of \mathbf{R}_{22} and \mathbf{R}_{21} . Note that this is a special kind of non-convex problem (“one constraint away” from the Rayleigh quotient) that can be solved exactly [7].

Case 2—sTx Does Not Know τ_p : Here we assume that the sTx does not have any knowledge about τ_p . In this case, by overhearing the primary ACK-NACKs, the sTx infers the following inequalities

$$\mathbf{w}_i^H \mathbf{R}_{21} \mathbf{w}_i \leq \mathbf{w}_j^H \mathbf{R}_{21} \mathbf{w}_j$$

$$\iff \text{Tr}(\mathbf{W}_i \mathbf{R}_{21}) \leq \text{Tr}(\mathbf{W}_j \mathbf{R}_{21}), \quad \forall i \in \mathcal{G}_{p1}, \forall j \in \mathcal{G}_{p0}. \quad (16)$$

From the feedback bits sent by the sRx to the sTx, the sTx infers the inequalities mentioned in (7). As before, the sTx updates $\hat{\mathbf{R}}_{22}(t+1)$ as the analytic center of the associated linear inequalities and the positive semidefinite cone as shown in (8). For designing the beamforming vector \mathbf{w}_{t+1} and updating $\hat{\mathbf{R}}_{21}(t+1)$, we propose the following formulation:

$$\mathbf{\Pi}_7 : \max_{\mathbf{w}, \mathbf{R}_{21} \succeq 0}$$

$$\left[\mathbf{w}^H \hat{\mathbf{R}}_{22}(t+1) \mathbf{w} - \mu \mathbf{w}^H \mathbf{R}_{21} \mathbf{w} - \lambda_t \mathbf{w}^H \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H \mathbf{w} \right]$$

$$s.t. \quad \|\mathbf{w}\|^2 \leq P_w \quad (17)$$

$$\text{Tr}(\mathbf{W}_i \mathbf{R}_{21}) \leq \text{Tr}(\mathbf{W}_j \mathbf{R}_{21}), \quad \forall i \in \mathcal{G}_{p1}, \forall j \in \mathcal{G}_{p0} \quad (18)$$

where $\mu \in \mathcal{R}_+$ and λ_t is a non-increasing function of t . The motivation for the objective function in $\mathbf{\Pi}_7$ is as follows. The first term is the sTx-side estimate of the signal power received at the sRx. For a given \mathbf{w} , the second term in the objective of $\mathbf{\Pi}_7$ selects from the admissible $\hat{\mathbf{R}}_{21}(t+1)$ the one that is most favorable from the sTx point of view; that is, the one that is annoyed the least by \mathbf{w}_{t+1} . The third term diversifies the choice of \mathbf{w} to explore the channel correlations space from as many directions as possible to gather information about \mathbf{R}_{22} and \mathbf{R}_{21} , and eventually improve the accuracy of the estimates.

Since $\mathbf{\Pi}_7$ is not jointly convex in \mathbf{w} and \mathbf{R}_{21} , it can be tackled using alternating optimization. The update of \mathbf{R}_{21} is semi-definite programming (SDP). For a fixed $\mathbf{R}_{21} = \hat{\mathbf{R}}_{21}(t+1)$, the update of \mathbf{w} is simply the unit vector along the principal eigenvector of $\hat{\mathbf{R}}_{22}(t) - \mu \hat{\mathbf{R}}_{21}(t+1) - \lambda_t \mathbf{V}_{w,t} \mathbf{V}_{w,t}^H$. The threshold γ_{t+1} for the sRx at time $t+1$ can be obtained using ACCPM, namely $\gamma_{t+1} = \text{Tr}(\mathbf{W}_{t+1} \hat{\mathbf{R}}_{22}(t+1))$, where $\mathbf{W}_{t+1} = \mathbf{w}_{t+1} \mathbf{w}_{t+1}^H$.

Since the update of $\hat{\mathbf{R}}_{22}(t)$ is independent of $\hat{\mathbf{R}}_{21}(t)$ and there is no dependency in terms of the constraints, convergence of $\hat{\mathbf{R}}_{22}(t)$ follows from the ACCPM, as before. On the other hand, the constraints determining the feasible region for \mathbf{R}_{21} are all homogeneous, $\text{Tr}(\mathbf{W}_m \mathbf{R}_{21}) \leq \text{Tr}(\mathbf{W}_n \mathbf{R}_{21})$, $\forall m \in \mathcal{G}_{p1}, \forall n \in \mathcal{G}_{p0}$ and $\mathbf{R}_{21} \succeq 0$. The hyperplanes corresponding to the linear inequalities all pass through the origin, so the feasible region remains unbounded for all t . As a result, there is no hope that $\hat{\mathbf{R}}_{21}(t)$ will converge to \mathbf{R}_{21} , or that the interference to the pRx will converge below its (unknown) interference tolerance level. On the other hand, with appropriate choice of the sequence of the \mathbf{W}_i 's (\iff the \mathbf{w}_i 's), there is hope that

