A Sparseness-Preserving Virtual MIMO Channel Model

Phil Schniter
Dept. ECE, The Ohio State University
Columbus, OH 43210
schniter.1@osu.edu

Akbar Sayeed
Dept. ECE, University of Wisconsin
Madison, WI 53706
akbar@engr.wisc.edu

***Abstract*— Through the use of antenna arrays at both ends of a wireless link, a channel with significant delay and Doppler spread can be decomposed into parallel virtual channels with much smaller delay and Doppler spreads, thereby simplifying the tasks of channel tracking and ISI mitigation. While one might hope that a sparse physical environment (i.e., few, well-localized scatterers) would lead to a sparse virtual channel representation, this is not necessarily the case. Here we describe how a simple scaling of antenna inputs and outputs can yield a sparseness-preserving virtual channel model.**[1]

I. Introduction

Certain wireless communication scenarios are well characterized by a sparse physical structure, where the propagation between transmitter and receiver occurs primarily over a small number of paths. With even a few paths, however, the overall channel delay spread can be long if the path delays are disparate. Furthermore, the overall channel response can change in a somewhat complicated fashion when the path delays vary in time. The tracking and mitigation of inter-symbol interference (ISI) induced by a quickly varying multipath delay profile is, in many applications, the most expensive function performed by the receiver, and, in some applications, the limiting factor to achieving high throughput [1].

The use of antenna arrays at the transmitter and receiver can simplify the multipath tracking task, especially for sparse channels [2]. For example, consider the case where the arrays can form narrow beams. Since *per-beam* delay and Doppler spreads will be much smaller than those of the overall channel, per-beam channel tracking can be made relatively simple. If, in addition, only a few beams are active, then overall channel tracking can be made quite manageable.

In the applications we consider, it is not usually the case that the transmitter knows the active scattering angles. Thus it is convenient to consider a *uniform* transmitter beam spacing, with the number of beams matching the number of transmit antennas. The same uniform-beam-spacing can be applied at the receiving array without loss of generality, since, as we will see, the corresponding signal transformation is lossless. This uniform approach has been studied by Sayeed and is the basis for his "virtual channel" model [3]. One weakness of the virtual channel model is that a sparse *physical* scattering environment does not necessarily lead to a sparse *virtual* channel representation. In this paper, we show how a simple scaling of the antenna inputs and outputs can preserve sparseness in the virtual multiple-input multiple-output (MIMO) channel model.

[1]This work was supported by NSF CAREER Award CCR-0237037.

Notation: We use $(\cdot)^t$ to denote transpose, $(\cdot)^*$ conjugate, and $(\cdot)^H$ conjugate transpose. $\mathcal{D}(\boldsymbol{b})$ denotes the diagonal matrix created from vector $\boldsymbol{b}$, and $\boldsymbol{I}$ the identity matrix. We use $[\boldsymbol{B}]_{m,n}$ to denote the element in the m^{th} row and n^{th} column of $\boldsymbol{B}$, where row/column indices begin with zero. $\mathrm{E}\{\cdot\}$ denotes expectation, δ_k denotes the Kronecker delta with respect to $k \in \mathbb{Z}$, $\mathbb{Z}$ the set of integers, and $\mathrm{sinc}(x) := \sin(x)/x$.

II. Background

We consider a MIMO system with P antennas at the transmitter and Q antennas at the receiver. In the absence of noise, the input/output relation can be described by

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s}$$

where the matrix $\boldsymbol{H} \in \mathbb{C}^{Q\times P}$ describes the effects of multipath propagation.

We restrict ourselves to the case of uniform linear arrays operating under the "narrowband" assumption. With L propagation paths, where the l^{th} path departs at angle $\phi_{\mathsf{T},l}$ and arrives at angle $\phi_{\mathsf{R},l}$ with path gain β_l, we have [3]

$$\begin{aligned}
\boldsymbol{H} &= \sum_{l=1}^{L} \beta_l \boldsymbol{\alpha}_{\mathsf{R}}(\theta_{\mathsf{R},l})\boldsymbol{\alpha}_{\mathsf{T}}^H(\theta_{\mathsf{T},l}) \qquad (1)\\
\boldsymbol{\alpha}_{\mathsf{T}}(\theta_{\mathsf{T}}) &= \frac{1}{\sqrt{P}}\left[1 \;\; e^{-j2\pi\theta_{\mathsf{T}}} \;\; \cdots \;\; e^{-j2\pi(P-1)\theta_{\mathsf{T}}}\right]^t\\
\boldsymbol{\alpha}_{\mathsf{R}}(\theta_{\mathsf{R}}) &= \frac{1}{\sqrt{Q}}\left[1 \;\; e^{-j2\pi\theta_{\mathsf{R}}} \;\; \cdots \;\; e^{-j2\pi(Q-1)\theta_{\mathsf{R}}}\right]^t\\
\theta_{\mathsf{T}} &= \frac{d_{\mathsf{T}}}{\lambda}\sin(\phi_{\mathsf{T}})\\
\theta_{\mathsf{R}} &= \frac{d_{\mathsf{R}}}{\lambda}\sin(\phi_{\mathsf{R}}).
\end{aligned}$$

Here d_{T} and d_{R} denote the antenna spacings at transmitter and receiver, respectively, and λ denotes the propagation wavelength. It is typical to define the normalized antenna spacing as $\alpha = d/\lambda$. Whereas ϕ represents the *physical* angle of departure/arrival, θ can be interpreted as the *spatial* angle. For simplicity we restrict ourselves to the critical spacing $\boldsymbol{\alpha}_{\mathsf{T}} = \boldsymbol{\alpha}_{\mathsf{R}} = 0.5$, so that there exists a one-to-one mapping between $\theta \in [-0.5, 0.5)$ and $\phi \in [-\pi/2, \pi/2)$. It is typical to

write the L-path physical model in matrix form as

$$\boldsymbol{H} = \underbrace{\left[\boldsymbol{\alpha}(\theta_{\mathsf{R},1}) \ \cdots \ \boldsymbol{\alpha}(\theta_{\mathsf{R},L})\right]}_{\boldsymbol{A}_{\mathsf{R}}(\underline{\theta}_{\mathsf{R}})} \underbrace{\mathcal{D}\big([\beta_1,\ldots,\beta_L]\big)}_{\boldsymbol{H}_P} \times \underbrace{\left[\boldsymbol{\alpha}(\theta_{\mathsf{T},1}) \ \cdots \ \boldsymbol{\alpha}(\theta_{\mathsf{T},L})\right]^H}_{\boldsymbol{A}_{\mathsf{T}}^H(\underline{\theta}_{\mathsf{T}})},$$

where $\boldsymbol{H}_P$ is an $L \times L$ diagonal matrix. Note the dependence of $\boldsymbol{A}_{\mathsf{R}}(\underline{\theta}_{\mathsf{R}})$ and $\boldsymbol{A}_{\mathsf{T}}(\underline{\theta}_{\mathsf{T}})$ on the spatial angles.

