

Compressive Channel Estimation and Tracking for Large Arrays in mm-Wave Picocells

Zhinus Marzi, Dinesh Ramasamy, and Upamanyu Madhow, *Fellow, IEEE*

Abstract—We propose and investigate a compressive architecture for estimation and tracking of sparse spatial channels in millimeter (mm) wave picocellular networks. The base stations are equipped with antenna arrays with a large number of elements (which can fit within compact form factors because of the small carrier wavelength) and employ radio frequency (RF) beamforming, so that standard least squares adaptation techniques (which require access to individual antenna elements) are not applicable. We focus on the downlink, and show that “compressive beacons,” transmitted using pseudorandom phase settings at the base station array, and compressively processed using pseudorandom phase settings at the mobile array, provide information sufficient for accurate estimation of the two-dimensional (2D) spatial frequencies associated with the directions of departure of the dominant rays from the base station, and the associated complex gains. This compressive approach is compatible with coarse phase-only control, and is based on a near-optimal sequential algorithm for frequency estimation which approaches the Cramér Rao Lower Bound. The algorithm exploits the geometric continuity of the channel across successive beaconing intervals to reduce the overhead to less than 1% even for very large (32×32) arrays. Compressive beaconing is essentially omnidirectional, and hence does not enjoy the SNR and spatial reuse benefits of beamforming obtained during data transmission. We therefore discuss system level design considerations for ensuring that the beacon SNR is sufficient for accurate channel estimation, and that inter-cell beacon interference is controlled by an appropriate reuse scheme.

Index Terms—Compressive, mm wave, picocells, RF beamforming, 60 GHz.

I. INTRODUCTION

THE explosive growth in demand for wireless mobile data, driven by the proliferation of ever more sophisticated handhelds creating and consuming rich multimedia, calls for orders of magnitude increases in the capacity of cellular data

Manuscript received June 01, 2015; revised November 21, 2015; accepted January 05, 2016. Date of publication January 22, 2016; date of current version April 14, 2016. This work was supported in part by the National Science Foundation under Grants CNS-1317153 and CNS-1518812, in part by the Institute for Collaborative Biotechnologies under the Grant W911NF-09-0001 from the U.S. Army Research Office, and in part by the Systems on Nanoscale Information fabriCs (SONIC), one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP), a Semiconductor Research Corporation program sponsored by MARCO and DARPA. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Sundeeep Rangan.

The authors are with the Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: zh_marzi@ece.ucsb.edu; dineshr@ece.ucsb.edu; madhow@ece.ucsb.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/JSTSP.2016.2520899

networks [1]. Millimeter wave communication from picocellular base stations to mobile devices is a particularly promising approach for meeting this challenge because of two reasons. First, there are huge amounts of available spectrum, enabling channel bandwidths of the order of GHz, 1–2 orders of magnitude higher than those in existing systems at lower carrier frequencies. Indeed, channel bandwidths could potentially increase even further with advances in transceiver technology such as bandwidth/power consumption/linearity tradeoffs for ultra high-speed analog electronics, and speed/precision/power consumption tradeoffs for analog-to-digital converters. Second, the small carrier wavelength enables the realization of highly directive steerable arrays, with a large number of antenna elements, in compact form factors, thus significantly enhancing spatial reuse. In this paper, we address fundamental signal processing challenges associated with channel estimation and tracking for such large arrays, placed within the context of system design for a mm wave picocellular network.

While the signal processing and system design concepts presented here are of rather general applicability, our numerical results are for a particular setting that we feel has great promise, as also discussed in some of our prior publications [2], [3]. We propose to employ the 60 GHz unlicensed band for base station to mobile communication in outdoor picocells: an example urban canyon deployment is depicted in Fig. 1 (discussed further in Section III). At the base station, we consider very large 32×32 arrays (such 1000-element arrays are still only palm-sized at a carrier wavelength of 5 mm) targeting the long term, as well as “moderately sized” 8×8 arrays (which can fit within an area of about half a square inch) which are currently realizable. Note that 16-element arrays were reported several years ago [4], and are already deployed in existing 60 GHz products, while 32-element arrays have been prototyped [5]. We assume that mobile devices are equipped with smaller 4×4 antenna arrays. We focus on downlink 60 GHz communication, with the goal of enabling base station arrays to perform transmit beamforming towards mobile devices, despite the challenges posed by mobility and blockage (which occurs more easily at smaller wavelengths). We do not count on reciprocity. The uplink could be a standard LTE or WiFi link at lower carrier frequencies, used both for uplink data (not modeled here) and feedback for enabling spatial channel estimation at the transmitter.

Multiple antenna systems at lower carrier frequencies have a relatively small number of elements, each with its own radio frequency (RF) chain. This provides control of the individual baseband signals associated with each element, enabling sophisticated adaptation, including frequency-selective spatiotemporal processing (e.g., per subcarrier beamforming in

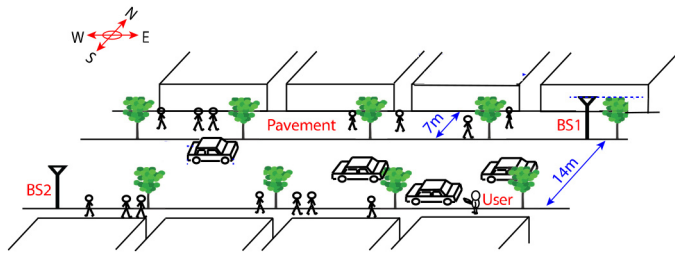


Fig. 1. Picocellular network deployed along an urban canyon.

OFDM systems). This approach does not scale when we have a large number of antenna elements packed into a tiny form factor. Instead, we consider RF beamforming, in which a common baseband signal is routed to/from the antenna elements, and we can only control the amplitude and phase for each element. Indeed, we go even further, assuming that the amplitude for each element is fixed, and that we can only apply coarse four-phase control for each element. Standard least squares array adaptation and channel estimation techniques, which require access to the baseband signals associated with each element, do not apply in this setting. Instead, we consider here a *compressive* approach which exploits the sparsity of the mm wave channel, so that relatively few measurements are required for channel estimation despite the large number of array elements.

1) *Contributions*: Our contributions are summarized as follows:

Architecture: We propose a novel architecture in which base stations send out compressive beacons, with a different set of pseudorandom phases used to transmit each beacon. Each mobile measures the complex gains associated with each beacon compressively, using pseudorandom control of the phases of its receive array. The scheme, described in more detail later, is depicted in Fig. 2. The base station sends M beacons, repeated L times, which permits the mobile to use L different settings of its own array to measure the associated complex gains. Each mobile feeds back appropriately chosen functions of its measurements to the base stations on the uplink (which can be a standard LTE link). Each base station uses this information to estimate and track the dominant paths to each mobile that it receives feedback from.

Algorithms: For the regular two-dimensional (2D) arrays considered here, directions of arrival/departure map to 2D spatial frequencies. The base station estimates the spatial frequencies to each mobile using a simple sequential algorithm, shown to be near-optimal (in terms of approaching the Cramer-Rao Bound) in related publications. The algorithm exploits the geometric continuity of the channel across successive beaconing intervals to reduce the required number of compressive measurements.

