

# An Upper Bound on the Capacity of Vector Dirty Paper with Unknown Spin and Stretch

David T.H. Kao  
Cornell University – Ithaca, NY  
davidkao@cornell.edu

Ashutosh Sabharwal  
Rice University – Houston, TX  
ashu@rice.edu

**Abstract**—Dirty paper codes are a powerful tool for combating known interference. However, there is a significant difference between knowing the transmitted interference sequence and knowing the received interference sequence, especially when the channel modifying the interference is uncertain.

We present an upper bound on the capacity of a compound vector dirty paper channel where although an additive Gaussian sequence is known to the transmitter, the channel matrix between the interferer and receiver is uncertain but known to lie within a bounded set. Our bound is tighter than previous bounds in the low-SIR regime for the scalar version of the compound dirty paper channel and employs a construction that focuses on the relationship between the dimension of the message-bearing signal and the dimension of the additive state sequence. Additionally, a bound on the high-SNR behavior of the system is established.

## I. INTRODUCTION

The benefit of transmitter side information has been studied in many forms. The case of causal state information for the discrete memoryless channel was first studied in [1], and noncausal state information was subsequently considered in [2], [3]. The noncausal case was later specialized to the point-to-point AWGN channel with additive Gaussian state in [4], the so-called dirty paper channel, wherein it was shown that, surprisingly, interference perfectly known at the transmitter may be completely mitigated through a clever binning scheme.

This dirty paper coding (DPC) approach is especially applicable to certain scenarios in multiuser wireless communications: In the vector Gaussian broadcast channel, DPC enables a capacity achieving encoding scheme [5], [6] which bins each successive receiver's message against interference from messages intended for preceding receivers. In the cognitive interference channel, DPC is a useful tool for exploiting cognitive knowledge to mitigate interference [7], [8].

A known limitation of DPC is its reliance on exact knowledge of channel state. Unfortunately, in large and distributed wireless networks, providing transmitters with full channel state knowledge incurs high overhead. Hence, often in practice, channel state is known with some uncertainty. In this paper we study how such channel state uncertainty inhibits the usefulness of transmitter side information. Specifically, we study a compound vector dirty paper channel, where side information about an additive vector Gaussian interference sequence is provided to the transmitter. Channel uncertainty is

modeled by a set of possible channel matrices that transform the interference sequence prior to reception. The set studied contains all matrices with singular values less than a parameter  $a_{\max}$ , representing a known maximum amplification. The transmitter, unaware of the exact channel state, must reliably convey a message to the receiver.

Our compound formulation captures the subtle but important distinction between noncausal knowledge of transmitted interference and noncausal knowledge of received interference. With our model, we may better understand, for example, cognitive interference channels where the interference channel gains are unmeasured [9], as well as MIMO broadcast where elements of the channel matrix are unknown [10].

Compound versions of transmitter side information channels have been examined previously, with the most general formulation being [11]. In the work of [12], a more precise model with a finite number of compound states was studied and an approach termed “carbon copying” was defined. Further extensions of [12] to specific Gaussian channels may be found in [9] where phase was uncertain, and [13] where the message-bearing and interfering signals scaled proportionally.

The main result of this paper is an upper bound on capacity of dirty paper channels when the interference signal may have both dimension greater than one and the potential to undergo a wide range of amplification. For scalar dirty paper, our bound is tighter than known bounds when maximum potential amplification of interference is high, i.e., the low signal-to-interference ratio (SIR) regime. For vector dirty paper in the low-SIR regime, we find that uncertainty incurs an approximate prelog loss in capacity. When the unknown amplification is unbounded, the loss is exact and signifies a prelog capacity loss *at all finite signal to noise ratios* (SNR). A degrees of freedom (DOF) upper bound also results, thus providing insight into the high-SNR behavior of the system.

Additionally, to our knowledge, this work represents the first treatment of compound channels for *vector* dirty paper, and the focus on vector channels reveals a relationship between the dimension of the interference, the dimension of the message bearing signal, and the resulting upper bound on DOF.

The paper is structured as follows. After defining our model in Section II, we state and prove our upper bound in its most general form in Section III. Section IV presents the bound applied to the more concrete cases of MISO and SIMO channels, and includes a comparison to bounds previously

This work has been supported in part by NSF CNS award 1012921, and was completed while David T.H. Kao was at Rice University.