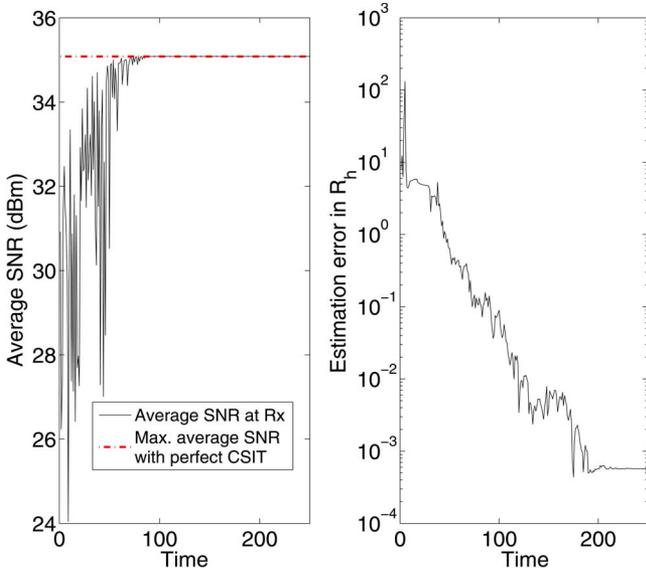


Fig. 2. Evolution of average link SNR (left) and $\|\mathbf{R}_h - \hat{\mathbf{R}}_h\|_F$ (right) for ACCPM, $N = 5$.

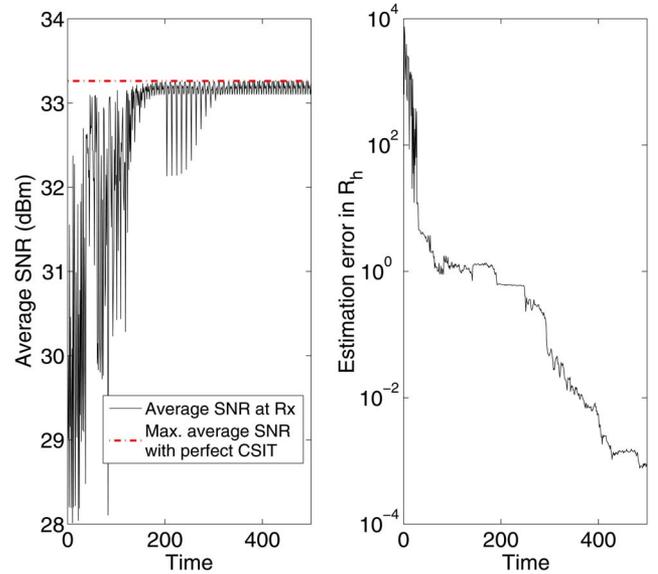


Fig. 3. Evolution of average link SNR (left) and $\|\mathbf{R}_h - \hat{\mathbf{R}}_h\|_F$ (right) for ACCPM, $N = 10$.

the *direction* of $\text{vec}(\hat{\mathbf{R}}_{21}(t))$ will align with that of $\text{vec}(\mathbf{R}_{21})$, despite the fact that the scale cannot be recovered. This suggests using an additional interference management mechanism to limit sTx to pRx interference, when needed. Towards this end, we may: (a) *Fix μ and vary P* at slot $t + 1$ based on whether an ACK or NACK was heard from the pRx at slot t ; i.e., \mathbf{w}_{t+1} is scaled by $\sqrt{\alpha_{t+1}}$, where $\alpha_{t+1} = \alpha_t \alpha$, if a NACK was heard, else $\alpha_{t+1} = \max(\alpha_t/\alpha, P_w)$, with back-off parameter $\alpha < 1$. Alternatively, we may (b) *Fix P and vary μ* in Π_5 , thereby changing the relative preference to directions that cause lower estimated average interference power to the pRx. Whenever the sTx hears a NACK from the pRx, it sets $\mu_{t+1} = \mu_t \delta$, while an ACK results in $\mu_{t+1} = \mu_t / \delta$, with $\delta > 1$.

As we will see, simulations show that the fix μ vary P power back-off scheme can learn the primary interference threshold and approach optimal secondary link performance—albeit we do not have proof of convergence in this case.

V. SIMULATION RESULTS

A. MISO Link

Fig. 2 shows simulation results for the average link SNR and the estimation error $\|\mathbf{R}_h - \hat{\mathbf{R}}_h\|_F$ for a point to point MISO link using $N = 5$ transmit antennas, using ACCPM. The x-axis is labeled (slow) “time,” i.e., the slot index. For simulation purposes, the channel vector \mathbf{h} is drawn independently during each time slot from a complex normal distribution with zero mean and covariance matrix \mathbf{R}_h and $\lambda_t = \frac{5}{[0.1t]}$. The channel correlation matrix \mathbf{R}_h was obtained by generating a random orthonormal matrix \mathbf{U} , a random diagonal matrix \mathbf{D} with positive real numbers along the main diagonal, and setting $\mathbf{R}_h = \mathbf{U}\mathbf{D}\mathbf{U}^H$. The dotted line in the figure on the left represents the maximum achievable average SNR with perfect knowledge of \mathbf{R}_h at the Tx. The solid line represents the average received SNR at each time t . It takes approximately 80 time slots (or $(80/(5(5 + 1)/2)) \approx 6$ feedback bits