System Design: While we do not provide a complete system design centered around our compressive strategy, we do provide preliminary results addressing some of the most important issues. We show that the overhead incurred by our beaconing scheme is very small (less than 1%). Furthermore, while compressive beacons are essentially omnidirectional (in contrast to the highly directive beams employed for communication), the link budget suffices for accurate channel estimation, and

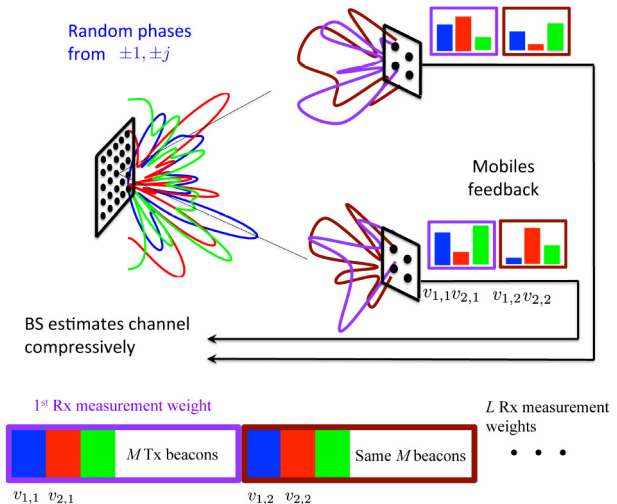


Fig. 2. Channel sounding scheme: The transmitter repeats the M transmit beacons L times so that receivers can measure the channel $v_{i,j}$ corresponding to the i th setting at the transmitter and the j th at the receiver.

a simple beacon reuse strategy suffices to control inter-beacon interference across picocells.

II. BACKGROUND AND RELATED WORK

There is a significant body of recent literature on the potential for mm wave communication for next generation mobile cellular networks [3], [6]–[10]. For these outdoor settings, most of the attention has focused on bands other than 60 GHz; for example, [10] studies blockage probabilities and achievable throughput based on measurements in the 28, 38, 71–76 and 81–86 GHz bands. However, as pointed out in our prior work [3], 60 GHz has immense potential at the relatively short ranges of interest in urban picocells, with the oxygen absorption characteristic of this band only having a marginal impact on the link budget (e.g., only 1.6 dB at 100 m). While studies such as [7]–[10] focus on coarse-grained channel statistics (and their implications on system design and performance), our focus here is on signal processing for fine-grained channel estimation.

We assume that the channel is well described by a relatively small number of discrete rays with delays and angles of arrival/departure taking values in a continuum. The key contribution of this paper is to provide a super-resolution framework for estimating and tracking these rays, with model-based estimation allowing us to go beyond (spatial) bandwidth based resolution limits. To the best of our knowledge, the existing literature on mm wave measurements does not attempt to super-resolve channels in this fashion, hence we do not know, for example, whether the continuous power-delay profiles reported in [10] are consistent with a parsimonious channel model such as ours. However, preliminary experimental results [11], which use compressive measurements to successfully recover power-angle profiles for a controlled experiment (a small number of reflectors in an anechoic chamber), indicate that a simple ray-based model like ours may well suffice. Validating this assertion would require application of the more sophisticated compressive estimation techniques discussed here, as opposed

to the standard basis pursuit algorithms employed in [11]. Note that existing models for simulation-based evaluations, such as the channel models standardized for IEEE 802.11ad indoor 60 GHz channels, typically *assume* more complex models which are variants of the Saleh-Valenzuela model, with a number of clusters, each consisting of multiple closely spaced rays. Such cluster-based models could be motivated by the roughness of reflecting surfaces such as walls, but they have not been experimentally demonstrated. While these issues fall beyond the scope of the present paper, which aims to make a fundamental contribution to signal processing for sparse spatial channels, a sustained effort in measurement-based validation of our model and approach is a critically important topic for future work.

An alternative approach to spatial channel estimation with RF beamforming is exhaustive search, but its overhead does not scale for the large arrays of interest to us. Hierarchical codebook search is more efficient [12]–[15], but it cannot guarantee overhead reduction in a multiuser setting. Furthermore, the multi-resolution beam patterns used in hierarchical search require more than one RF chain [15], and these methods only provide binned estimates of path directions. In contrast, the method presented here provides fine-grained estimates while using only a single RF chain.

A hardware enhancement to pure RF beamforming as considered here (where a single RF chain serves all antenna elements) is to employ hybrid analog-digital beamforming, with a number of RF chains smaller than the number of antenna elements. We may term this an *array of subarrays*, with RF-level control for subarrays, and digital processing of subarray outputs. Such an approach is used in [15], [16] for spatial channel estimation. Our work shows, however, that a single RF beamformed array suffices for this purpose. Of course, arrays of subarrays are certainly required for more advanced functionalities such as multiuser MIMO [17], spatial multiplexing [18]–[22], and spatial diversity [20], [23], [24]. Integrating the compressive approach proposed in this paper within an array of subarrays architecture is an interesting area for future work.

The present paper builds on our prior conference papers on compressive array adaptation [2], [25], but goes well beyond them in several respects. In addition to a more detailed development of the analytical framework underlying our estimation algorithm, we now explicitly model the receive array at the mobile, which requires a generalization of the beaconing and feedback strategy. We also address system level design for compressive tracking in far greater detail, discussing link budget and overhead, and accounting for inter-cell beacon interference. Our initial work on compressive array adaptation [25] subsequently led to a general framework for compressive estimation [26], [27], which identify the isometries required to preserve fundamental bounds such as the Ziv-Zakai (ZZB) and Cramer-Rao (CRB), and use the relationship between these bounds to provide criteria for determining the minimum number of compressive measurements required to preserve geometry and to permit accurate parameter estimation based on a signal corrupted by an AWGN. We now adapt these general results here to develop guidelines for system-level parameter choices. We have also pointed out recently [28], [29] that the ideas behind our

algorithm apply to the fundamental problem of sparse modeling using a mixture of sinusoids, which has a range of applications, including super-resolution in the context of radar and (time domain) multipath channel estimation. The setting in the present paper is more complicated than in [28], [29] due to the use of compressive weights at both the transmitter and the receiver, but there are a common set of key ideas, including orthogonal matching pursuit with Newton refinement, and a Constant False Alarm Rate (CFAR) stopping criterion.

III. SYSTEM MODEL

Given the high demand for wireless data in dense urban environments, we focus our modeling and performance evaluation on the urban canyon setting depicted in Fig. 1, with streets flanked by buildings on both sides. Picocellular base stations are deployed on lampposts in a zig-zag pattern on both sides of the street. We consider mm wave transmission on the downlink (for beaconing and downlink data) and LTE or WiFi at lower carrier frequencies on the uplink (for feedback and uplink data). In terms of channel estimation and tracking, this could be viewed as a worst-case assumption, since two-way transmission on the same mm wave band could potentially be exploited using channel reciprocity. For the east-west street shown, each base station has two “faces,” facing east and west, respectively. Each face can have multiple antenna arrays, but for simplicity, we consider a single antenna array for each face, used for both compressive beaconing and downlink data communication. Mobile stations are modeled as either pedestrians walking along sidewalks, or cars moving along the street.

In our simulations, we model $K = 4$ dominant paths from base station to mobile in our simulations: the line of sight (LoS), and the single bounce reflections from the ground or the side walls. Some of these paths may be blocked by obstacles (diffraction around obstacles is limited for small carrier wavelengths). We ignore multiple bounces, since such paths get attenuated significantly, especially because each bounce sees a smaller angle of incidence than a typical single bounce. However, our compressive estimation algorithm does not use the preceding assumptions on number of dominant paths as prior information, and automatically discovers and tracks paths.