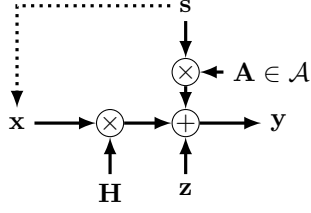


Fig. 1. Channel Model. An interference sequence is known to the transmitter, however the linear transformation of the interference prior to reception is known only to lie within some set.

applied to the scalar compound dirty paper channel. We comment on high-SNR behavior of the system in Section V, and summarize in Section VI.

## II. PROBLEM STATEMENT

We consider the channel depicted by Figure 1 with input-output relationship characterized by

$$\mathbf{y}[t] = \mathbf{H}\mathbf{x}[t] + \mathbf{A}\mathbf{s}[t] + \mathbf{z}[t], \quad (1)$$

where  $\mathbf{x}[t]$  and  $\mathbf{y}[t]$  represent channel input and output respectively of a vector channel at time index  $t$ , and  $\mathbf{s}[t]$  and  $\mathbf{z}[t]$  are zero-mean Gaussian random vectors i.i.d. across time with covariance matrices  $\mathbf{Q}_s$ , assumed to be full rank, and  $\mathbf{Q}_z = \mathbf{I}$  respectively. The dimension of  $\mathbf{x}$ ,  $\mathbf{s}$ ,  $\mathbf{z}$ , and  $\mathbf{y}$  are  $M_t$ ,  $M_s$ ,  $M_r$ , and  $M_r$  respectively. An  $M_t \times M_r$  matrix  $\mathbf{H}$  and an  $M_s \times M_r$  matrix  $\mathbf{A}$  are both assumed to be quasistatic in the sense that for any length- $n$  codeword they are constant. We use the exponent  $n$ , e.g.,  $\mathbf{x}^n \triangleq (\mathbf{x}[1], \mathbf{x}[2], \dots, \mathbf{x}[n])$ , to denote  $n$  uses of the channel. On the input covariance matrix  $\mathbf{Q}_x \triangleq \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ , we impose the average power constraint  $\text{tr}(\mathbf{Q}_x) \leq P$ .

The transmitter is given the additive vector state (interference) sequence  $\mathbf{s}^n$  noncausally, but knows only that  $\mathbf{A}$  lies within an uncertainty set,  $\mathbf{A} \in \mathcal{A} \subseteq \mathbb{R}^{M_s \times M_r}$  ( $\mathbb{C}^{M_s \times M_r}$  for the complex channel). The uncertainty set  $\mathcal{A}$  is defined as the set of all matrices with largest singular value bounded above by a known maximum amplification parameter  $a_{\max} \in [0, \infty]$ . Notice that  $a_{\max} = \infty$  implies  $\mathcal{A} = \mathbb{R}^{M_s \times M_r}$ . Furthermore, the set  $\mathcal{A}$  as defined is symmetric ( $\mathbf{A} \in \mathcal{A}$  implies  $-\mathbf{A} \in \mathcal{A}$ ), convex, and compact. When  $|\mathcal{A}| = 1$ , we have exactly the canonical vector dirty paper channel.

Our analyses apply to both real- and complex-valued channels, and we note differences in assuming one or the other only as needed. For notation, we use boldfaced lowercase to represent vectors, boldfaced uppercase for matrices,  $\dagger$  to denote Hermitian transpose, and all logarithms are base-2.

## III. UPPER BOUND ON CAPACITY

**Theorem 1** (Upper Bound). *Define  $M_0 \triangleq \text{rank}(\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger)$  and a matrix  $\mathbf{U}$  that projects the received signal onto the subspace spanned by  $\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger$ . The capacity,  $C$ , is bounded above by  $\bar{C}$  given in (2), where  $\kappa = \frac{1}{2}$  for real channels and  $\kappa = 1$  for complex channels, and supremum and infimum are*