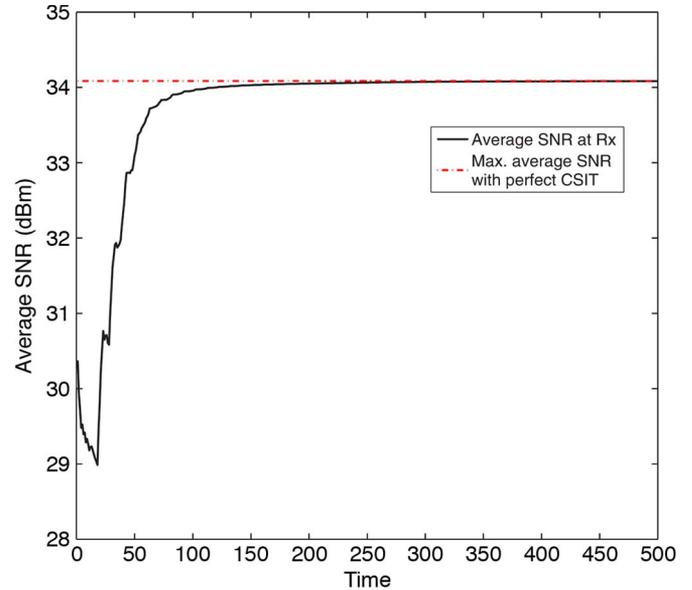


Fig. 4. Monte Carlo simulation for evolution of average link SNR in isolated MISO link, $N = 5$.

per complex entry of \mathbf{R}_h) for the algorithm to converge to the maximum achievable SNR. Fig. 3 shows corresponding results for $N = 10$ transmit antennas. The time taken by the algorithm to converge to the maximum achievable SNR at the Rx increases as N increases (approx 200 time slots for $N = 10$, or 7 bits per complex entry of \mathbf{R}_h). Furthermore, it can be seen from the figures on the right that the channel correlation matrix estimation error also decreases with t . Fig. 4 plots the Monte-Carlo simulation of the average link SNR for $N = 5$ by averaging over 100 random realizations of \mathbf{R}_h .

Fig. 5 compares the average SNR performance of the a) proposed ACCPM algorithm with b) the distributed beamforming algorithm in [12], b) the gradient sign algorithm in [13], and the one-bit null space learning algorithm in [16], for an isolated MISO link with $N = 5$ transmit antennas. The figure

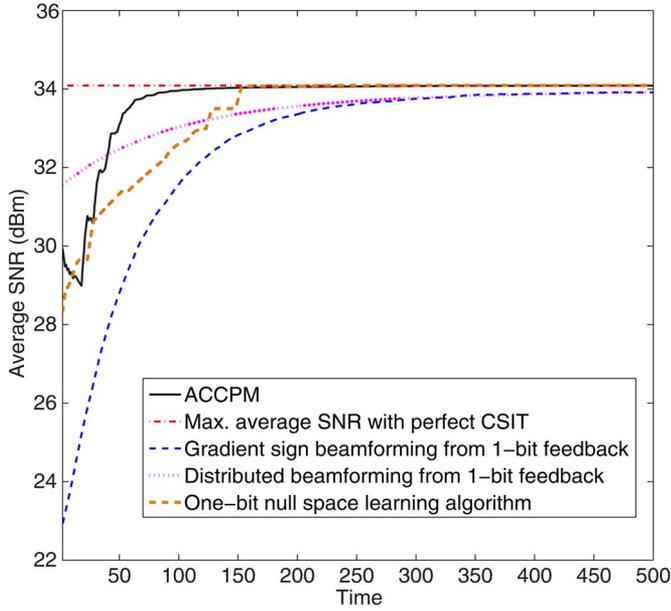


Fig. 5. Performance comparison of average link SNR in isolated MISO link for a) proposed ACCPM algorithm, b) distributed beamforming from 1-bit feedback, c) gradient sign algorithm, and d) one-bit null space learning algorithm, $N = 5$.

report results averaged over 100 Monte Carlo draws of \mathbf{R}_h . It can be seen that the proposed algorithm converges much faster to the maximum SNR than the other algorithms (100 versus 500, 500 and 150). The performance of the proposed algorithm is superior to algorithms b) and c) because it uses a better exploration technique. The performance gain of algorithm a) over d) can be attributed to the faster convergence of the ACCPM in comparison to the Cyclic-Jacobi algorithm ($\mathcal{O}(N^2)$ versus $\mathcal{O}(N^2 \log N)$ [16]). Furthermore, it can also be seen that there is a slight gap in the maximum SNR and the maximum value attained by the algorithms in [12] and [13], even after 500 iterations.