The finite dimensionality of the signal space also allows the *virtual* channel model [3]

$$\boldsymbol{H} = \underbrace{\left[\boldsymbol{\alpha}(\tfrac{-\tilde{Q}}{Q}) \ \boldsymbol{\alpha}(\tfrac{-\tilde{Q}+1}{Q}) \ \cdots \ \boldsymbol{\alpha}(\tfrac{\tilde{Q}}{Q})\right]}_{\tilde{\boldsymbol{A}}_{\mathsf{R}}} \boldsymbol{H}_V \times \underbrace{\left[\boldsymbol{\alpha}(\tfrac{-\tilde{P}}{P}) \ \boldsymbol{\alpha}(\tfrac{-\tilde{P}+1}{P}) \ \cdots \ \boldsymbol{\alpha}(\tfrac{\tilde{P}}{P})\right]^H}_{\tilde{\boldsymbol{A}}_{\mathsf{T}}^H}$$

where $\tilde{P} := \frac{P-1}{2}$ and $\tilde{Q} := \frac{Q-1}{2}$ (assuming that P and Q are odd), and where $\boldsymbol{H}_V \in \mathbb{C}^{Q\times P}$ is no longer diagonal. Note that $\tilde{\boldsymbol{A}}_{\mathsf{R}}$ and $\tilde{\boldsymbol{A}}_{\mathsf{T}}$ are channel-invariant unitary DFT matrices. Using the fact that $\boldsymbol{H}_V = \tilde{\boldsymbol{A}}_{\mathsf{R}}^H \boldsymbol{H} \tilde{\boldsymbol{A}}_{\mathsf{T}}$, it is possible to show [3] that

$$[\boldsymbol{H}_V]_{\tilde{Q}+q,\tilde{P}+p} = \sum_{l=1}^{L} \beta_l f_{\mathsf{R}}(\theta_{\mathsf{R},l} - \tfrac{q}{Q}) f_{\mathsf{T}}^*(\theta_{\mathsf{T},l} - \tfrac{p}{P}) \quad (2)$$

$$\text{for } -\tilde{Q} \le q \le \tilde{Q}, -\tilde{P} \le p \le \tilde{P}$$

$$f_{\mathsf{R}}(\theta) = \frac{\sin(\pi Q\theta)}{Q\sin(\pi\theta)} e^{-j2\pi\tilde{Q}\theta} \quad (3)$$

$$f_{\mathsf{T}}(\theta) = \frac{\sin(\pi P\theta)}{P\sin(\pi\theta)} e^{-j2\pi\tilde{P}\theta} \quad (4)$$

While the virtual channel representation is attractive in many ways (see [3]), the "smoothing functions" $f_{\mathsf{R}}(\theta)$ and $f_{\mathsf{T}}(\theta)$ yield the undesirable behavior that the effects of localized scattering may be distributed over many entries in $\boldsymbol{H}_V$. In other words, *a sparse physical scattering environment does not necessarily lead to a sparse virtual channel representation*. For example, Fig. 1 plots $|f_{\mathsf{T}}(\theta)|$ versus θ, which can be interpreted as the response of virtual coefficient p to a single scatterer at departure angle $\theta_{\mathsf{T}} = \theta + p/P$. Note that the response does not decay quickly in θ. In this paper, we propose a relatively simple transformation that corrects this behavior.

III. The Shaped Virtual Channel

In (1), $\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s}$ was used to relate the antenna outputs to the antenna inputs in the case of a multipath propagation channel. Consider now scaling the p^{th} transmit antenna input by $d_{\mathsf{T},p}^*$ and scaling the q^{th} receive antenna output by $d_{\mathsf{R},q}$. Defining the *shaping coefficient* vectors $\boldsymbol{d}_{\mathsf{T}} := [d_{\mathsf{T},1},\ldots,d_{\mathsf{T},P}]^t$ and $\boldsymbol{d}_{\mathsf{R}} := [d_{\mathsf{R},1},\ldots,d_{\mathsf{R},Q}]^t$, the *shaped* MIMO output $\bar{\boldsymbol{x}}$ can be written

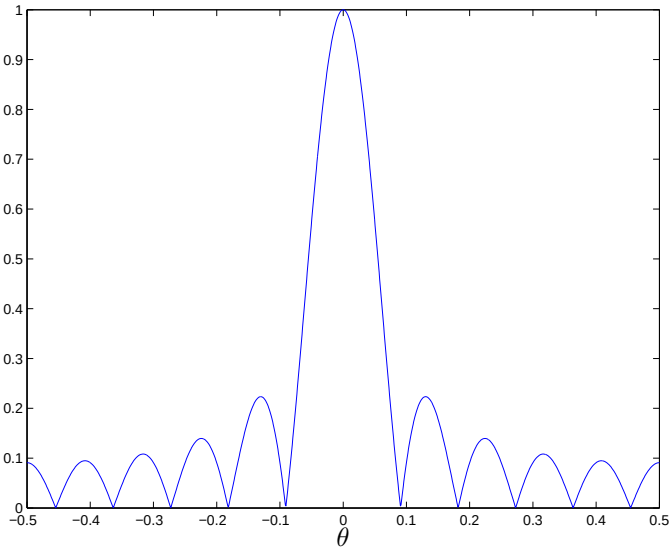

Fig. 1. Smoothing function magnitude $|f_{\mathsf{T}}(\theta)|$ for $P = 11$.

$$\begin{aligned}\bar{\boldsymbol{x}} &= \mathcal{D}(\boldsymbol{d}_{\mathsf{R}})\boldsymbol{H}\,\mathcal{D}(\boldsymbol{d}_{\mathsf{T}}^*)\boldsymbol{s} \\ &= \tilde{\boldsymbol{A}}_{\mathsf{R}} \underbrace{\tilde{\boldsymbol{A}}_{\mathsf{R}}^H \mathcal{D}(\boldsymbol{d}_{\mathsf{R}})\boldsymbol{H}\,\mathcal{D}(\boldsymbol{d}_{\mathsf{T}}^*)\tilde{\boldsymbol{A}}_{\mathsf{T}}}_{:=\bar{\boldsymbol{H}}_V} \tilde{\boldsymbol{A}}_{\mathsf{T}}^H \boldsymbol{s} \\ &= \tilde{\boldsymbol{A}}_{\mathsf{R}} \bar{\boldsymbol{H}}_V \tilde{\boldsymbol{A}}_{\mathsf{T}}^H \boldsymbol{s}\end{aligned}$$