*subject to the constraints*

$$\begin{aligned} \mathbf{Q}_x &\succeq 0 \\ \text{tr}(\mathbf{Q}_x) &\leq P \\ \mathbf{A}_i &\in \mathcal{A} \quad \forall i \in \left\{1, \dots, \left\lceil \frac{M_s}{M_0} \right\rceil\right\} \\ \mathbf{A}_i \mathbf{Q}_s \mathbf{A}_j^\dagger &= 0 \quad \forall i \neq j. \end{aligned}$$

Before presenting the proof, we point out that the received signal dimension is a *chosen* integer, which affects the denominator of (2). Consequently, the objective function of (2) may be discontinuous with respect to  $\mathbf{Q}_x$ , implying that convex methods may not suffice to solve (2) and that the maximin formulation is not necessarily interchangeable with a minimax formulation [14]. Discontinuities also necessitate the use of supremum and infimum operations.

*Proof:* The proof begins in a manner similar to [12] and [9], however differs in emphasis of potential state gain values, and is extended to vector state through an inductive argument. We assume real-valued channels for the presentation below.

We first fix the choice of  $\mathbf{Q}_x$  and notice that at most  $\left\lceil \frac{M_s}{M_0} \right\rceil$  rank- $M_0$  matrices,  $\{\mathbf{A}_i\}_{i=1, \dots, \left\lceil \frac{M_s}{M_0} \right\rceil}$ , may be chosen such that  $\mathbf{A}_i \mathbf{Q}_s \mathbf{A}_j^\dagger = 0$ . If  $M_s$  is not evenly divisible by  $M_0$ , to collection  $\{\mathbf{A}_i\}_{i=1, \dots, \left\lceil \frac{M_s}{M_0} \right\rceil}$  we add a final matrix  $\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}$ , which incorporates the remaining independent dimensions of  $\mathbf{s}$ . In the following, we denote the message as  $W$ , the channel output given state transformation matrix  $\mathbf{A}$  as  $\mathbf{y}_\mathbf{A}$ , and the projected channel output as  $\mathbf{v}_\mathbf{A} \triangleq \mathbf{U}\mathbf{y}_\mathbf{A}$ . Noting that  $\mathbf{A}_i \in \mathcal{A}$  implies  $-\mathbf{A}_i \in \mathcal{A}$ , we begin from Fano's inequality:

$$\begin{aligned} nr &\leq \min_{\mathbf{A} \in \mathcal{A}} I(W; \mathbf{y}_\mathbf{A}^n) \\ &\stackrel{(a)}{\leq} \min_{\mathbf{A} \in \mathcal{A}'} I(W; \mathbf{v}_\mathbf{A}^n) \\ &\stackrel{(b)}{\leq} \frac{2I(W; \mathbf{v}_0^n) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} I(W; \mathbf{v}_{-\mathbf{A}_i}^n) + I(W; \mathbf{v}_{\mathbf{A}_i}^n)}{2 \left\lceil \frac{M_s}{M_0} \right\rceil + 2} \\ &= \frac{2h(\mathbf{v}_0^n) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} h(\mathbf{v}_{-\mathbf{A}_i}^n) + h(\mathbf{v}_{\mathbf{A}_i}^n)}{2 \left\lceil \frac{M_s}{M_0} \right\rceil + 2} \\ &= \frac{2h(\mathbf{v}_0^n|W) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} h(\mathbf{v}_{-\mathbf{A}_i}^n|W) + h(\mathbf{v}_{\mathbf{A}_i}^n|W)}{2 \left\lceil \frac{M_s}{M_0} \right\rceil + 2}, \quad (4) \end{aligned}$$

where in step (a) we chose a reduced uncertainty set

$$\mathcal{A}' = \left\{ -\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}, \dots, -\mathbf{A}_1, \mathbf{0}, \mathbf{A}_1, \dots, \mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil} \right\},$$

and project onto the subspace containing the message, and in (b) we note that the arithmetic mean is greater than the minimum of a set.