The average received SNR and the estimation error for \mathbf{R}_h using the MLE formulation is plotted in Fig. 6 for $N = 5$ and $\sigma_n = 0.01$. As mentioned in the captions, there were 86 bit flips among the 500 feedback bits (17%). It can be seen that even in the presence of bit flips, $\hat{\mathbf{R}}_h^{MLE}(t)$ approaches \mathbf{R}_h , resulting in the average received SNR at the user's side approaching the maximum achievable SNR with perfect knowledge of \mathbf{R}_h . However, the time taken for convergence of $\hat{\mathbf{R}}_h^{MLE}(t)$ to \mathbf{R}_h is higher than in the case without errors (Fig. 2), i.e., 250 with errors versus 80 without errors. One reason for the slower convergence rate of the MLE is the bit flips occurring as a result of the noise. Another reason is that, unlike the ACCPM, the MLE is not explicitly designed to quickly cut down the feasible region. Fig. 7 shows the average SNR obtained using the MLE update for different σ_n values. The performance of the ACCPM algorithm assuming perfect data (no bit flips) is also plotted as a baseline for comparison. The plots have been averaged over 100 Monte Carlo runs. It can be seen that as σ_n increases, the time taken for convergence also increases. This is because of the increased number of bit flips due to higher noise variance. However, it is interesting to note that even with a small number of bit flips (i.e., when $\sigma_n = 10^{-4}$, average bit-flips = 4), the

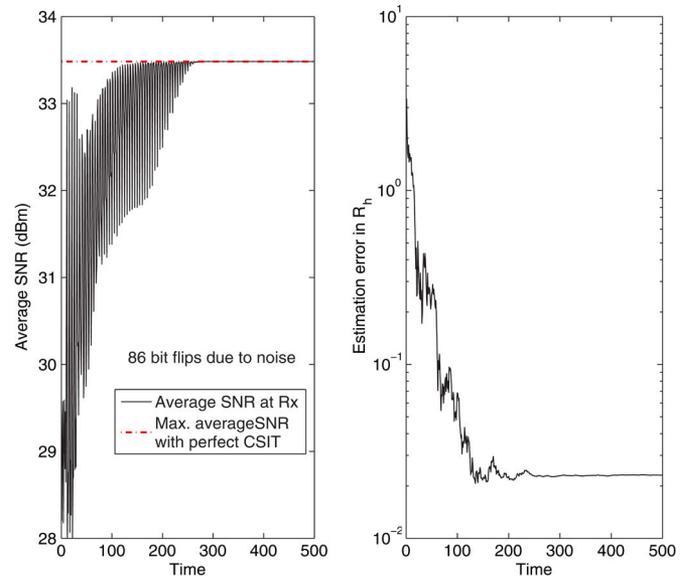


Fig. 6. Evolution of average link SNR (left) and $\|\mathbf{R}_h - \hat{\mathbf{R}}_h\|_F$ (right) for MLE, $N = 5$.

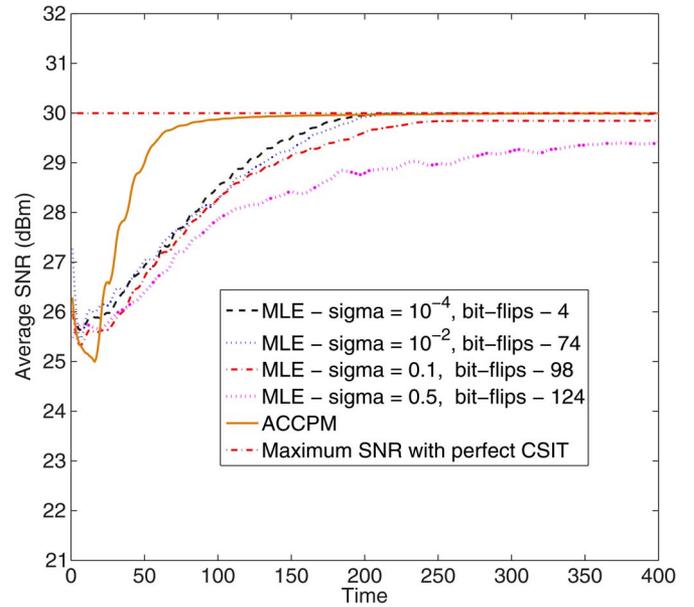


Fig. 7. Performance comparison of average SNR using MLE update for different σ_n , $N = 5$.

convergence of the MLE algorithm is slower compared to the ACCPM algorithm. On the other hand, it should also be noted that the noise can render the set of inequalities infeasible—in which case the ACCPM is no longer applicable, but MLE still works and manages to approach the optimum solution, without explicitly rejecting the conflicting inequalities. This is pretty remarkable, as it addresses an important practical concern in our context.

Fig. 8 plots a Monte-Carlo average link SNR using the MLE for $N = 5$, averaging over multiple random realizations of \mathbf{R}_h .