Our aim is to design $\boldsymbol{d}_{\mathsf{T}}$ and $\boldsymbol{d}_{\mathsf{R}}$ so that the *shaped virtual channel matrix* $\bar{\boldsymbol{H}}_V$ is sparse whenever the physical scattering clusters are well localized. Loosely speaking, *the shaping coefficients act as lenses through which the channel can be brought into focus.*

It is instructive to note that we can write $\bar{\boldsymbol{H}}_V = \boldsymbol{C}_{\mathsf{R}}\boldsymbol{H}_V\boldsymbol{C}_{\mathsf{T}}^H$ for circulant (i.e., circular convolution) matrices $\boldsymbol{C}_{\mathsf{R}} := \tilde{\boldsymbol{A}}_{\mathsf{R}}^H \mathcal{D}(\boldsymbol{d}_{\mathsf{R}})\tilde{\boldsymbol{A}}_{\mathsf{R}}$ and $\boldsymbol{C}_{\mathsf{T}} := \tilde{\boldsymbol{A}}_{\mathsf{T}}^H \mathcal{D}(\boldsymbol{d}_{\mathsf{T}})\tilde{\boldsymbol{A}}_{\mathsf{T}}$. Left multiplication by $\boldsymbol{C}_{\mathsf{R}}$ can be interpreted as a filtering operation on the columns of $\boldsymbol{H}_V$, and right multiplication by $\boldsymbol{C}_{\mathsf{T}}^H$ can be interpreted as a filtering operation on the rows of $\boldsymbol{H}_V$. This is reminiscent of linear filtering for ISI-channel shortening prior to symbol detection in single carrier systems [4] and of circular filtering for ICI-channel shortening in multi-carrier systems [5].

A. Max-SIR Shaping Coefficients

To design the shaping vectors $\boldsymbol{d}_{\mathsf{T}}$ and $\boldsymbol{d}_{\mathsf{R}}$, we first consider an single scattering cluster that consists of L paths uniformly distributed in a region around the center angles $\bar{\theta}_{\mathsf{T}}$ and $\bar{\theta}_{\mathsf{R}}$. Specifically, we model the l^{th} path parameters as follows:

$$\begin{aligned}\beta_l &\sim \mathcal{N}(0,\sigma_l^2) \\ \theta_{\mathsf{R},l} &\sim \mathcal{U}\left[\bar{\theta}_{\mathsf{R}} - \tfrac{W_{\mathsf{R}}}{2}, \bar{\theta}_{\mathsf{R}} + \tfrac{W_{\mathsf{R}}}{2}\right) \\ \theta_{\mathsf{T},l} &\sim \mathcal{U}\left[\bar{\theta}_{\mathsf{T}} - \tfrac{W_{\mathsf{T}}}{2}, \bar{\theta}_{\mathsf{T}} + \tfrac{W_{\mathsf{T}}}{2}\right).\end{aligned}$$

Furthermore, $\{\beta_l\}$, $\{\theta_{\mathsf{T},l}\}$, and $\{\theta_{\mathsf{R},l}\}$ are assumed to be statistically independent. In the distributions above, W_{T} and

W_{R} denote the cluster widths. The virtual bin indices nearest the cluster center will be denoted by (m,n), and the distance between cluster and bin centers will be denoted by Δ_{T} and Δ_R. Together we have

$$\bar{\theta}_{\mathsf{R}} = \frac{m+\Delta_{\mathsf{R}}}{Q}, \quad \Delta_{\mathsf{R}} \in [-\tfrac{1}{2},\tfrac{1}{2}),\ m \in \{-\tilde{Q},\ldots,\tilde{Q}\},$$
$$\bar{\theta}_{\mathsf{T}} = \frac{n+\Delta_{\mathsf{T}}}{P}, \quad \Delta_{\mathsf{T}} \in [-\tfrac{1}{2},\tfrac{1}{2}),\ n \in \{-\tilde{P},\ldots,\tilde{P}\}.$$

We quantify the "sparseness" of $\bar{\boldsymbol{H}}_V$ in terms of signal-to-interference ratio (SIR) defined by the ratio of energy captured by the $(m,n)^{th}$ virtual bin to energy in other virtual bins. To allow for the possibility that a scattering cluster cannot be well-squeezed into a single bin, we allow a *don't care region* of $\pm D_{\mathsf{T}}$ adjacent transmit bins and $\pm D_{\mathsf{R}}$ adjacent receive bins. Thus we define $\mathrm{SIR}_{D_{\mathsf{T}},D_{\mathsf{R}}} := \mathcal{E}_{\mathsf{s}}/\mathcal{E}_{\mathsf{i}}$, where

$$\mathcal{E}_{\mathsf{s}} := \mathrm{E}\left|\left[\bar{\boldsymbol{H}}_V\right]_{\tilde{Q}+m,\tilde{P}+n}\right|^2$$
$$\mathcal{E}_{\mathsf{i}} := \underbrace{\sum_{q=-\tilde{Q}}^{\tilde{Q}}\sum_{p=-\tilde{P}}^{\tilde{P}} \mathrm{E}\left|\left[\bar{\boldsymbol{H}}_V\right]_{\tilde{Q}+q,\tilde{P}+p}\right|^2}_{\mathcal{E}_{\mathsf{t}}} - \underbrace{\sum_{q=m-D_{\mathsf{R}}}^{m+D_{\mathsf{R}}}\sum_{p=n-D_{\mathsf{T}}}^{n+D_{\mathsf{T}}} \mathrm{E}\left|\left[\bar{\boldsymbol{H}}_V\right]_{\tilde{Q}+q,\tilde{P}+p}\right|^2}_{\mathcal{E}_{\mathsf{d}}}.$$

In the previous equation, $\mathcal{E}_{\mathsf{t}}$ denotes the total energy and $\mathcal{E}_{\mathsf{d}}$ the non-interference energy.