$$\overline{C} \triangleq \sup_{\mathbf{Q}_x} \inf_{\{\mathbf{A}_i\}} \left[ \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil - 1} \log \frac{\det(\mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger + \mathbf{I} + \mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger)}{\det(\mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger)} + \log \det(\mathbf{I} + \mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger) + g\left(\mathbf{Q}_x, \mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\right) \right] \quad (2)$$

$$g(\mathbf{Q}_x, \mathbf{A}) = \begin{cases} \log \frac{\det(\mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger + \mathbf{I} + \mathbf{U}\mathbf{A}\mathbf{Q}_s\mathbf{A}^\dagger\mathbf{U}^\dagger)}{\det(\mathbf{U}\mathbf{A}\mathbf{Q}_s\mathbf{A}^\dagger\mathbf{U}^\dagger)} & \text{if } \left\lceil \frac{M_s}{M_0} \right\rceil = \frac{M_s}{M_0} \\ \log \frac{\det(\mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger + \mathbf{I} + \mathbf{U}\mathbf{A}\mathbf{Q}_s\mathbf{A}^\dagger\mathbf{U}^\dagger)}{\det(\mathbf{U}\mathbf{A}\mathbf{Q}_s\mathbf{A}^\dagger\mathbf{U}^\dagger + \frac{1}{2}\mathbf{I})} + 2M_0 & \text{else} \end{cases} \quad (3)$$

First, we bound the unconditioned entropy terms of antipodal state channel matrices:

$$\begin{aligned} & h(\mathbf{v}_{-\mathbf{A}_i}^n) + h(\mathbf{v}_{\mathbf{A}_i}^n) \\ & \stackrel{(a)}{\leq} \sum_{t=1}^n \frac{1}{2} \log \det(\mathbf{I} + \mathbf{U}\mathbf{H}\mathbf{Q}_{x[t]}\mathbf{H}^\dagger\mathbf{U}^\dagger + \Psi_t + \mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger) \\ & \quad + \frac{1}{2} \log \det(\mathbf{I} + \mathbf{U}\mathbf{H}\mathbf{Q}_{x[t]}\mathbf{H}^\dagger\mathbf{U}^\dagger - \Psi_t + \mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger) \\ & \quad + nM_0 \log(2\pi e) \\ & \stackrel{(b)}{\leq} n \log \det(\mathbf{I} + \mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger + \mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger) \\ & \quad + nM_0 \log(2\pi e), \end{aligned} \quad (5)$$

where (a) uses maximum entropy principles and expansion of covariance terms, (b) uses the concavity of log-determinant, and  $Q_{x[t]}$  and  $\Psi_t \triangleq \mathbb{E}[\mathbf{U}\mathbf{H}\mathbf{x}[t]\mathbf{s}^\dagger\mathbf{A}_i^\dagger\mathbf{U}^\dagger + \mathbf{U}\mathbf{A}_i\mathbf{s}\mathbf{x}[t]^\dagger\mathbf{H}^\dagger\mathbf{U}^\dagger]$  denote input covariance and cross correlation of  $t$ -th channel use respectively. Additionally, we note

$$h(\mathbf{v}_0^n) \leq \frac{n}{2} \log \det(\mathbf{I} + \mathbf{U}\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger\mathbf{U}^\dagger) + \frac{nM_0}{2} \log(2\pi e). \quad (6)$$

We lower bound terms conditioned on the message  $W$ :

$$\begin{aligned} & h(\mathbf{v}_0^n|W) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} h(\mathbf{v}_{\mathbf{A}_i}^n|W) \\ & \geq h(\mathbf{v}_0^n, \mathbf{v}_{\mathbf{A}_1}^n, \dots, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W) \\ & \stackrel{(a)}{=} h\left(\frac{\mathbf{v}_{\mathbf{A}_1}^n + \mathbf{v}_0^n}{\sqrt{2}}, \frac{\mathbf{v}_{\mathbf{A}_1}^n - \mathbf{v}_0^n}{\sqrt{2}}, \mathbf{v}_{\mathbf{A}_2}^n, \dots, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W\right) \\ & = h\left(\frac{\mathbf{v}_{\mathbf{A}_1}^n + \mathbf{v}_0^n}{\sqrt{2}}, \mathbf{v}_{\mathbf{A}_2}^n, \dots, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n \middle| \frac{\mathbf{v}_{\mathbf{A}_1}^n - \mathbf{v}_0^n}{\sqrt{2}}, W\right) \\ & \quad + h\left(\frac{\mathbf{v}_{\mathbf{A}_1}^n - \mathbf{v}_0^n}{\sqrt{2}}|W\right) \\ & \stackrel{(b)}{=} h\left(\frac{\mathbf{v}_{\mathbf{A}_1}^n + \mathbf{v}_0^n}{\sqrt{2}}, \mathbf{v}_{\mathbf{A}_2}^n, \dots, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n \middle| \frac{\mathbf{U}\mathbf{A}_1\mathbf{s}^n}{\sqrt{2}}, W\right) \\ & \quad + h\left(\frac{\mathbf{U}\mathbf{A}_1\mathbf{s}^n}{\sqrt{2}}\right) \\ & \stackrel{(c)}{\geq} h(\mathbf{v}_0^n, \mathbf{v}_{\mathbf{A}_2}^n, \dots, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W) \\ & \quad + \frac{n}{2} \log \det(\mathbf{U}\mathbf{A}_1\mathbf{Q}_s\mathbf{A}_1^\dagger\mathbf{U}^\dagger) + \frac{nM_0}{2} \log(2\pi e), \end{aligned} \quad (7)$$