Fig. 9 shows the average SNR performance for $N = 3$ when \mathbf{R}_h changes with time and the discounted MLE formulation is used. For simulation purposes, \mathbf{R}_h is changed to a new correlation matrix at $t = 400$. The channel correlation matrices

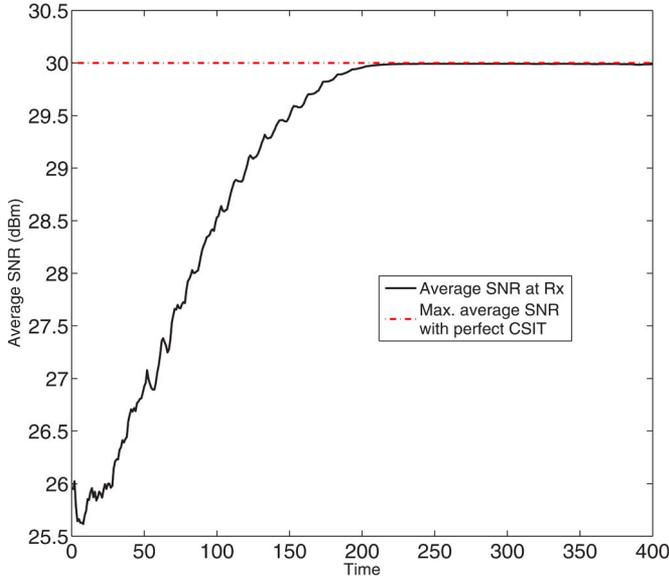


Fig. 8. Monte Carlo simulation for evolution of average link SNR using MLE formulation in isolated MISO link, $N = 5$.

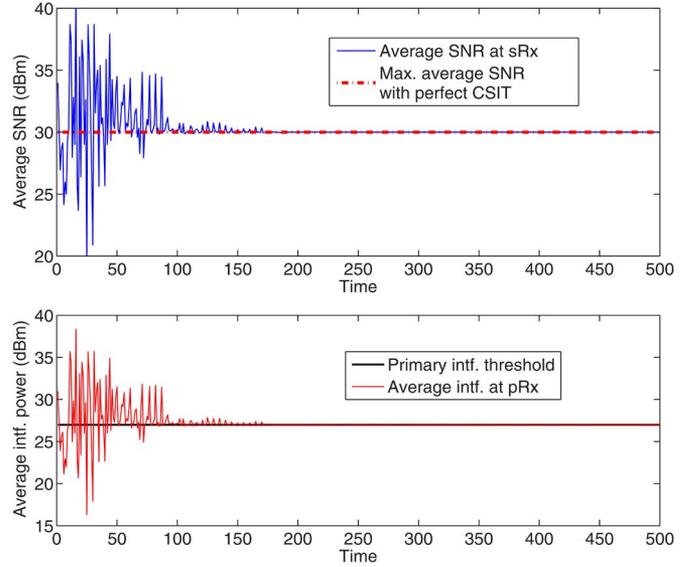


Fig. 10. τ_p is known at sTx—Avg. SNR at sRx (top) and avg. interference power at pRx (bottom), $N = 5$, $\tau_p = 0.5$.

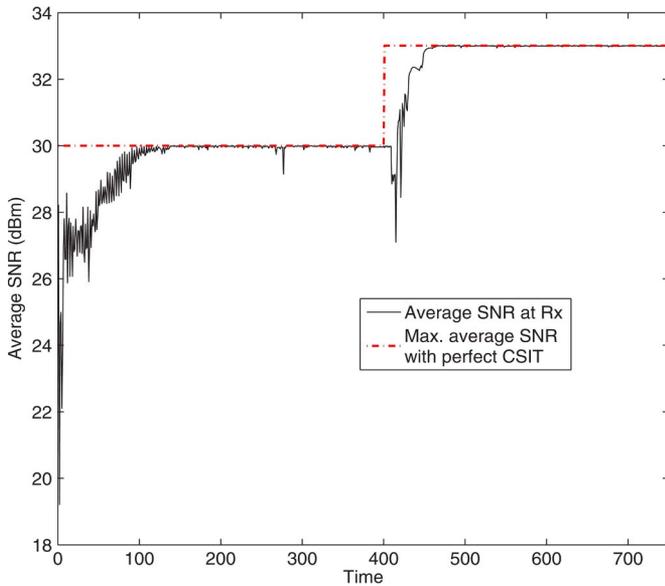


Fig. 9. Channel correlation tracking performance using the discounted MLE, $N = 3$.

for the simulation were generated in the same fashion as in the point to point MISO case. For the plots shown in Fig. 9, the value of $\beta = 0.95$. It can be seen from Fig. 9 that the discounted MLE formulation is able to track the changes in \mathbf{R}_h and adapt the beamforming vectors to achieve the maximum SNR at the receiver.

B. Cognitive Radio Underlay Simulations

Fig. 10 shows the simulation results for the case when τ_p is known to the sTx. The top plot shows the average received SNR at the sRx and the bottom plot shows the average interference power at the pRx for $N = 5$, $\lambda_t = \frac{5}{[0.1t]}$, $\tau_p = 0.5$ and $P_w = 5$. For simulation purposes, $\mathbf{h}_{ij} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{ij})$, $i, j \in \{1, 2\}$. \mathbf{R}_{22} and \mathbf{R}_{21} were obtained by generating random orthonormal