In the Appendix we show that

$$\left[\bar{\boldsymbol{H}}_V\right]_{\tilde{Q}+q,\tilde{P}+p} = \sum_{l=1}^{L}\beta_l\,\boldsymbol{\alpha}_{\mathsf{R}}^H(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\boldsymbol{d}_{\mathsf{R}} \times \boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{\alpha}_{\mathsf{T}}(\frac{p}{P}-\theta_{\mathsf{T},l}) \quad (5)$$

$$\mathrm{E}\left|\left[\bar{\boldsymbol{H}}_V\right]_{\tilde{Q}+q,\tilde{P}+p}\right|^2 = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}} \times \boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}} \quad (6)$$

$$\left[\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})\right]_{c,b} := \tfrac{1}{P}e^{-j\frac{2\pi}{P}(p-n-\Delta_{\mathsf{T}})(c-b)} \times \mathrm{sinc}\big(\pi W_{\mathsf{T}}(c-b)\big) \quad (7)$$

$$\left[\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})\right]_{c,b} := \tfrac{1}{Q}e^{-j\frac{2\pi}{Q}(q-m-\Delta_{\mathsf{R}})(c-b)} \times \mathrm{sinc}\big(\pi W_{\mathsf{R}}(c-b)\big) \quad (8)$$

for $\sigma^2 := \sum_{l=1}^{L}\sigma_l^2$. Equation (6) implies that $\mathcal{E}_{\mathsf{s}} = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T}}(n,n,\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R}}(m,m,\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}}$. Thus, from (7) and (8) we see that $\mathcal{E}_{\mathsf{s}}$ is invariant to (m,n). To make this invariance explicit, we write

$$\mathcal{E}_{\mathsf{s}} = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T,s}}(\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R,s}}(\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}} \quad (9)$$

$$\left[\boldsymbol{B}_{\mathsf{T,s}}(\Delta_{\mathsf{T}})\right]_{c,b} := \tfrac{1}{P}e^{j\frac{2\pi}{P}\Delta_{\mathsf{T}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{P}) \quad (10)$$

$$\left[\boldsymbol{B}_{\mathsf{R,s}}(\Delta_{\mathsf{R}})\right]_{c,b} := \tfrac{1}{Q}e^{j\frac{2\pi}{Q}\Delta_{\mathsf{R}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{Q}). \quad (11)$$

Equation (6) also implies that

$$\mathcal{E}_{\mathsf{d}} = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\left(\sum_{p=n-D_{\mathsf{T}}}^{n+D_{\mathsf{T}}}\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})\right)\boldsymbol{d}_{\mathsf{T}} \times \boldsymbol{d}_{\mathsf{R}}^H\left(\sum_{q=m-D_{\mathsf{R}}}^{m+D_{\mathsf{R}}}\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})\right)\boldsymbol{d}_{\mathsf{R}} = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T,n}}(\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R,n}}(\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}} \quad (12)$$

where Euler's identity yields

$$\left[\boldsymbol{B}_{\mathsf{T,n}}(\Delta_T)\right]_{c,b} = \tfrac{1}{P}e^{j\frac{2\pi}{P}\Delta_{\mathsf{T}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{P}) \times\left(1+2\sum_{p=1}^{D_{\mathsf{T}}}\cos(2\pi\frac{p}{P}(c-b))\right) \quad (13)$$

$$\left[\boldsymbol{B}_{\mathsf{R,n}}(\Delta_R)\right]_{c,b} = \tfrac{1}{Q}e^{j\frac{2\pi}{Q}\Delta_{\mathsf{T}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{Q}) \times\left(1+2\sum_{q=1}^{D_{\mathsf{R}}}\cos(2\pi\frac{q}{Q}(c-b))\right). \quad (14)$$

From (6) again we have

$$\mathcal{E}_{\mathsf{t}} = \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\left(\sum_{p=-\tilde{P}}^{\tilde{P}}\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_T)\right)\boldsymbol{d}_{\mathsf{T}} \times \boldsymbol{d}_{\mathsf{R}}^H\left(\sum_{q=-\tilde{Q}}^{\tilde{Q}}\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_R)\right)\boldsymbol{d}_{\mathsf{R}}.$$

Exploiting the P-periodicity of $\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})$ and the Q-periodicity of $\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})$,

$$\sum_{p=-\tilde{P}}^{\tilde{P}}\left[\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})\right]_{c,b} = e^{j\frac{2\pi}{P}\Delta_{\mathsf{T}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{P})\frac{1}{P}\sum_{p'=0}^{P-1}e^{-j2\pi\frac{p'}{P}(c-b)} = \delta_{c-b}$$

$$\sum_{q=-\tilde{Q}}^{\tilde{Q}}\left[\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})\right]_{c,b} = e^{j\frac{2\pi}{Q}\Delta_{\mathsf{R}}(c-b)}\,\mathrm{sinc}(\tfrac{\pi(c-b)}{Q})\frac{1}{Q}\sum_{q'=0}^{Q-1}e^{-j2\pi\frac{q'}{Q}(c-b)} = \delta_{c-b}$$

so that

$$\mathcal{E}_{\mathsf{t}} = \sigma^2\|\boldsymbol{d}_{\mathsf{T}}\|^2\|\boldsymbol{d}_{\mathsf{R}}\|^2. \quad (15)$$

Combining (9), (12), and (15), we have

$$\mathrm{SIR}_{D_{\mathsf{T}},D_{\mathsf{R}}}(\boldsymbol{d}_{\mathsf{T}},\boldsymbol{d}_{\mathsf{R}}) = \frac{\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T,s}}(\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R,s}}(\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}}}{\|\boldsymbol{d}_{\mathsf{T}}\|^2\|\boldsymbol{d}_{\mathsf{R}}\|^2 - \boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T,n}}(\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R,n}}(\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}}} \quad (16)$$

Note that $\mathrm{SIR}_{D_{\mathsf{T}},D_{\mathsf{R}}}$ is invariant to $\|\boldsymbol{d}_{\mathsf{T}}\|$, $\|\boldsymbol{d}_{\mathsf{R}}\|$, σ^2, and L.