where (a) results from a basis transformation, (b) results from perfectly correlated noise terms between the two channel outputs and independence between message and state, and (c) results from factorization of matched scaling constants. The analysis for (7) is repeated inductively to arrive at

$$\begin{aligned} & h(\mathbf{v}_0^n|W) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} h(\mathbf{v}_{\mathbf{A}_i}^n|W) \\ & \geq h(\mathbf{v}_0^n, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W) + \frac{nM_0 \left(\left\lceil \frac{M_s}{M_0} \right\rceil - 1\right)}{2} \log(2\pi e) \\ & \quad + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil - 1} \frac{n}{2} \log \det(\mathbf{U}\mathbf{A}_i\mathbf{Q}_s\mathbf{A}_i^\dagger\mathbf{U}^\dagger). \end{aligned}$$

If  $M_s$  is evenly divisible by  $M_0$  the same induction may be applied to decouple the final two potential channel outputs:

$$\begin{aligned} & h(\mathbf{v}_0^n, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W) \\ & \geq \frac{nM_0}{2} \log(2\pi e) + \frac{n}{2} \log \det\left(\mathbf{U}\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{Q}_s\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}^\dagger\mathbf{U}^\dagger\right). \end{aligned} \quad (8)$$

If  $M_s$  is not evenly divisible by  $M_0$ , the matrix  $\mathbf{U}\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{Q}_s\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}^\dagger\mathbf{U}^\dagger$  is rank deficient, and thus we only partially correlate the two noise terms  $\mathbf{z}_{\left\lceil \frac{M_s}{M_0} \right\rceil}$  and  $\mathbf{z}_0^n$ :

$$\begin{aligned} & h(\mathbf{v}_0^n, \mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n|W) \\ & = h\left(\frac{\mathbf{v}_{\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}}^n + \mathbf{v}_0^n}{\sqrt{2}} \middle| \frac{\mathbf{U}}{\sqrt{2}} \left(\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{s}^n + \mathbf{z}_{\left\lceil \frac{M_s}{M_0} \right\rceil} - \mathbf{z}_0^n\right), W\right) \\ & \quad + h\left(\frac{\mathbf{U}}{\sqrt{2}} \left(\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{s}^n + \mathbf{z}_{\left\lceil \frac{M_s}{M_0} \right\rceil} - \mathbf{z}_0^n\right) \middle| W\right) \\ & \geq h\left(\frac{\mathbf{U}}{\sqrt{2}} \left(\mathbf{z}_{\left\lceil \frac{M_s}{M_0} \right\rceil} + \mathbf{z}_0^n\right)\right) \\ & \quad + h\left(\frac{\mathbf{U}}{\sqrt{2}} \left(\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{s}^n + \mathbf{z}_{\left\lceil \frac{M_s}{M_0} \right\rceil} - \mathbf{z}_0^n\right) \middle| W\right) \\ & \geq \frac{nM_0}{2} \log(\pi e) + \frac{n}{2} \log \det\left(\mathbf{U}\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}\mathbf{Q}_s\mathbf{A}_{\left\lceil \frac{M_s}{M_0} \right\rceil}^\dagger\mathbf{U}^\dagger + \frac{1}{2}\mathbf{I}\right). \end{aligned} \quad (9)$$