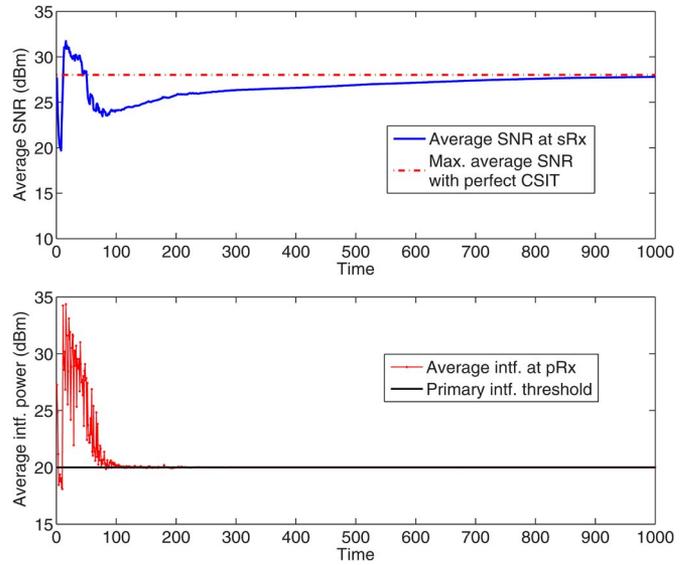


Fig. 11. τ_p is known at sTx—Monte Carlo simulation for Avg. SNR at sRx (top) and avg. interference power at pRx (bottom), $N = 5$, $\tau_p = 0.1$.

matrices \mathbf{U}_{22} and \mathbf{U}_{21} , random diagonal matrices \mathbf{D}_{22} and \mathbf{D}_{21} with positive diagonals, and setting $\mathbf{R}_{22} = \mathbf{U}_{22}\mathbf{D}_{22}\mathbf{U}_{22}^H$, $\mathbf{R}_{21} = \mathbf{U}_{21}\mathbf{D}_{21}\mathbf{U}_{21}^H$. The dotted horizontal line in the top plot is the maximum achievable average received SNR at the sRx (with perfect knowledge of \mathbf{R}_{22} , \mathbf{R}_{21} and τ_p at the sTx). The horizontal line in the bottom in Fig. 10 represents the primary interference power threshold τ_p . It can be seen that the average SNR at sRx converges to the maximum achievable SNR value (obtained with perfect knowledge of \mathbf{R}_{22} and \mathbf{R}_{21} at sTx) and the average interference power at the pRx converges to τ_p . Fig. 11 plots the Monte-Carlo simulation of the average SNR at sRx for $N = 5$, $\tau_p = 0.1$ by averaging over multiple random realizations of \mathbf{R}_{22} and \mathbf{R}_{21} . It can be seen from Fig. 11 that the average SNR at the sRx attains the maximum average SNR and the average interference power at the pRx is limited to τ_p for every random realization of \mathbf{R}_{22} and \mathbf{R}_{21} .

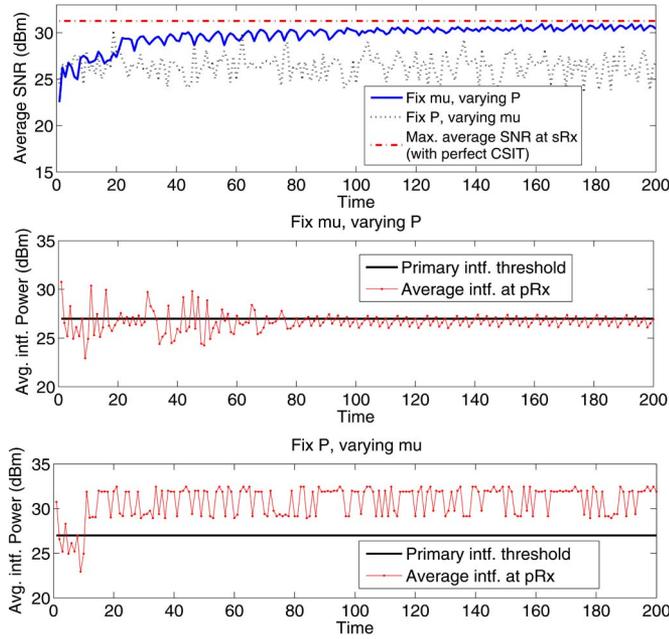


Fig. 12. τ_p is unknown at sTx—Avg. SNR at sRx (top) and interference power at pRx for two candidate back-off schemes, $N = 5$, $\tau_p = 0.5$.

Fig. 12 shows simulation results for the average received SNR at the sRx and the average interference power at the pRx when τ_p is unknown to the sTx with $N = 5$, $\lambda_t = \frac{5}{[0.1t]}$, $\tau_p = 0.5$, $P_w = 5$, and back-off parameters $\alpha = 0.8$ and $\delta = 2$. The dotted straight line in the top plot is the maximum achievable average received SNR at sRx (with perfect knowledge of \mathbf{R}_{22} , \mathbf{R}_{21} and τ_p at the sTx). The solid horizontal line in the two lower plots in Fig. 12 represents the primary interference power threshold which is not known to the sTx. It is very pleasing and intriguing to see that the proposed power back-off mechanism (a) approaches optimal performance in terms of sRx SNR, while the interference it causes to the pRx converges to the primary interference threshold τ_p , which is unknown! At the same time, the indirect back-off mechanism (b) clearly fails in this case—which speaks for the importance of choosing the right back-off scheme. Fig. 13 plots the Monte-Carlo simulation of average SNR at sRx and average interference power at pRx using power back-off mechanism (a) for $N = 5$, $\tau_p = 0.1$ by averaging over multiple random realizations of \mathbf{R}_{22} and \mathbf{R}_{21} . Note that there is a small gap relative to optimal performance in this case (32 versus 33.5 dBm).