While the SIR expression (16) is invariant to (n,m), the virtual bin location nearest to the cluster center, it depends on

the *offset* between the cluster center and this nearest bin (i.e., Δ_T and Δ_R). But since

$$\begin{aligned}\boldsymbol{B}_{\mathsf{R,s}}(\Delta_\mathsf{R}) &= \boldsymbol{M}_\mathsf{R}(\Delta_\mathsf{R})^H \boldsymbol{B}_{\mathsf{R,s}}(0)\boldsymbol{M}_\mathsf{R}(\Delta_\mathsf{R})\\ \boldsymbol{B}_{\mathsf{T,s}}(\Delta_\mathsf{T}) &= \boldsymbol{M}_\mathsf{T}(\Delta_\mathsf{T})^H \boldsymbol{B}_{\mathsf{T,s}}(0)\boldsymbol{M}_\mathsf{T}(\Delta_\mathsf{T})\\ \boldsymbol{B}_{\mathsf{R,n}}(\Delta_\mathsf{R}) &= \boldsymbol{M}_\mathsf{R}(\Delta_\mathsf{R})^H \boldsymbol{B}_{\mathsf{R,n}}(0)\boldsymbol{M}_\mathsf{R}(\Delta_\mathsf{R})\\ \boldsymbol{B}_{\mathsf{T,n}}(\Delta_\mathsf{T}) &= \boldsymbol{M}_\mathsf{T}(\Delta_\mathsf{T})^H \boldsymbol{B}_{\mathsf{T,n}}(0)\boldsymbol{M}_\mathsf{T}(\Delta_\mathsf{T})\end{aligned}$$

for diagonal "modulation" matrices

$$\begin{aligned}\boldsymbol{M}_\mathsf{T}(\Delta_\mathsf{T}) &:= \mathcal{D}\big([e^{j\frac{2\pi}{P}\Delta_T\cdot 0}, e^{j\frac{2\pi}{P}\Delta_T\cdot 1}, \dots e^{j\frac{2\pi}{P}\Delta_T(P-1)}]\big)\\ \boldsymbol{M}_\mathsf{R}(\Delta_\mathsf{R}) &:= \mathcal{D}\big([e^{j\frac{2\pi}{Q}\Delta_R\cdot 0}, e^{j\frac{2\pi}{Q}\Delta_R\cdot 1}, \dots e^{j\frac{2\pi}{Q}\Delta_R(Q-1)}]\big),\end{aligned}$$

the SIR-maximizing $\boldsymbol{d}_\mathsf{T}$ and $\boldsymbol{d}_\mathsf{R}$ for the case of nonzero Δ_T and Δ_R are modulated versions of the SIR-maximizing $\boldsymbol{d}_\mathsf{T}$ and $\boldsymbol{d}_\mathsf{R}$ for $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0$. The effect of the modulation is to spatially *de-rotate* the effective cluster center to the middle of the nearest virtual bin. In practice, it is unlikely that Δ_T and Δ_R will be known, especially at the transmitter. Furthermore, when multiple clusters are present, it will not be possible to individually de-rotate each cluster using a single set of shaping coefficients. In either of these cases, it would be most appropriate to design the shaping coefficients based on the assumption that $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0$.

A closed form solution to the joint maximization of (16) appears difficult to obtain, so we propose to alternate

$$\boldsymbol{d}_\mathsf{T}^{(i)} = \arg\max_{\boldsymbol{d}_\mathsf{T}} \frac{\boldsymbol{d}_\mathsf{T}^H \boldsymbol{B}_{\mathsf{T,s}} \boldsymbol{d}_\mathsf{T}}{\boldsymbol{d}_\mathsf{T}^H \left(\boldsymbol{I} - \frac{\boldsymbol{d}_\mathsf{R}^{(i-1)H}\boldsymbol{B}_{\mathsf{R,n}}\boldsymbol{d}_\mathsf{R}^{(i-1)}}{\|\boldsymbol{d}_\mathsf{R}^{(i-1)}\|^2}\boldsymbol{B}_{\mathsf{T,n}}\right)\boldsymbol{d}_\mathsf{T}} \quad (17)$$

$$\boldsymbol{d}_\mathsf{R}^{(i)} = \arg\max_{\boldsymbol{d}_\mathsf{R}} \frac{\boldsymbol{d}_\mathsf{R}^H \boldsymbol{B}_{\mathsf{R,s}} \boldsymbol{d}_\mathsf{R}}{\boldsymbol{d}_\mathsf{R}^H \left(\boldsymbol{I} - \frac{\boldsymbol{d}_\mathsf{T}^{(i)H}\boldsymbol{B}_{\mathsf{T,n}}\boldsymbol{d}_\mathsf{T}^{(i)}}{\|\boldsymbol{d}_\mathsf{T}^{(i)}\|^2}\boldsymbol{B}_{\mathsf{R,n}}\right)\boldsymbol{d}_\mathsf{R}} \quad (18)$$

for $i = 1, 2, 3, \dots$ with $\boldsymbol{d}_\mathsf{R}^{(0)} = \mathbf{1}^T$. The optimizations (17) and (18) follow from the solution of a generalized eigenvalue problem. In our numerical experiments, (17)-(18) converges within two iterations. Without loss of generality, we scale the shaping coefficients so that $\|\boldsymbol{d}_\mathsf{T}\| = \|\boldsymbol{d}_\mathsf{R}\| = 1$.

B. The Shaped Smoothing Functions

Comparing (2) and (5), we see that the unshaped smoothing functions $f_\mathsf{T}(\theta)$ and $f_\mathsf{R}(\theta)$ are generalized by the shaped smoothing functions $\bar{f}_\mathsf{T}(\theta, \boldsymbol{d}_\mathsf{T})$ and $\bar{f}_\mathsf{R}(\theta, \boldsymbol{d}_\mathsf{R})$:

$$\bar{f}_\mathsf{T}(\theta, \boldsymbol{d}_\mathsf{T}) := \boldsymbol{\alpha}_\mathsf{T}^H(\tfrac{p}{P} - \theta)\boldsymbol{d}_\mathsf{T} \quad (19)$$

$$\bar{f}_\mathsf{R}(\theta, \boldsymbol{d}_\mathsf{R}) := \boldsymbol{\alpha}_\mathsf{R}^H(\tfrac{q}{Q} - \theta)\boldsymbol{d}_\mathsf{R}. \quad (20)$$

The shaped smoothing functions are plotted in Fig. 2 (linear scale) and Fig. 3 (log scale) for the case of SIR-maximizing $\boldsymbol{d}_\mathsf{T}$ and $\boldsymbol{d}_\mathsf{R}$, $P = Q = 11$, $W_\mathsf{T} = W_\mathsf{R} = \frac{1}{P}$, $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0$, and various choices of $D = D_\mathsf{T} = D_\mathsf{R}$. Also plotted is the unshaped smoothing function. Note that larger values of D trade sidelobe height for mainlobe width. It is interesting to note that the case $D = 0$ is relatively close to the unshaped case. Figure 4 shows the corresponding shaping coefficients $\boldsymbol{d}_\mathsf{T}$. Note that the shaping vector is even-symmetric and real-valued, and that, in this case, $\boldsymbol{d}_\mathsf{T} = \boldsymbol{d}_\mathsf{R}$ by symmetry.