An identical analysis may be performed for  $h(\mathbf{v}_0^n|W) + \sum_{i=1}^{\left\lceil \frac{M_s}{M_0} \right\rceil} h(\mathbf{v}_{-\mathbf{A}_i}^n|W)$ , and by substituting (5), (6), (8), and

(9) for both positive and negative  $\mathbf{A}_i$  into (4) and allowing minimization over collections of  $\{\mathbf{A}_i\}$ , we arrive at (2). ■

**Remark 1:** Although the optimization in (2) is potentially difficult to compute, a simpler bound can be arrived at which chooses as each  $\mathbf{A}_i$  a matrix that aligns eigenvectors of  $\mathbf{Q}_s$  with eigenvectors of  $\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger$ .

**Remark 2:** For the case where both the channel input and interference are scalars, the bound of [9] provided evidence that, with unknown phase, correlation between the input and state provided no benefit, and comparison between a zero additive state with a high-variance additive state is a special case studied in [12]. Incorporating both of these emphases into a single analysis provides new insight into how the unknown phase and unknown amplitude jointly reduce capacity.

**Remark 3:** The primary innovation in the construction of (2) is emphasis of the effect that the dimension of interference may have on the prelog factor of capacity. In particular, if large but finite interference power is considered, i.e.,  $P \ll a_{\max}^2 < \infty$ , the terms in the sum of (2) tend towards zero signifying an approximate prelog capacity loss. The error in this approximation depends primarily on the SIR of the system. If  $a_{\max} = \infty$ , i.e. the set  $\mathcal{A} = \mathbb{R}^{M_s \times M_r}$ , then the prelog loss becomes exact and the system exhibits what resembles a degrees of freedom loss at all SNR.

**Remark 4:** The choice of correlation of noise terms in (9) is not optimized, and thus the bound is potentially loose. This optimization however depends on  $\mathbf{Q}_s$  and choice of  $\{\mathbf{A}_i\}$  which in turn depends on choice of  $\mathbf{Q}_x$ . As discussed prior to presentation of the proof, it not immediately apparent how these choices interact to tighten or loosen the bound.

#### IV. MISO & SIMO CHANNELS WITH VECTOR STATE

The question of the dimension of the received message bearing signal  $\mathbf{H}\mathbf{x}$ , and subsequent optimization of the input covariance  $\mathbf{Q}_x$  prevents a more explicit characterization of (2) in general. On the other hand, if the signal is necessarily one-dimensional (e.g., when either the transmitter or receiver in a wireless communication link has one antenna) the upper bound on capacity may be simplified considerably:

**Corollary 2.** Let  $v_{si}$  denote the eigenvalues of  $\mathbf{Q}_s$ , and  $\mathbf{h}$  the channel vector modifying  $\mathbf{x}$ . The capacity,  $C$ , is bounded above by  $\bar{C}_1$  given by

$$\bar{C}_1 = \frac{\kappa \left[ \sum_{i=1}^{M_s} \log \left( \frac{\|\mathbf{h}\|^2 P + 1 + a_{\max}^2 v_{si}}{a_{\max}^2 v_{si}} \right) + \log(1 + \|\mathbf{h}\|^2 P) \right]}{M_s + 1}, \quad (10)$$

where  $\kappa = \frac{1}{2}$  for real channels and  $\kappa = 1$  for complex channels.

*Sketch of Proof:* The full proof is omitted due to limited space, however it relies only on beamforming (transmit or receive) for the message-bearing signal and a sequence of  $\mathbf{A}_i$  matrices that project individual eigenvectors of  $\mathbf{Q}_s$  onto the subspace containing the message-bearing signal. □