For a regular d -spaced square $N_{1D} \times N_{1D}$ antenna array and a point transmitter in the far-field, the channel seen by the array is a mixture of 2D sinusoids, each corresponding to a propagation path, and is given by

$$h_{m,n} = \sum_{k=1}^K g_k e^{j(\omega_{x,k}m + \omega_{z,k}n)}, g_k \in \mathbb{C}, 1 \leq m, n, \leq N_{1D},$$

where g_k is the propagation gain of the k -th path, $\omega_k = (\omega_{x,k}, \omega_{z,k})$ are the spatial frequencies corresponding to the k -th path (w.l.o.g. we have assumed that the square array is placed in the x - z plane, with its sides aligned to the coordinate axes) and $h_{m,n}$ refers the channel seen by the m, n -th antenna element. The spatial frequencies of the k -th path are given by $\omega_{x,k} = 2\pi(d/\lambda) \sin \theta_k \cos \phi_k$ and $\omega_{z,k} = 2\pi(d/\lambda) \sin \theta_k \sin \phi_k$, where d denotes the array spacing, λ the carrier wavelength and (θ_k, ϕ_k) the inclination and

azimuthal angles of the k -th path relative to $x - z$ plane. We vectorize the 2D sinusoid $[e^{j(\omega_x m + \omega_z n)}, 0 \leq m, n \leq N_{1D} - 1]$ and denote the resulting N_{1D}^2 long vector by $\mathbf{x}(N_{1D}, \boldsymbol{\omega})$, where $\boldsymbol{\omega} = (\omega_x, \omega_z)$ is the frequency of the 2D sinusoid. Vectorizing $[h_{m,n}, 1 \leq m, n \leq N_{1D}]$ in an identical manner gives us

$$\mathbf{h} = \sum_{k=1}^K g_k \mathbf{x}(N_{1D}, \boldsymbol{\omega}_k).$$

Now, consider a base station transmitter with a regularly spaced 2D array of size $N_{t,1D} \times N_{t,1D}$, and a mobile receiver with a regular 2D antenna array of size $N_{r,1D} \times N_{r,1D}$. Let \mathbf{H} denote the corresponding $N_{t,1D}^2 \times N_{r,1D}^2$ channel matrix. Denoting by \mathbf{h}_i the i th row of this matrix, \mathbf{h}_i^T is the response of the receive antenna array to the i th transmit antenna. Denoting $\mathbf{x}(N_{t,1D}, \boldsymbol{\omega})$ by $\mathbf{x}_t(\boldsymbol{\omega})$ and $\mathbf{x}(N_{r,1D}, \boldsymbol{\omega})$ by $\mathbf{x}_r(\boldsymbol{\omega})$, under the far-field assumption, it can be shown that

$$\mathbf{H} = \sum_{k=1}^K g_k \mathbf{x}_t(\boldsymbol{\omega}_k^t) \mathbf{x}_r^T(\boldsymbol{\omega}_k^r), g_k \in \mathbb{C}. \quad (1)$$

Since we know the array geometries (in this case, a regularly spaced 2D array), an estimate of the $N_{t,1D}^2 \times N_{r,1D}^2$ MIMO channel matrix \mathbf{H} can be efficiently arrived at by estimating the spatial frequencies and the associated gains: $\{(g_k, \boldsymbol{\omega}_k^t, \boldsymbol{\omega}_k^r), k = 1, \dots, K\}$. Such a parametric approach is far more efficient than direct estimation of individual entries of \mathbf{H} , and enables us to drastically reduce the number of measurements required.

IV. COMPRESSIVE CHANNEL ESTIMATION

We now describe our compressive approach for spatial channel estimation, which consists of a channel sounding strategy and an estimation algorithm which allows a base station to estimate the propagation gains $\{|g_k|\}$ and the spatial frequencies $\{\boldsymbol{\omega}_k^t\}$ in parallel for *all* mobiles in the picocell.

A. Channel Sounding

The basestation sounds the channel using M compressive beacons. Each beacon is a known signal sent using a different set of transmit weights. The weights are chosen uniformly and independently at random from a small set of coarse phase shifts (for e.g. from the set $\{\pm 1, \pm j\}$, where $j = \sqrt{-1}$). We may therefore view each beacon as being transmitted from a different “virtual antenna” with a quasi-omnidirectional pattern. Each of the M transmit beacons are repeated L times by the basestation (see Fig. 2). For each of these M transmit beacons, a mobile receiver employs L “virtual antennas” to measure the channel response, using receive weights chosen uniformly at random from $\{\pm 1, \pm j\}$. Let $y(m, n)$ denote the response at the (m, n) th receive element due to a given transmit beacon. Letting $b(m, n, l) \in \{\pm 1, \pm j\}$ denote the weight for (m, n) th receive element for the l th virtual receive antenna, the response seen by the l th virtual receive antenna is given by

$$r(l) = \sum_{1 \leq m, n \leq N_{r,1D}} b(m, n, l) \times y(m, n), 1 \leq l \leq L. \quad (2)$$

These measurements are used to construct the $M \times L$ Multiple Input Multiple Output (MIMO) “virtual channel” matrix \mathbf{V} between the virtual transmit and receive antennas. Note that we do not require that the base station know the receive weights used by the mobile, or that the mobile know the transmit weights used by the base station.

Denoting the vectorized version of weights of the i -th virtual transmit antenna by \mathbf{a}_i (a vector of shape $N_{t,1D}^2 \times 1$) and that of the j -th virtual receive antenna by \mathbf{b}_j ($N_{r,1D}^2 \times 1$), the i, j -th element of \mathbf{V} (the channel between the (i, j) -th virtual transmit-receive pair) is given by $v_{i,j} = \mathbf{a}_i^T \mathbf{H} \mathbf{b}_j$. Letting $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_M]^T$ and $\mathbf{B} = [\mathbf{b}_1 \dots \mathbf{b}_L]^T$, it is easy to see that $\mathbf{V} = \mathbf{A} \mathbf{H} \mathbf{B}^T$. Using (1), we have that

$$\mathbf{V} = \sum_{k=1}^K g_k (\mathbf{A} \mathbf{x}_t(\boldsymbol{\omega}_k^t)) (\mathbf{B} \mathbf{x}_r(\boldsymbol{\omega}_k^r))^T. \quad (3)$$

The channel measurements are perturbed by i.i.d measurement noise, and are given by

$$y_{i,j} = \sqrt{P_e} v_{i,j} + z_{i,j}, z_{i,j} \sim \mathcal{CN}(0, \sigma^2),$$

where P_e is the per-element transmit power. Letting \mathbf{Y} and \mathbf{Z} denote $M \times L$ matrices with their i, j -th entries given by $y_{i,j}$ and $z_{i,j}$ respectively, the “measured virtual channel” is given by

$$\begin{aligned} \mathbf{Y} &= \sqrt{P_e} \mathbf{V} + \mathbf{Z} \\ &= \sqrt{P_e} \sum_{k=1}^K g_k (\mathbf{A} \mathbf{x}_t(\boldsymbol{\omega}_k^t)) (\mathbf{B} \mathbf{x}_r(\boldsymbol{\omega}_k^r))^T + \mathbf{Z}. \end{aligned} \quad (4)$$

B. Feedback Strategies

Our goal is to track the mm-wave spatial channel as seen from the basestation, described by the parameters $P = \{(|g_k|, \boldsymbol{\omega}_k^t), k = 1, \dots, K\}$. To this end, every mobile in the picocell needs to feed back a portion of the measured virtual channel \mathbf{Y} to the basestation. From (4), we see that the information regarding the spatial channel as seen from the basestation, given by P , is available in the *column space* of \mathbf{Y} . Building on this observation, we propose two feedback strategies:

- (i) The receiver feeds back the entire matrix \mathbf{Y} .
- (ii) A more elaborate strategy involves feeding back $Q \leq L$ strongest left singular vectors of \mathbf{Y} scaled by their corresponding singular values. i.e., if $\mathbf{Y} = \sum_{l=1}^L \sigma_l \mathbf{u}_l \mathbf{v}_l^H$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_L \geq 0$, the receivers feed back $\mathbf{D} \equiv [\sigma_1 \mathbf{u}_1 \dots \sigma_Q \mathbf{u}_Q]$. This strategy identifies the Q -dimensional *subspace* of the column space of \mathbf{Y} with maximum energy.