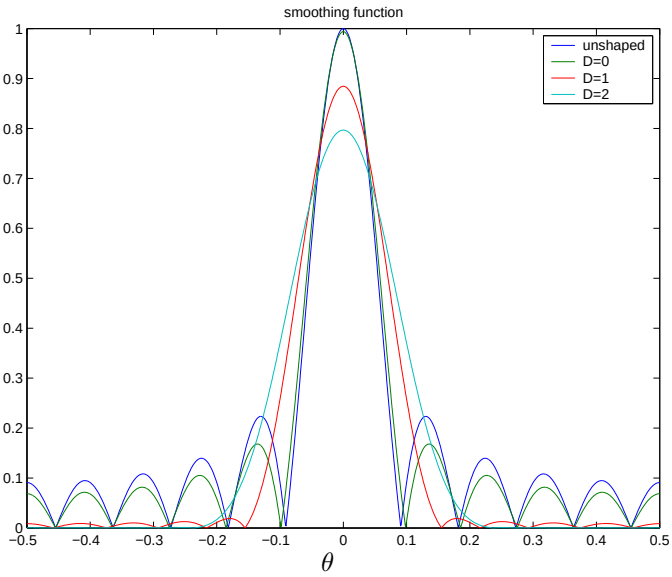

Fig. 2. Shaped smoothing function magnitude $|\bar{f}_\mathsf{T}(\theta, \boldsymbol{d}_\mathsf{T})|$ for SIR-maximizing $\boldsymbol{d}_\mathsf{T}$ and $\boldsymbol{d}_\mathsf{R}$, $P = Q = 11$, $W_\mathsf{T} = W_\mathsf{R} = \frac{1}{P}$, $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0$, and various choices of $D = D_\mathsf{T} = D_\mathsf{R}$, plotted on a linear scale.

IV. Numerical Experiments

In Figs. 5-8 we show examples of virtual channel coefficient (i.e., $\bar{\boldsymbol{H}}_V$) magnitudes for a randomly generated 2-cluster environment with $L = 100$ paths per cluster. As in the other examples, we used $P = Q = 11$. Cluster widths were chosen to be $W_\mathsf{T} = W_\mathsf{R} = \frac{1}{P}$. The first cluster was centered at $\bar{\theta}_\mathsf{T} = \bar{\theta}_\mathsf{R} = \frac{2.25}{11}$ while the second was centered at $\bar{\theta}_\mathsf{T} = \bar{\theta}_\mathsf{R} = -\frac{1.75}{11}$, corresponding to $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0.25$ and $\Delta_\mathsf{T} = \Delta_\mathsf{R} = -0.25$, respectively. In determining the max-SINR shaping coefficients, we assumed $\Delta_\mathsf{T} = \Delta_\mathsf{R} = 0$ as per the discussion in Section III-A. Figure 5 plots the *unshaped* virtual channel coefficients, which are non-sparse even though the clusters are well separated. Figures 6-8 plot the *shaped* virtual channel coefficients for $D = \{0, 1, 2\}$, respectively. While shaping with $D = 0$ helps to concentrate the physical cluster into a single virtual bin, it still leaves a significant amount of energy in other virtual bins. Shaping with $D = 1$ squeezes each physical cluster almost completely into 3-by-3 virtual blocks; the shaped virtual representation is indeed sparse. With $D = 2$, the shaped virtual clusters grow larger in order to further suppress out-of-cluster energy, and so does not further enhance sparseness for this example. For larger cluster widths, though, larger values of D would be appropriate.

V. Conclusion

In this paper we demonstrated that a simple scaling of antenna inputs and outputs has the potential to create a MIMO channel whose uniform-beam-space (i.e., "virtual" [3]) representation is sparse when the physical scattering environment has well localized scattering clusters. Specifically, sparseness was quantified in terms of SIR, and a max-SIR shaping

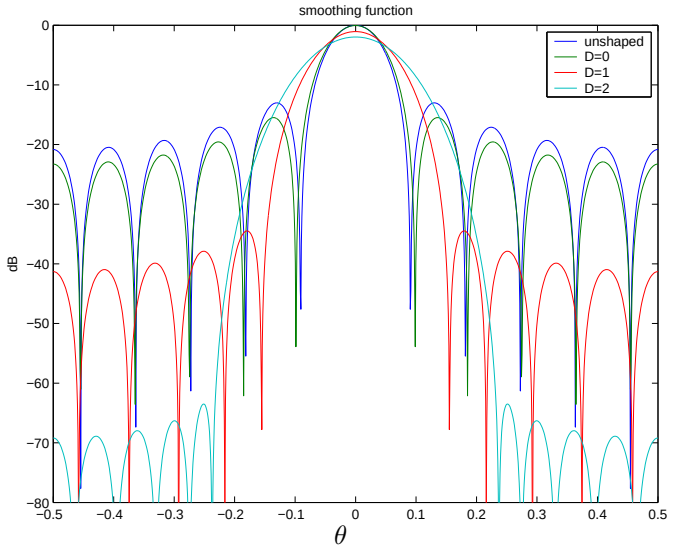

Fig. 3. Shaped smoothing function magnitude $|\bar{f}_{\mathsf{T}}(\theta, \boldsymbol{d}_{\mathsf{T}})|$ for SIR-maximizing $\boldsymbol{d}_{\mathsf{T}}$ and $\boldsymbol{d}_{\mathsf{R}}$, $P = Q = 11$, $W_{\mathsf{T}} = W_{\mathsf{R}} = \frac{1}{P}$, $\Delta_{\mathsf{T}} = \Delta_{\mathsf{R}} = 0$, and various choices of $D = D_{\mathsf{T}} = D_{\mathsf{R}}$, plotted on a log scale.