Fig. 14 highlights the average SNR performance at the secondary Rx of a cognitive radio using interference management scheme (a) for two extreme cases. For case 1, \mathbf{R}_{22} has been chosen proportional to \mathbf{R}_{21} . This can occur when the secondary Rx and the primary Rx are aligned to each other when viewed from the secondary Tx. In this case it is generally difficult for the secondary Tx to design transmit beamforming vectors for providing high average SNR at the secondary Rx without causing excessive interference to the primary Rx. For case 2, \mathbf{R}_{22} and \mathbf{R}_{21} have been designed such that the principal eigenvector of \mathbf{R}_{22} is aligned with the minor eigenvector of \mathbf{R}_{21} , i.e., the direction of the eigenvector corresponding to the minimum eigenvalue of \mathbf{R}_{21} . This is a very desirable scenario for the sec-

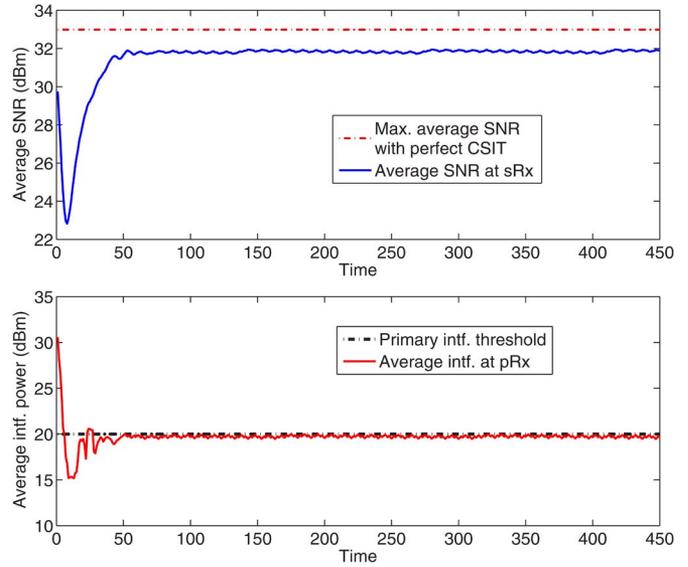


Fig. 13. τ_p is unknown at sTx—Monte Carlo simulation for Avg. SNR at sRx (top) and avg. interference power at pRx (bottom), $N = 5$, $\tau_p = 0.1$.

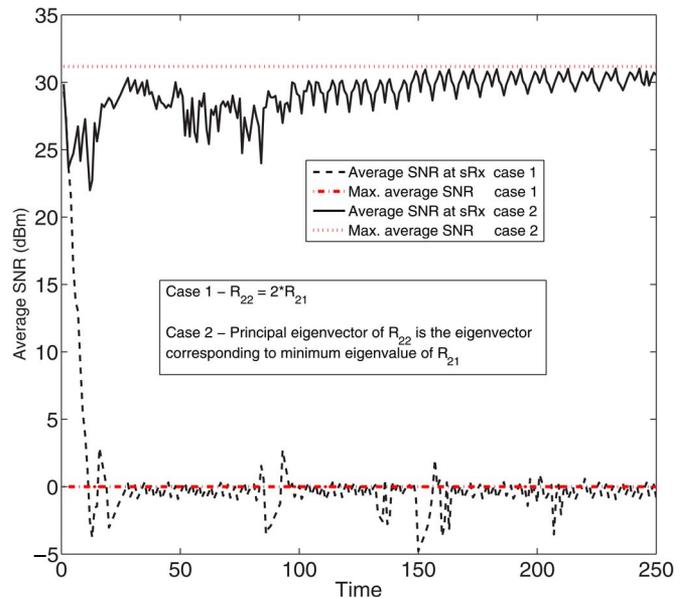


Fig. 14. Comparison of average SNR performance at sRx for $\tau_p = 5 \cdot 10^{-4}$.

ondary Tx because if it aligns the transmit beamforming vector along the principal eigenvector of \mathbf{R}_{22} , then it can achieve high average SNR at the secondary Rx as well as cause the least possible interference to the primary Rx simultaneously. For this simulation, the correlation matrix \mathbf{R}_{22} was generated in the same fashion as mentioned in the previous paragraph and \mathbf{R}_{21} was generated based on the conditions required for cases 1 and 2 mentioned here. It can be seen that there is approximately 30 dB difference in the maximum average SNR between these two cases, and that the secondary Tx achieves the maximum possible average SNR in both cases.