V. ESTIMATION ALGORITHM

We now present an algorithm to estimate the parameters $\{(|g_k|, \boldsymbol{\omega}_k^t), k = 1, \dots, K\}$ characterizing the mm-wave channel as seen from the basestation. The same algorithm applies for

both forms of feedback discussed in Section IV-B: the entire measured virtual MIMO matrix \mathbf{Y} or the dominant weighted left singular vectors \mathbf{D} .

The l th column of \mathbf{Y} is given by

$$\mathbf{y}_l = \sum_{k=1}^K h_{k,l} \mathbf{A} \mathbf{x}_t(\omega_k^t) + \mathbf{z}_l, l = 1, \dots, L \quad (5)$$

where $\mathbf{z}_l \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbb{I}_M)$ denotes the l -th column of \mathbf{Z} and $h_{k,l} = \sqrt{P_e} g_k \mathbf{b}_l^T \mathbf{x}_r(\omega_k^r)$. We assume the weight sequence $\{\mathbf{b}_l\}$ used to construct the receive virtual antennas at the receive antenna array is not available at the transmitter, and hence cannot jointly estimate ω_k^t and ω_k^r . However, since $\{\mathbf{b}_l, l = 1, \dots, L\}$ were picked in an i.i.d manner, we have that $\{h_{k,l}, l = 1, \dots, L\}$ are i.i.d realizations of a random variable with $\mathbb{E}|h_{k,l}|^2 = P_e \mathbb{E}|g_k \mathbf{b}_l^T \mathbf{x}_r(\omega_k^r)|^2 = P_e N_{r,1D}^2 |g_k|^2$. This allows us to estimate $|g_k|^2$ as follows:

$$P_e |\hat{g}_k|^2 = (1 / (LN_{r,1D}^2)) \sum_{l=1}^L |\hat{h}_{k,l}|^2.$$

From here on, in Section V, we use the notation ω_k to refer to ω_k^t and $\mathbf{x}(\omega_k)$ to refer to $\mathbf{A} \mathbf{x}_t(\omega_k^t)$. Thus, the measurements can be written as

$$\mathbf{y}_l = \sum_{k=1}^K h_{k,l} \mathbf{x}(\omega_k) + \mathbf{z}_l, l = 1, \dots, L. \quad (6)$$

We now provide an algorithm to estimate $\{\omega_k, \{h_{k,l}\}\}$.

A. Single Path

We first present an algorithm for estimating a single path $K = 1$, which forms the basis for our sequential estimation algorithm for $K > 1$. Omitting the path index k in (6), we have

$$\mathbf{y}_l = h_l \mathbf{x}(\omega) + \mathbf{z}_l, l = 1, \dots, L.$$

Given that $\{\mathbf{z}_l\}$ are independent realizations of $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbb{I}_M)$, the maximum likelihood (ML) estimates of ω , $\{h_l\}$ are given by:

$$\hat{\omega}, \{\hat{h}_l\} = \arg \min_{\omega, \{h_l\}} \sum_{l=1}^L \|\mathbf{y}_l - h_l \mathbf{x}(\omega)\|^2. \quad (7)$$

For any ω , the optimal h_l -s are given by least-squares estimates:

$$h_l^*(\omega) = \langle \mathbf{x}(\omega), \mathbf{y}_l \rangle / \|\mathbf{x}(\omega)\|^2 \quad (8)$$

where $\langle x, y \rangle$ denotes $x^H y$. Plugging into (7), the ML estimate of ω is given by:

$$\hat{\omega} = \arg \max_{\omega} \frac{1}{\|\mathbf{x}(\omega)\|^2} \sum_{l=1}^L |\langle \mathbf{x}(\omega), \mathbf{y}_l \rangle|^2 \quad (9)$$

and the ML estimate of h_l is given by $h_l^*(\hat{\omega})$. We employ a two-step algorithm to arrive the ML estimates: a “detection” phase followed by a “refinement” phase.

1) *Detection Phase*: Using M two-dimensional FFT computations, we precompute $\mathbf{x}(\omega)$ for frequencies of the form $\omega \in \Phi \equiv \{(2\pi i/T, 2\pi j/T), 0 \leq i, j \leq T-1\}$, $T = RN_{1D,t}$, where R is the oversampling factor. We pick the frequency $\hat{\omega}$ from Φ which maximizes (9). The corresponding gains are given by $\hat{h}_l = h_l^*(\hat{\omega})$. We remove the contribution of the newly detected sinusoid from the measured channel response and this residual measurement is given by

$$\mathbf{r}_l = \mathbf{y}_l - \hat{h}_l \mathbf{x}(\hat{\omega}). \quad (10)$$

(This residue is used for sequential detection for $K > 1$, as discussed shortly.)

2) *Refinement Phase*: Our estimate from the detection phase is restricted to the discrete set Φ and consequently we do not expect $\hat{\omega}$ to be equal to the ML estimate given by (9) (where the maximization is over $[-2\pi d/\lambda, 2\pi d/\lambda]^2$ with d being the spacing between transmitter antennas). However, if we make the grid fine enough, the best estimate of ω in Φ is expected to be close enough to the optimal solution to allow refinement via local optimization. In order to do this, we first fix the gain estimates $\{\hat{h}_l\}$ and refine only the estimate of the spatial frequency ω by seeking the minimizer of the ML cost function

$$C(\omega) = \sum_{l=1}^L \left\| \mathbf{y}_l - \hat{h}_l \mathbf{x}(\omega) \right\|^2$$

in the neighborhood of the current estimate $\hat{\omega}$ using the Newton method. This involves evaluating the gradient vector $\mathcal{G}(\omega)$ and the Hessian matrix $\mathcal{H}(\omega)$ of $C(\omega)$ at the current estimate $\hat{\omega}$. The corresponding expressions are given by:

$$\begin{aligned} \mathcal{G}_i(\omega) &= \frac{\partial C(\omega)}{\partial \omega_i} = -2 \sum_{l=1}^L \Re \left\{ \left\langle \mathbf{r}_l, \hat{h}_l \frac{\partial \mathbf{x}(\omega)}{\partial \omega_i} \right\rangle \right\}, \\ \mathcal{H}_{ij}(\omega) &= \frac{\partial^2 C(\omega)}{\partial \omega_i \partial \omega_j} = -2 \sum_{l=1}^L \Re \left\{ \left\langle \mathbf{r}_l, \hat{h}_l \frac{\partial^2 \mathbf{x}(\omega)}{\partial \omega_i \partial \omega_j} \right\rangle \right. \\ &\quad \left. - |\hat{h}_l|^2 \left\langle \frac{\partial \mathbf{x}(\omega)}{\partial \omega_i}, \frac{\partial \mathbf{x}(\omega)}{\partial \omega_j} \right\rangle \right\}, 1 \leq i, j \leq 2 \end{aligned}$$

where $\omega = [\omega_1, \omega_2]$. $\{\mathbf{r}_l\}$ -s are the residual measurements given by (10). The Newton update for $\hat{\omega}$ is

$$\hat{\omega} \leftarrow \hat{\omega} - \mathcal{H}^{-1}(\hat{\omega}) \mathcal{G}(\hat{\omega}). \quad (11)$$

We follow this up by updating our estimates $\{\hat{h}_l\}$ by plugging the new value of $\hat{\omega}$ in (8), i.e.,

$$\hat{h}_l \leftarrow h_l^*(\hat{\omega}) = \langle \mathbf{x}(\hat{\omega}), \mathbf{y}_l \rangle / \|\mathbf{x}(\hat{\omega})\|^2 \quad (12)$$

and modifying the residues $(\{\mathbf{r}_l\})$ accordingly using (10). The algorithm alternates between the updates in (11) and (12) for a few iterations.