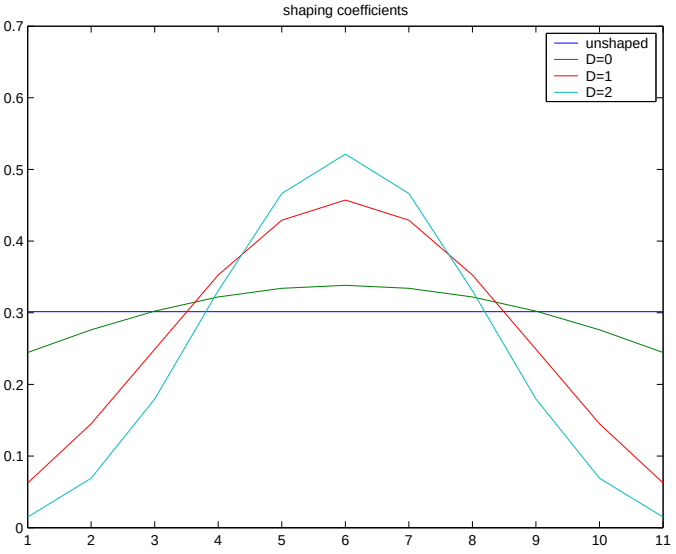

Fig. 4. SIR-maximizing Shaping coefficients $\boldsymbol{d}_{\mathsf{T}}$ for $P = Q = 11$, $W_{\mathsf{T}} = W_{\mathsf{R}} = \frac{1}{P}$, $\Delta_{\mathsf{T}} = \Delta_{\mathsf{R}} = 0$, and various choices of $D = D_{\mathsf{T}} = D_{\mathsf{R}}$.

coefficient design procedure was proposed assuming a single cluster with uniformly distributed scatterers.

APPENDIX

First note that

$$\begin{aligned}
&\big[\bar{\boldsymbol{H}}_V\big]_{\tilde{Q}+q,\tilde{P}+p} \\
&= \boldsymbol{\alpha}_{\mathsf{R}}^H(\tfrac{q}{Q})\,\mathcal{D}(\boldsymbol{d}_{\mathsf{R}})\boldsymbol{H}\,\mathcal{D}(\boldsymbol{d}_{\mathsf{T}}^*)\boldsymbol{\alpha}_{\mathsf{T}}(\tfrac{p}{P}) \\
&= \sum_{l=1}^{L}\beta_l \sum_{c=0}^{Q-1} e^{j2\pi\frac{q}{Q}c} d_{\mathsf{R},c} e^{-j2\pi\theta_{\mathsf{R},l}c} \sum_{b=0}^{P-1} e^{j2\pi\theta_{\mathsf{T},l}b} d_{\mathsf{T},b}^* e^{-j2\pi\frac{p}{P}b} \\
&= \sum_{l=1}^{L}\beta_l\,\boldsymbol{\alpha}_{\mathsf{R}}^H(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\boldsymbol{d}_{\mathsf{R}}\cdot\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{\alpha}_{\mathsf{T}}(\tfrac{p}{P}-\theta_{\mathsf{T},l})
\end{aligned}$$

Then

$$\begin{aligned}
&\mathrm{E}\big|\big[\bar{\boldsymbol{H}}_V\big]_{\tilde{Q}+q,\tilde{P}+p}\big|^2 \\
&= \sum_{l=1}^{L}\mathrm{E}\,|\beta_l|^2\,\boldsymbol{d}_{\mathsf{T}}^H\,\mathrm{E}\big\{\boldsymbol{\alpha}_{\mathsf{T}}(\tfrac{p}{P}-\theta_{\mathsf{T},l})\boldsymbol{\alpha}_{\mathsf{T}}^H(\tfrac{p}{P}-\theta_{\mathsf{T},l})\big\}\boldsymbol{d}_{\mathsf{T}} \\
&\quad\times\boldsymbol{d}_{\mathsf{R}}^H\,\mathrm{E}\big\{\boldsymbol{\alpha}_{\mathsf{R}}(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\boldsymbol{\alpha}_{\mathsf{R}}^H(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\big\}\boldsymbol{d}_{\mathsf{R}}.
\end{aligned}$$

Now, examining the expectation over $\theta_{\mathsf{T},l}$

$$\begin{aligned}
&\big[\mathrm{E}\big\{\boldsymbol{\alpha}_{\mathsf{T}}(\tfrac{p}{P}-\theta_{\mathsf{T},l})\boldsymbol{\alpha}_{\mathsf{T}}^H(\tfrac{p}{P}-\theta_{\mathsf{T},l})\big\}\big]_{c,b} \\
&= \frac{1}{P}\,\mathrm{E}\big\{e^{-j2\pi(\frac{p}{P}-\theta_{\mathsf{T},l})(c-b)}\big\} \\
&= \frac{1}{PW_{\mathsf{T}}}\int_{\tilde{\theta}_{\mathsf{T}}+W_{\mathsf{T}}/2}^{\tilde{\theta}_{\mathsf{T}}+W_{\mathsf{T}}/2} e^{-j2\pi(\frac{p}{P}-\theta_{\mathsf{T},l})(c-b)}d\theta_{\mathsf{T},l} \\
&= \frac{1}{P}e^{-j\frac{2\pi}{P}(p-n-\Delta_{\mathsf{T}})(c-b)}\,\mathrm{sinc}(\pi W_{\mathsf{T}}(c-b)).
\end{aligned}$$

A similar derivation leads to

$$\begin{aligned}
&\Big[\mathrm{E}\big\{\boldsymbol{\alpha}_{\mathsf{R}}(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\boldsymbol{\alpha}_{\mathsf{R}}^H(\tfrac{q}{Q}-\theta_{\mathsf{R},l})\big\}\Big]_{c,b} \\
&= \frac{1}{Q}e^{-j\frac{2\pi}{Q}(q-m-\Delta_{\mathsf{R}})(c-b)}\,\mathrm{sinc}(\pi W_{\mathsf{R}}(c-b)).
\end{aligned}$$

Then

$$\begin{aligned}
&\mathrm{E}\big|\big[\bar{\boldsymbol{H}}_V\big]_{\tilde{Q}+q,\tilde{P}+p}\big|^2 \\
&= \sigma^2\boldsymbol{d}_{\mathsf{T}}^H\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})\boldsymbol{d}_{\mathsf{T}}\cdot\boldsymbol{d}_{\mathsf{R}}^H\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})\boldsymbol{d}_{\mathsf{R}}
\end{aligned}$$

for $\boldsymbol{B}_{\mathsf{T}}(p,n,\Delta_{\mathsf{T}})$ and $\boldsymbol{B}_{\mathsf{R}}(q,m,\Delta_{\mathsf{R}})$ defined in (7) and (8).

REFERENCES

[1] D. B. Kilfoyle and A. B. Baggeroer, "The state of the art in underwater acoustic telemetry," vol. 25, Jan. 2000.

[2] D. Brady and J. C. Preisig, "Underwater acoustic communications," in *Wireless Communications: Signal Processing Perspectives* (H. V. Poor and G. W. Wornell, eds.), ch. 8, pp. 330–379, Prentice-Hall, 1998.

[3] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE Trans. on Signal Processing*, pp. 2563–2579, Oct. 2002.

[4] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell System Technical Journal*, vol. 52, pp. 1541–1562, Nov. 1973.

[5] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. on Signal Processing*, vol. 52, Apr. 2004.