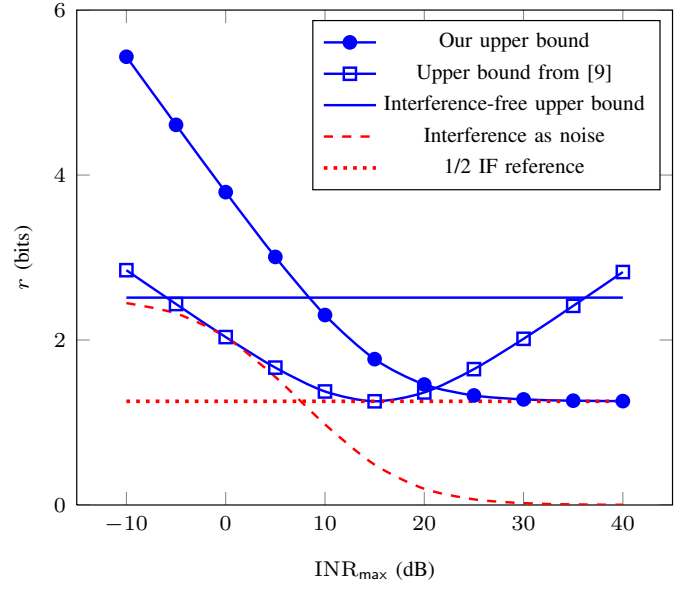


Fig. 2. Comparison of upper and lower bounds for the compound dirty paper channel with real scalar input and interference (SNR = 15 dB). The dotted trace represents half the interference-free rate and is included as reference; it does not represent an achievable scheme.

**Remark 5:** The point made in Remark 3 regarding approximation of the bound with a prelog factor is more clearly illustrated in rank-1 channels. If for example

$$\min_i v_{si} \geq \frac{1 + \|\mathbf{h}\|^2 P}{a_{\max}^2}, \quad (11)$$

then each log term in the sum of (10) is bounded above by 1, and the gap between the approximation  $\tilde{C}_1 \triangleq \frac{\kappa}{M_s+1} \log(1 + \|\mathbf{h}\|^2 P)$  and the actual bound is less than  $\frac{\kappa M_s}{M_s+1}$  bits.

**Remark 6:** One special case is the scalar channel with scalar additive state whose bound is shown in Figure 2 relative to prior work. Notice our bound is tighter at high INR (low SIR), and complements the result from [9]. The approximate prelog loss is illustrated as well: for the scalar case,  $M_s = 1$  so the prelog loss is approximately  $\frac{1}{2}$ .

#### V. HIGH-SNR BEHAVIOR

The behavior of wireless systems at high signal to noise ratios often provides insight into the spatial interaction of signals. One standard metric for high-SNR performance is the multiplexing gain or degrees of freedom (DOF) defined as

$$\text{DOF} \triangleq \lim_{\text{SNR} \rightarrow \infty} \frac{C(\text{SNR})}{\kappa \log(1 + \text{SNR})}, \quad (12)$$

where  $\text{SNR} = P$ , and  $\kappa = \frac{1}{2}$  or  $\kappa = 1$  for real or complex channels respectively. In our system, an upper bound on DOF may be deduced directly from (2). The form of the bound is contingent on how the INR, or equivalently the spectral properties of state sequence  $\mathbf{s}^n$ , scales with SNR:

**Corollary 3.** *The system has full ( $M^* \triangleq \min\{M_t, M_r\}$ ) degrees of freedom if and only if both of the following conditions hold:*

- 1) *The parameter  $a_{\max}$  is finite.*
- 2) *The interference power, or equivalently INR, grows sublinearly with respect to SNR.*

*If either condition does not hold then the degrees of freedom of the system is bounded above by*

$$\text{DOF} \leq \frac{M^* \left( \left\lceil \frac{M_s}{M^*} \right\rceil + 1 \right) - M_s}{\left\lceil \frac{M_s}{M^*} \right\rceil + 1}. \quad (13)$$

*Proof:* If both conditions hold, then the interference may be treated as noise and at high SNR the gap between the rate achieved and the interference free rate approaches a constant. On the other hand, if the first condition is false, then, as noted in Remark 2, the terms in the summation of (2) evaluate to 0. Therefore this proof focuses on the case where only the second condition is false.

If INR grows linearly with SNR, then there must exist some finite scalar  $\beta$  such that

$$\mathbf{H}\mathbf{Q}_x\mathbf{H}^\dagger \preceq \beta\mathbf{A}\mathbf{Q}_s\mathbf{A}^\dagger, \quad (14)$$

for all  $\mathbf{Q}_x$  where  $\text{tr}(\mathbf{Q}_x) \leq P$ . Consequently, the terms in the summation of (2) may be bounded by a constant which has vanishing contribution when normalized by  $\log(1+P)$  as  $P \rightarrow \infty$ . If INR grows superlinearly with respect to SNR, then a function  $\beta(P) \rightarrow 0$  as  $P \rightarrow \infty$  suffices to satisfy (14), and the terms in the summation of (2) vanish as  $P \rightarrow \infty$ .