B. Multiple Paths

We now build on the preceding single path algorithm for the general setting of $K \geq 1$. Suppose that our current estimate of the sinusoids/paths is given by $\mathcal{P}_K = \{(\hat{\omega}_k,$

$\{\hat{h}_{k,l}\}, k = 1, \dots, K$. The residual measurements corresponding to a set of estimated parameters \mathcal{P} is given by:

$$\mathbf{v}_l(\mathcal{P}) = \mathbf{y}_l - \sum_{\omega_k, \{h_{k,l}\} \in \mathcal{P}} h_{k,l} \mathbf{x}(\omega_k).$$

1) *Detect a New Path*: Assuming that the measurements \mathbf{y}_l are given by the current residue $\mathbf{v}_l(\mathcal{P}_q)$ (corresponding to the q detected paths), we use the single path algorithm in Section V-A to detect and refine a new sinusoid $(\hat{\omega}_{q+1}, \{\hat{h}_{q+1,l}\})$. Let \mathcal{P}_{q+1} denote the new set of estimated parameters $\mathcal{P}_q \cup \{(\hat{\omega}_{q+1}, \{\hat{h}_{q+1,l}\})\}$.

2) *Refine Existing Paths*: Once we add this new path, we refine the parameters of *all* $q+1$ sinusoids in \mathcal{P}_{q+1} one by one. Consider the parameters $(\hat{\omega}_k, \{\hat{h}_{k,l}\})$ of the k -th sinusoid. We use the refinement algorithm in Section V-A to refine $(\hat{\omega}_k, \{\hat{h}_{k,l}\})$ by assuming that the measurements \mathbf{y}_l are given by the residual measurements after excluding the sinusoid of interest. i.e., $\mathbf{v}_l(\mathcal{P}_{q+1} \setminus \{(\hat{\omega}_k, \{\hat{h}_{k,l}\})\})$. Sinusoids are refined in a round robin manner, and the process is repeated for a few rounds: $1 \rightarrow 2 \rightarrow \dots \rightarrow (q+1) \rightarrow 1 \rightarrow \dots \rightarrow (q+1)$.

3) *Stopping Criterion*: The algorithm terminates when the residues $\{\mathbf{v}_l(\mathcal{P}_q)\}$ after refinement satisfy:

$$\max_{\omega \in \text{DFT}} \frac{1}{\|\mathbf{x}(\omega)\|^2} \sum_{l=1}^L |\langle \mathbf{x}(\omega), \mathbf{v}_l(\mathcal{P}_q) \rangle|^2 \leq \tau, \quad (13)$$

with the threshold τ given by $\sigma^2 \gamma^{-1}(L, \Gamma(L)(1 - P_{\text{fa}})^{1/N_{1D,t}^2})$ where $\gamma^{-1}(S, y)$ is the inverse of incomplete gamma function $\gamma(S, x)$ with respect to integral limit x and $\text{DFT} = \{2\pi i/N_{1D,t}, 2\pi j/N_{1D,t} : 0 \leq i, j \leq N_{1D,t} - 1\}$. This choice of τ corresponds to a Constant False Alarm Rate (CFAR) test (the false alarm rate being P_{fa}) for whether the residual measurements $\{\mathbf{v}_l(\mathcal{P}_q)\}$ can be explained by another path. In arriving at this expression for τ we make two assumptions: (i) channel sounding is not compressive: i.e., $\mathbf{A} = \mathbb{I}_{N_{1D,t}^2}$ and (ii) new paths correspond to specific beam orientations given by $\omega \in \text{DFT}$. However, our simulation results show that, this is a good approximation for our case with compressive measurements and continuous values of beam directions. We refer the reader to Appendix A for details on the stopping criterion.

C. Tracking

We sound the channel often enough so that between any two successive channel estimation cycles, the *geometry* of the mm-wave channel, given by the spatial frequencies $\{\omega_k\}$ of the paths, does not change “significantly,” even if the path gains $\{g_k\}$ do. This ensures that angle of departure estimates from the prior sounding round do not become stale over the course of the communication phase during which they are needed for beamforming purposes. For example, if we do not wish to tolerate a beamforming loss of 3 dB or more, then our estimate from the previous round $\hat{\omega}$ should be close enough to the current ω so that

$$|\langle \mathbf{x}(\hat{\omega}), \mathbf{x}(\omega) \rangle|^2 / \|\mathbf{x}(\omega)\|^2 > 0.5$$

over the entire communication phase. This condition is met if $\|\omega - \hat{\omega}\|_\infty < 0.5 \times (2\pi/N_{t,1D})$. Therefore, the estimates of spatial frequencies from the previous sounding round are good approximations of their current true value (within a DFT spacing of $2\pi/N_{t,1D}$). We exploit this by using $\{\hat{\omega}_k, k = 1, \dots, K\}$ from the prior round to initialize our algorithm (as opposed to using the empty set $\{\}$). We do this by constructing the matrix $\mathbf{X} = [\mathbf{x}(\hat{\omega}_1) \dots \mathbf{x}(\hat{\omega}_K)]$ and setting $\hat{h}_{k,l}$ to be the (k, l) -th entries of $(\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{Y}$, where $\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_L]$. We refine all parameters in $\mathcal{P}_K = \{(\hat{\omega}_k, \{\hat{h}_{k,l}\}), k = 1, \dots, K\}$ using the refinement algorithm in Section V-B before proceeding to seek for new paths using the algorithm in Section V-B.

1) *Deleting Weak Paths*: Paths estimated in prior rounds may not be viable at the current time instant (e.g., because of blockage). Therefore, we need means to remove such stale paths. We use the stopping criterion (13) as a means to delete weak paths. If deleting the path under question (say k) and optimizing other parameters $\mathcal{P}_{k,\text{opt}} \leftarrow \text{Refine}(\mathcal{P} \setminus \{(\hat{\omega}_k, \{\hat{h}_{k,l}\})\})$ using our refinement algorithm results in residual measurements $\{\mathbf{v}_l(\mathcal{P}_{k,\text{opt}})\}$ that can be explained by noise (i.e., satisfies (13)), we delete the path permanently. Otherwise we keep the path.

VI. PROTOCOL PARAMETER CHOICES

In this section, we give a principled approach to choosing parameters of the compressive channel estimation protocol, namely the number of unique transmit beacons M , the number of receive measurement weights L and the minimum effective Signal to Noise Ratio (SNR) needed for channel estimation, which we use to choose the sounding bandwidth W_s . We then turn to the question of how frequently the channel has to be sounded. In Section VIII, we take two scenarios and apply this recipe to arrive at system level parameters for the protocol.

A. Number of Compressive Transmit Beacons

Classical compressive sensing aims to reconstruct signals which are sparse in a discrete basis, based on a small number of projections. In order for reconstruction to be successful, these projections must preserve the underlying geometry [30]. In [27], these ideas are extended to compressive estimation, in which a small number of projections are used to estimate continuous-valued parameters.