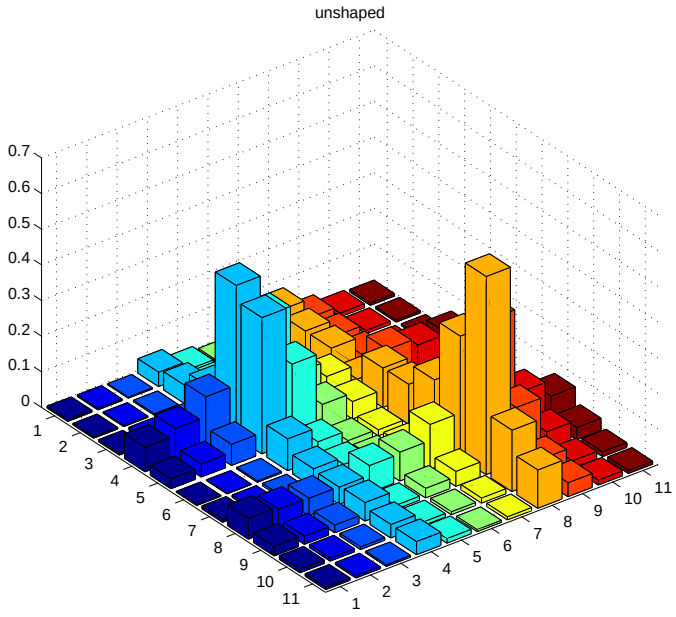

Fig. 5. $\boldsymbol{H}_V$ coefficient magnitudes for 2-cluster scattering. (No shaping.)

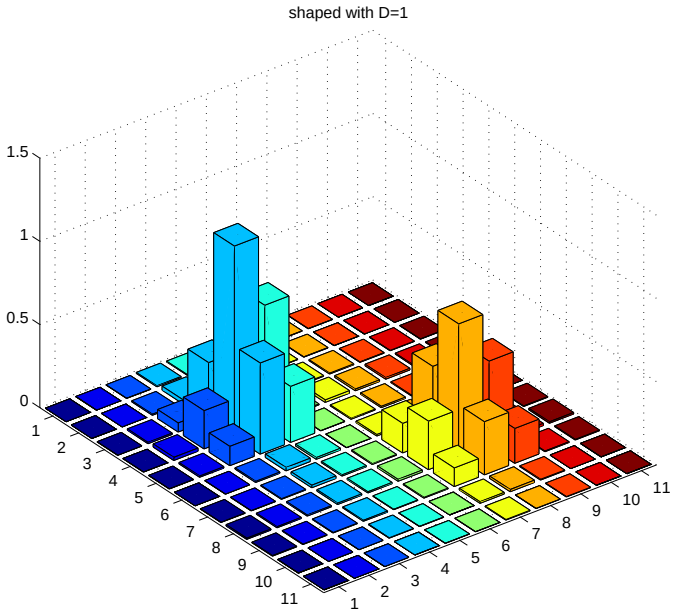

Fig. 7. $\bar{\boldsymbol{H}}_V$ coefficient magnitudes for 2-cluster scattering with SIR-optimal shaping for $D = D_{\mathsf{T}} = D_{\mathsf{R}} = 1$.

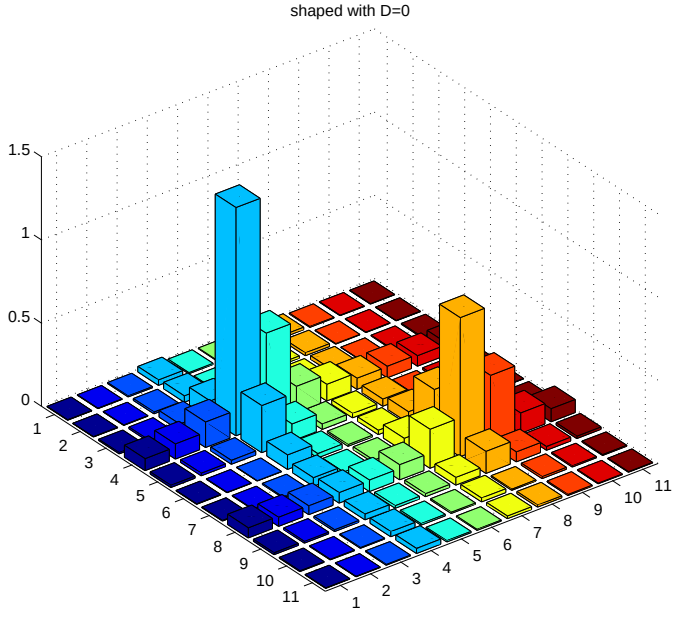

Fig. 6. $\bar{\boldsymbol{H}}_V$ coefficient magnitudes for 2-cluster scattering with SIR-optimal shaping for $D = D_{\mathsf{T}} = D_{\mathsf{R}} = 0$.

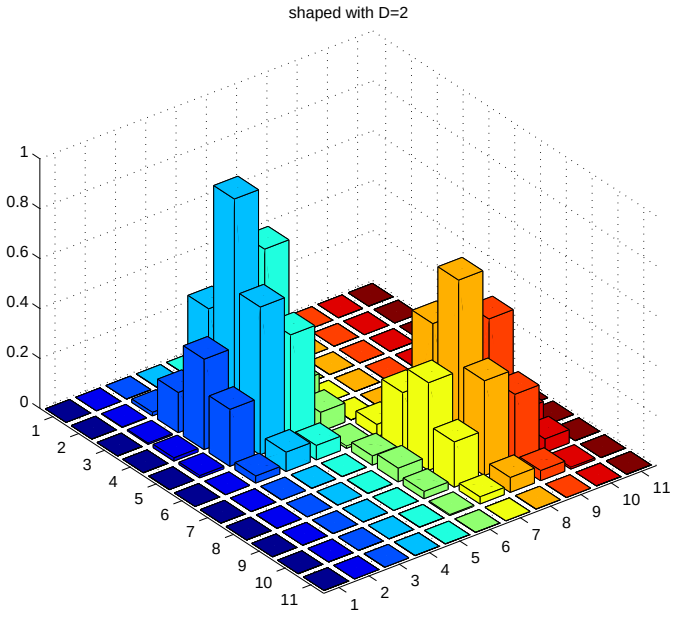

Fig. 8. $\bar{\boldsymbol{H}}_V$ coefficient magnitudes for 2-cluster scattering with SIR-optimal shaping for $D = D_{\mathsf{T}} = D_{\mathsf{R}} = 2$.