By counting the number of dimensions of interference relative to message bearing signal, the asymptotic behavior of the remaining two terms in the numerator of (2) results in the DOF upper bound for fixed  $M_0$

$$\text{DOF} \leq \frac{M_0 \left( \left\lceil \frac{M_s}{M_0} \right\rceil + 1 \right) - M_s}{\left\lceil \frac{M_s}{M_0} \right\rceil + 1},$$

which is maximized when  $M_0 = M^*$ . ■

**Remark 7:** It is important to note cases where the two conditions posed in Corollary 3 hold in the context of common wireless network applications of DPC. With respect to the first condition, often the known sequence represents encoded messages intended for other receivers that interfere with the DPC transmission. In these cases, perhaps when the INR is high enough, i.e., the singular values of  $\mathbf{A}$  are above  $a_{\max}$ , the interference may be decoded and the nature of the system changes. Alternatively, the vector  $\mathbf{s}$  might represent multiple known sequences whose exact linear transformation at the receiver is unknown, but whose magnitude may be bounded based on a measurement of aggregate INR.

For the second condition to hold, we must assume that the nature of increased SNR is a result of increased transmission power at the transmitter rather than a decrease in thermal noise at the receiver.

**Remark 8:** Unlike the finite-SNR behavior, the DOF loss exhibited when Condition 2 is false is not confined to any specific SIR regime. Even if  $a_{\max}$  is small, if the INR scales linearly with SNR then the statement holds. Moreover, the exact covariance structure is less relevant at high-SNR than the dimension or rank of  $\mathbf{Q}_s$ .

## VI. SUMMARY

In this paper, we studied a compound channel model for vector dirty paper where the linear transformation spinning and stretching the dirty paper is unknown. We presented an upper bound on the capacity of this compound vector dirty paper channel. Our bound is tighter than previous bounds in the low-SIR regime for the case of scalar input and interference, and extends intuitions regarding prelog loss in capacity to the vector dirty paper channel. The bound offers insight into the high-SNR behavior of systems modeled by vector dirty paper, and a relationship between the dimension of the message-bearing signal, the dimension of the interference, and the degrees of freedom of the system was revealed.

## REFERENCES

- [1] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 289–293, Oct. 1958.
- [2] S. I. Gel'fand and M. S. Pinsker, "Coding for channels with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [3] C. Heegard and A. Gamal, "On the capacity of computer memory with defects," *Information Theory, IEEE Transactions on*, vol. 29, no. 5, pp. 731–739, Sep. 1983.
- [4] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, May 1983.
- [5] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *Information Theory, IEEE Transactions on*, vol. 49, no. 7, pp. 1691–1706, July 2003.
- [6] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [7] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *Information Theory, IEEE Transactions on*, vol. 52, no. 5, pp. 1813–1827, May 2006.
- [8] S. Rini, D. Tuninetti, and N. Devroye, "Inner and outer bounds for the gaussian cognitive interference channel and new capacity results," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 820–848, Feb. 2012.
- [9] P. Grover and A. Sahai, "On the need for knowledge of the phase in exploiting known primary transmissions," in *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, April 2007, pp. 462–471.
- [10] D. T. Kao and A. Sabharwal, "Node cooperation with local views in the two-user interference channel," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 1748–1752.
- [11] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side information at the transmitter," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1745–1755, April 2006.
- [12] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon copying onto dirty paper," *Information Theory, IEEE Transactions on*, vol. 53, no. 5, pp. 1814–1827, May 2007.
- [13] W. Zhang, S. Kotagiri, and J. N. Laneman, "Writing on dirty paper with resizing and its application to quasi-static fading broadcast channels," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, June 2007, pp. 381–385.
- [14] K. Fan, "On a theorem of Weyl concerning the eigenvalues of linear transformations II," *Proc. Nat. Acad. Sci., U.S.A.*, vol. 36, pp. 31–35, 1950.