We now translate these concepts to our present context. Our goal is to estimate the spatial frequencies $\{\omega_k^t\}$ from measurements of the form

$$\mathbf{y}_l = \sum_{k=1}^K h_{k,l} \mathbf{A} \mathbf{x}_t(\omega_k^t) + \mathbf{z}_l, l = 1, \dots, L.$$

The algorithm in Section V aims to estimate parameters $\{h_{k,l}, \omega_k\}$ by minimizing the ML cost function:

$$\begin{aligned} & \sum_{l=1}^L \left\| \mathbf{y}_l - \sum_{k=1}^K \hat{h}_{k,l} \mathbf{A} \mathbf{x}_t(\hat{\omega}_k) \right\|^2 \\ &= \sum_{l=1}^L \left\| \mathbf{A} \times \sum_{k=1}^K \left(h_{k,l} \mathbf{x}_t(\omega_k^t) - \hat{h}_{k,l} \mathbf{x}_t(\hat{\omega}_k^t) \right) + \mathbf{z}_l \right\|^2, \end{aligned}$$

where $\{\hat{h}_{k,l}, \hat{\omega}_k^t\}$ refer to our estimates of $\{h_{k,l}, \omega_k^t\}$. If the compressive measurement matrix \mathbf{A} ensures that

$$\begin{aligned} & \left\| \mathbf{A} \times \sum_{k=1}^K \left(h_{k,l} \mathbf{x}_t(\omega_k^t) - \hat{h}_{k,l} \mathbf{x}_t(\hat{\omega}_k^t) \right) \right\|^2 \\ & \approx M \left\| \sum_{k=1}^K \left(h_{k,l} \mathbf{x}_t(\omega_k^t) - \hat{h}_{k,l} \mathbf{x}_t(\hat{\omega}_k^t) \right) \right\|^2, \forall h_{k,l}, \hat{h}_{k,l}, \end{aligned} \quad (14)$$

for relevant $(\{\omega_k\}, \{\hat{\omega}_k\})$ -pairs, the cost structure of the estimation problem is roughly preserved. Therefore, estimation using compressive measurements is similar to estimation with all $N_{t,1D}^2$ measurements (except for a reduction in effective SNR, given by $M/N_{t,1D}^2$) [27].

We now state an isometry property relevant for our purpose. For some fixed ϵ , the matrix $\Phi \in \mathbb{C}^{m \times p}$ is said to enjoy the s -isometry property for the basis B (of size $p \times n$) if there exists a constant $C > 0$ such that

$$C(1 - \epsilon) \leq |\Phi B \mathbf{u}|^2 / |B \mathbf{u}|^2 \leq C(1 + \epsilon),$$

for all s -sparse \mathbf{u} in \mathbb{C}^n . It can be shown using the celebrated Johnson-Lindenstrauss Lemma (JL Lemma) [31] (with arguments similar to those in [30]) that if $m = O(s\epsilon^{-2} \log n)$, a randomly picked Φ (with elements drawn from proper distributions, e.g., $Uniform\{\pm 1, \pm j\}$ [27], [31]) satisfies this isometry property with high probability.

In order to provide a concrete definition of geometry preservation in our setting, we consider a discretized version of the problem, when spatial frequencies are restricted to an oversampled DFT grid G of size $R = O(N_{t,1D}^2)$. The condition (14) reduces to a $2K$ -isometry property of the measurement matrix \mathbf{A} relative to the basis \mathbf{X} , where \mathbf{X} is the $N_{t,1D}^2 \times R$ matrix with its columns given by $\{\mathbf{x}_t(\omega) : \omega \in G\}$. We can now apply the preceding isometry property to conclude that, when the number of unique transmitter beacons scales as $M = O(K\epsilon^{-2} \log R) = O(K\epsilon^{-2} \log N_{t,1D})$, then the $2K$ -pairwise isometry criterion w.r.t the basis \mathbf{X} is met by the randomly picked sounding matrix \mathbf{A} (it can be shown that $C = M$ for our choice of scale), thereby ensuring that the geometry of the spatial frequency estimation problem is preserved.

While the preceding calculations give order estimates for the number of measurements M required, in order to provide numerical values for our protocol design, we employ simulations. We consider the example of the 32×32 transmitter array and plot the maximum and minimum values of $(1/M)\|\mathbf{A}\mathbf{X}\mathbf{u}\|^2/\|\mathbf{X}\mathbf{u}\|^2$ from 5×10^6 random realizations of a $2K = 8$ -sparse \mathbf{u} (which represent different realizations of $K = 4$ paths and their corresponding potential estimates) in Fig. 3 using the 64-times oversampled DFT grid as the choice of basis \mathbf{X} . We see that this ratio is within $[-5, 3]$ dB when $M \geq 30$. This illustrates that for estimating $K = 4$ paths using a 32×32 array, measuring the response corresponding to $M = 30$ random beacons approximates the effect of measuring all $32 \times 32 = 1024$ antenna elements individually.

B. Number of Compressive Receive Measurements

While we do not track $\{\omega_k^r, k = 1, \dots, K\}$, the spatial frequencies at the receiver, we need to ensure that the set of

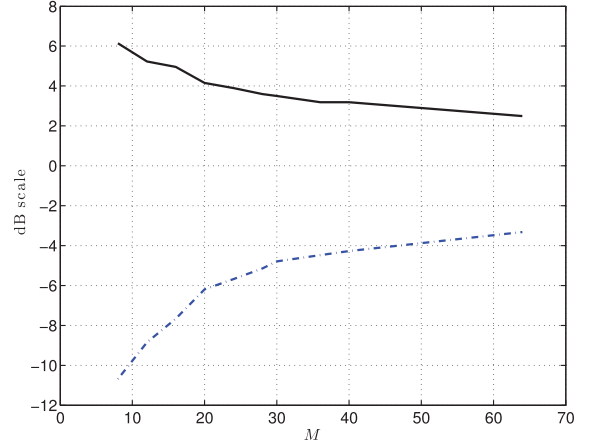


Fig. 3. Maximum and minimum values of $\|\mathbf{A}\mathbf{X}\mathbf{u}\|^2/(M\|\mathbf{X}\mathbf{u}\|^2)$ for different values of M , the number of transmitter beacons, across 5×10^6 random realizations of 8-sparse \mathbf{u} . The basis \mathbf{X} corresponds to the responses for a 32×32 array evaluated uniformly over a $R = 64 \times 32^2$ -sized grid.

measurements made at the receiver have sufficient information to estimate transmitter spatial frequencies. Suppose that $|\mathbf{B}\mathbf{x}_r(\omega_k^r)| \approx 0$, it follows from $h_{k,l} = g_k \sqrt{P_e} \mathbf{b}_l^T \mathbf{x}_r(\omega_k^r)$ that all L measurements $\{\mathbf{y}_l, 1 \leq l \leq L\}$ will have very small contributions from the k -th path. i.e., $|h_{k,l}| \approx 0, 1 \leq l \leq L$. To see this observe that:

$$\begin{aligned} \sum_{l=1}^L |h_{k,l}|^2 &= P_e |g_k|^2 \sum_{l=1}^L |\mathbf{b}_l^T \mathbf{x}_r(\omega_k^r)|^2 \\ &= P_e |g_k|^2 \|\mathbf{B}\mathbf{x}_r(\omega_k^r)\|^2. \end{aligned} \quad (15)$$

Again, restricting the receive spatial frequencies to an oversampled DFT grid G of size $R = O(N_{r,1D}^2)$, it can be shown that for $L = O(\log R) = O(\log N_{r,1D})$, $|\mathbf{B}\mathbf{x}_r(\omega)|^2 \approx L|\mathbf{x}_r(\omega)|^2 = LN_{r,1D}^2$ w.h.p. This ensures that

$$\sum_{l=1}^L |h_{k,l}|^2 \approx P_e LN_{r,1D}^2 |g_k|^2 \text{ w.h.p,}$$

thereby capturing power along the k -th path. We perform computations for the maximum power lost across spatial frequencies when using a 4×4 array and plot the results in Fig. 4. This shows that around 5 carefully chosen projections (we pick the best measurement matrix from 10^4 random instances) suffice to ensure that SNR degradation (relative to the nominal value of L) is no greater than 3 dB for a 4×4 receive array.