VI. CONCLUSION

In this paper, we have proposed an efficient way to accurately estimate the channel correlation matrix at the Tx of a

MISO link based on binary feedback from the Rx, obtained by comparing the average received SNR with a threshold that is varied adaptively by the Tx and communicated to the Rx. This algorithm is used for designing transmit beamforming vectors. The proposed technique is promising because the Tx starts without any CSI, and as time progresses, not only does it obtain an accurate estimate of the correlation matrix and the maximum-SNR beamformer without dedicated training, but it does so while transmitting payload in parallel with the learning process. A maximum likelihood formulation was proposed to accommodate measurement/feedback communication errors that can produce inconsistent inequalities, in which case the ACCPM is no longer applicable. A discounted maximum likelihood formulation was also proposed for tracking changes in the channel correlation matrix. Pertinent extensions to an underlay cognitive radio network setup were proposed for designing beamforming vectors at the secondary Tx to maximize the average received SNR at the secondary Rx without causing excessive interference to the primary Rx. Through relevant simulations, it was shown that the proposed algorithms for cognitive radio networks enable the secondary Tx to learn the relevant channel correlation matrices, starting from no CSI, and design beamformers to attain the maximum achievable SNR value at the secondary Rx, obtained when the secondary Tx has perfect knowledge of the primary interference threshold and channel correlation matrices to the secondary Rx and the primary Rx.

REFERENCES

- [1] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, 2010.
- [2] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.
- [3] R. A. Tannious and A. Nosratinia, "Cognitive radio protocols based on exploiting hybrid ARQ retransmissions," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2833–2841, Sep. 2010.
- [4] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
- [5] 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. 897–914, May 2009.
- [6] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
- [7] Y. Ye and S. Zhang, "New results on quadratic minimization," *SIAM J. Optim.*, vol. 14, no. 1, pp. 245–267, 2003.
- [8] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
- [9] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853–1863, May 2006.
- [10] R. Zhang, F. Gao, and Y. C. Liang, "Cognitive beamforming made practical: Effective interference channel and learning-throughput tradeoff," *IEEE Trans. Commun.*, vol. 8, no. 2, pp. 706–718, Feb. 2010.
- [11] F. Gao, R. Zhang, Y. C. Liang, and X. Wang, "Multi-antenna cognitive radio systems: Environmental learning and channel training," in *Proc. IEEE ICASSP*, Taipei, Taiwan, Apr. 19–24, 2009, pp. 2329–2332.
- [12] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 411–426, Jan. 2010.
- [13] B. C. Banister and J. R. Ziedler, "A simple gradient sign algorithm for transmit antenna weight adaptation with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1156–1171, May 2003.
- [14] R. Zhang, "On active learning and supervised transmission of spectrum sharing based cognitive radios by exploiting hidden primary radio feedback," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2960–2970, Oct. 2010.
- [15] Y. Noam and A. J. Goldsmith, "Blind null-space learning for MIMO underlay cognitive radio with primary user interference adaptation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1722–1734, Dec. 2013.
- [16] Y. Noam and A. J. Goldsmith, "The one-bit null space learning algorithm and its convergence," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6135–6149, Dec. 2013.
- [17] B. Gopalakrishnan and N. D. Sidiropoulos, "Cognitive transmit beamforming from binary link-feedback for point to point MISO channels," in *Proc. IEEE ICASSP*, Florence, Italy, May 4–9, 2014, pp. 7293–7297.
- [18] J. Xu and R. Zhang, "Energy beamforming with one-bit feedback," in *Proc. IEEE ICASSP*, Florence, Italy, May 4–9, 2014, pp. 3513–3517.
- [19] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.
- [20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [21] S. Boyd and L. Vandenberghe, "Notes on localization and cutting-plane methods," in *EE392 Lecture Notes*. Stanford, CA, USA: Stanford Univ. Press, Sep. 2003.
- [22] W. K. Newey and D. McFadden, "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, vol. 4, 1994.
- [23] E. Tsakonas, J. Jalden, N. D. Sidiropoulos, and B. Ottersten, "Sparse conjoint analysis through maximum likelihood estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5704–5715, Nov. 2013.
- [24] Y. Noam and J. Tabrikian, "Marginal likelihood for estimation and detection theory," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 3963–3974, Aug. 2007.
- [25] H. White, *Estimation, Inference and Specification Analysis*. New York, NY, USA: Cambridge Univ. Press, 1994.



Balasubramanian Gopalakrishnan received the B.Tech. degree in electronics and telecommunication engineering from the University of Kerala, Thiruvananthapuram, India, in 2006 and the M.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 2008. He is currently working towards the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA. His research interests are in wireless communications, signal processing, and optimization.



Nicholas D. Sidiropoulos (F'09) received the Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1988 and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, USA, in 1990 and 1992, respectively. He served as an Assistant Professor with the University of Virginia, Charlottesville, VA, USA, in 1997–1999, an Associate Professor with the University of Minnesota, Minneapolis, MN, USA, in 2000–2002, a Professor with the Technical University of Crete, Chania, Greece, in 2002–2011, and a Professor with the University of Minnesota since 2011. His current research primarily focuses on signal and tensor analytics, with applications in cognitive radio, big data, and preference measurement. Prof. Sidiropoulos is a Fellow of EURASIP. He has served as an IEEE SPS Distinguished Lecturer in 2008–2009 and the Chair of the IEEE Signal Processing for Communications and Networking Technical Committee in 2007–2008. He was a recipient of the NSF/CAREER Award in 1998, the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011, the IEEE SPS Meritorious Service Award in 2010, and the Distinguished Alumni Award from the Department of Electrical and Computer Engineering, University of Maryland.