C. SNR for Successful Estimation

The preceding criteria delineate the regime in which the geometry of the estimation problem is preserved approximately. We now turn to another factor which affects estimation performance, namely the SNR. Consider measurements of the form

$$y_{m,n} = e^{j(\omega_1 m + \omega_2 n + \phi)} + z_{m,n}, 0 \leq m, n \leq N_{t,1D} - 1, \quad (16)$$

where $z_{m,n}$ are i.i.d. $\mathcal{CN}(0, \sigma^2)$ and spatial frequencies ω_1, ω_2 and phase ϕ are parameters to be estimated. The Cramér

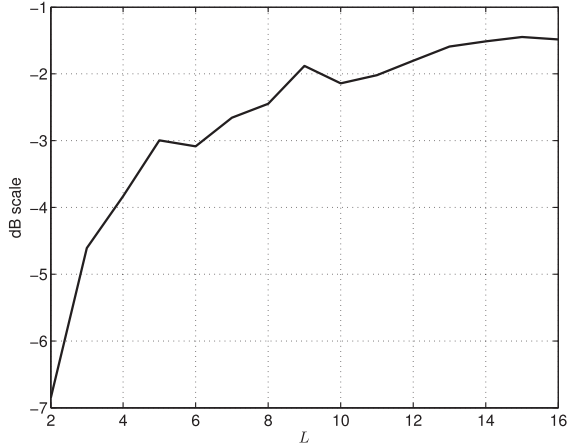


Fig. 4. Maximum SNR degradation $\min_{\omega} |\mathbf{B}\mathbf{x}_r(\omega)|^2 / (L|\mathbf{x}_r(\omega)|^2)$ for the most favorable realization (from 10^4 runs) of an $L \times N_{r,1D}^2$ matrix \mathbf{B} with $N_{r,1D} = 4$.

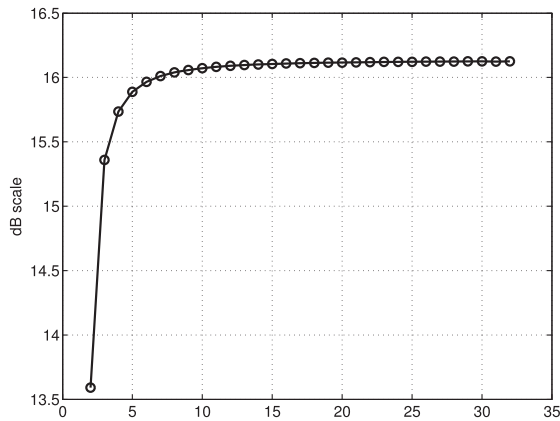


Fig. 5. ZZZB threshold SNR SNR_{th} for estimating the frequency of a $N_{t,1D} \times N_{t,1D}$ sinusoid as a function of $N_{t,1D}$.

Rao Bound [32] (CRB) for estimating ω_1 from measurements (16) is given by $C(\sigma^2) = 6/(\text{SNR}(N_{t,1D}^2 - 1))$, where $\text{SNR} = |\mathbf{x}_t(\omega)|^2 / \sigma^2 = N_{t,1D}^2 / \sigma^2$ (same expression holds for ω_2). Assuming an uniform prior over $[0, 2\pi]^3$ for the parameters $(\omega_1, \omega_2, \phi)$, the Ziv Zakai Bound (ZZB) with periodic distortion [33] for estimating ω_1 evaluates to:

$$Z(\text{SNR}) = \int_0^\pi Q \left(\sqrt{\text{SNR} \left(1 - \left| \frac{\sin(N_{t,1D}h/2)}{N_{t,1D} \sin(h/2)} \right| \right)} \right) h dh.$$

An indicator of the SNR needed for successful estimation is the convergence of the ZZB to the CRB [27]. We use the SNR beyond which the ZZB is within 0.1 dB of the CRB as a measure of this convergence. We plot this ZZB *threshold* SNR for different values of $N_{t,1D}$ in Fig. 5. e.g., $\text{SNR}_{\text{th}} = 16.04$ dB for an 8×8 array and $\text{SNR}_{\text{th}} = 16.13$ dB for a 32×32 array.

The total energy E_{tot} corresponding to the k -th path collected across the ML measurements $\{\mathbf{y}_l, 1 \leq l \leq L\}$ is given by:

$$E_{\text{tot}} = \|\mathbf{A}\mathbf{x}_t(\omega_k^t)\|^2 \times \sum_{l=1}^L |h_{k,l}|^2.$$

Using (15) in the above, we have that

$$\begin{aligned} E_{\text{tot}} &= \|\mathbf{A}\mathbf{x}_t(\omega_k^t)\|^2 \|\mathbf{B}\mathbf{x}_r(\omega_k^r)\|^2 P_e |g_k|^2 \\ &\approx MLN_{t,1D}^2 N_{r,1D}^2 P_e |g_k|^2 \\ &= MLPN_{r,1D}^2 |g_k|^2, \end{aligned}$$

where $P = N_{t,1D}^2 P_e$ is the total transmit power supplied to the $N_{t,1D} \times N_{t,1D}$ antenna array. The above approximation holds when M and L satisfy the preceding geometry preservation criteria in Sections VI-A and VI-B respectively. The effective SNR of the i -th path is given by $\text{SNR}_{\text{eff}} = E_{\text{tot}} / \sigma^2$. It is important to note that the per-measurement noise variance σ^2 is given by $\sigma^2 = N_{r,1D}^2 \sigma_e^2$, where σ_e^2 is the noise variance per antenna element. Assuming no interference (which we account for in Section VII-D), $\sigma_e^2 = N_0 W_s$ with W_s denoting the sounding bandwidth and N_0 the thermal noise floor. The reason for the scale factor $N_{r,1D}^2$ in the expression for σ^2 is the following: Our measurement process consists of multiplying the received signal at each antenna (of which there are $N_{r,1D}$) by phasors $b(m, n, l) \in \{\pm 1, \pm j\}$ and *adding* the resultant signal (as per (2)). Since thermal noise seen by the $N_{r,1D}$ isolated receive antennas are independent random variables, we have that $\sigma^2 = N_{r,1D}^2 \times (N_0 W_s)$. Therefore, the effective SNR of the i -th sinusoid is given by:

$$\text{SNR}_{\text{eff}} = MLP|g_k|^2 / N_0 W_s.$$

This must exceed the ZZB threshold SNR_{th} for successful estimation. Noting that ML/W_s is the time taken for channel sounding, the ZZB threshold SNR_{th} gives us the means to evaluate the minimum overhead in time to estimate the channel for a given path gain $|g_k|^2$:

$$\text{Time taken} = ML/W_s \geq \text{SNR}_{\text{th}} N_0 / P |g_k|^2 \quad (17)$$

The size of the picocell gives us a lower bound on $|g_k|^2$ and we later use this to guide us in choosing the sounding bandwidth W_s using (17).

D. Sounding Rate

We round off the discussion on choice of protocol parameters by giving a rule of thumb for the rate/frequency f_B at which the spatial channel $\{\omega_k\}$ needs to be *reestimated*. We use the estimated spatial frequency $\hat{\omega}$ for beamforming purposes in the time period between two channel sounding rounds (communication phase sandwiched between consecutive sounding phases; see Fig. 7). Following the discussion in Section V-C, we have that if $\|\omega(t) - \hat{\omega}\|_\infty < \pi/N_{t,1D}$ throughout the communication phase, where $\omega(t)$ denotes the true spatial frequency and $\hat{\omega}$ the estimate from the prior sounding round, then the *loss* in SNR, given by $\|\mathbf{x}_t(\omega(t))\|^2 / \langle \mathbf{x}_t(\hat{\omega}), \mathbf{x}_t(\omega(t)) \rangle$, is smaller than 3 dB. If we assume that the closest user to the basestation array is at a distance R meters and that the maximum speed of a user in the picocell is given by v_{max} meters per second, then the maximum change (in terms of the ℓ_∞ -norm) in spatial frequency $\Delta\omega$ between consecutive sounding phases, spaced $1/f_B$ apart, is given by $2\pi d v_{\text{max}} / f_B R \lambda$. The worst-case geometry which achieves this bound is when the user is

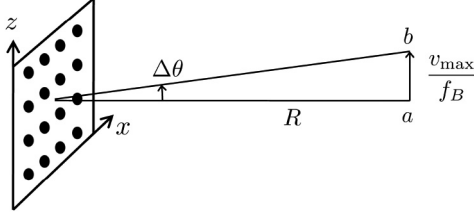


Fig. 6. Geometry corresponding to maximum change in ω_z : The user moves along the z axis at a speed of v_{\max} in the time interval $1/f_B$ between two consecutive channel sounding rounds.

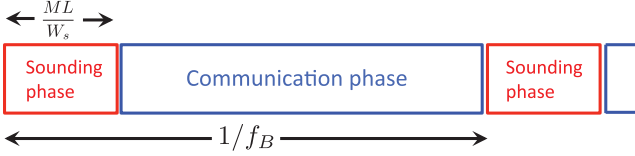


Fig. 7. Channel sounding and communication phases of the proposed system.

at a distance R along the bore-sight of the array and heading in a direction aligned with the one of the array axes. For this worst-case geometry (plotted in Fig. 6), we have that:

$$\Delta\omega \leq (2\pi d/\lambda) \sin \Delta\theta \approx \frac{2\pi d v_{\max}}{f_B \lambda R}.$$

Assuming that the estimate $\hat{\omega}$ from the previous sounding phase is accurate, if we ensure that $2\pi d v_{\max}/f_B R \lambda \leq \pi/N_{t,1D}$, we have that the beamforming losses in the intervening period are smaller than 3 dB. This tells us that channel needs to be sounded often enough so that

$$f_B \geq 2d v_{\max} N_{t,1D}/R\lambda. \quad (18)$$

In the following discussions, we use the preceding in conjunction with (17) to determine the overhead incurred in estimating the channel using the compressive architecture proposed herein.

VII. SYSTEM DESIGN

We now discuss some key aspects of downlink system design related to our compressive architecture. We start by choosing basestation transmit power based on rules set by regulatory authorities, and then filling in the other details of the protocol according to the prescriptions laid out in Section VI. Fixing the mobile array to be 4×4 ($N_{r,1D} = 4$), we consider two different choices for the base station array size: 8×8 ($N_{t,1D} = 8$) and 32×32 ($N_{t,1D} = 32$). All arrays are $d = \lambda/2$ -spaced. The total available bandwidth for communication and sounding is 2 GHz.

A. Transmit Power

We fix the effective isotropically radiated power (EIRP) to 40 dBm, consistent with Federal Communications Commission (FCC) regulations for 60 GHz unlicensed transmission. Accounting for transmit beamforming gain using an $N_{t,1D} \times N_{t,1D}$ array, the total transmit power

$$P = 40 - 20 \log N_{t,1D} \text{dBm} \quad (19)$$

which evaluates to 22 dBm and 10 dBm for $N_{t,1D} = 8$ and $N_{t,1D} = 32$ respectively. Assuming that this power is evenly split among the $N_{t,1D}^2$ transmit elements, the power per transmit element is given by $P_e = P - 20 \log N_{t,1D} \text{dBm} = 40 - 20 \log N_{t,1D} \text{dBm}$, which evaluates to 4 dBm and -20 dBm, respectively. Assuming that we design each element to operate at a fixed power, this is also the power per element used in the beaconing phase, even though the latter does not get the benefit of transmit beamforming.

B. Communication Range

In order to ensure that the SNR for compressive estimation is adequate over a picocell, we first determine the picocell size using a nominal communication link budget, and then calculate the overhead required for successful estimation at that range. Standard link budget calculations, assuming oxygen absorption of 16 dB/km, an EIRP of 40 dBm and a 4×4 receive array providing directivity gains of 12 dBi, can be used to show that we can attain a per-symbol SNR of 6 dB at a link margin of 10 dB for a symbol rate of 2 GHz at a range of 200 m. For omnidirectional free space propagation, the power gain in dB as a function of range r is given by

$$G_{dB}(r) = -\mu r + 20 \log_{10} \frac{\lambda}{4\pi r} \text{(dB)}$$

where $\mu = 0.016$ dB/m to account for oxygen absorption. Note that μ can be increased in order to account for rain. However, since our purpose is to ensure that channel estimation is successful whenever communication is successful, the contribution due to $G_{dB}(r)$ cancels out, as we show shortly. Thus, while the particular value of μ determines picocell size, we shall see that it does not affect the overhead for channel estimation.

The SNR per symbol is given by

$$\text{SNR}_c(\text{dB}) = \text{EIRP}(\text{dBm}) + G_{dB}(r) + 20 \log_{10} N_{r,1D} - 10 \log_{10} (N_0 W_c) - L_{\text{margin}}(\text{comm}) \quad (20)$$

where $L_{\text{margin}}(\text{comm})$ is the link margin (dB) for communication. Note that $10 \log_{10} N_0 = -174 + NF$ dBm over a bandwidth of 1 Hz, where NF denotes the receiver noise figure in dB. Plugging in $W_c = 2$ GHz, $N_{r,1D} = 4$, and $NF = 6$ dB, we obtain a per symbol SNR of 7 dB at a range of $r = 100$ meters.

C. Channel Sounding Protocol

Our channel sounding protocol is specified by four parameters: (i) bandwidth used by each basestation when sounding the channel, which we denote by W_s (ii) number of transmit beacons (or virtual transmit antennas) M (iii) number of receive measurements per transmit beacon (or virtual receive antennas) L and (iv) sounding rate f_B which determines how often the channel is sounded. The parameters M , L and W_s together determine the effective sounding SNR. This must exceed the ZZB threshold SNR for successful channel estimation. This gives rise to the condition in 17. Imposing an estimation link margin $L_{\text{margin}}(\text{est})$ (dB) and going to the dB domain, we have

$$10 \log_{10} (ML/W_s) \geq \text{SNR}_{\text{th}} + L_{\text{margin}}(\text{est}) + 10 \log_{10} N_0 - P - G_{dB}(r) \quad (21)$$

Adding (21) and (20) and simplifying, we obtain

$$10 \log_{10} (ML/W_s) \geq \text{SNR}_{\text{th}} - \text{SNR}_c + L_{\text{margin}}(\text{est}) - L_{\text{margin}}(\text{comm}) + 20 \log_{10} N_{t,1D} + 20 \log_{10} N_{r,1D} - 10 \log_{10} W_c$$

The key take-away is that ML/W_s must be large enough to compensate for the fact that we do not have the benefit of beamforming during the sounding phase. Notice that the range r (i.e., the dependence on picocell size) has cancelled out. Setting $L_{\text{margin}}(\text{est}) = 16$ dB (we use a higher link margin for channel sounding to account for power losses due to randomness of \mathbf{A} and \mathbf{B}), we obtain

$$\text{Time taken} = \frac{ML}{W_s} \geq \begin{cases} 16.34 \mu\text{s} & N_{t,1D} = 8 \\ 0.2669 \text{ ms} & N_{t,1D} = 32. \end{cases} \quad (22)$$

We choose the number of transmitter beacons for the 8×8 and 32×32 transmitter arrays based on the geometry preservation criterion for the transmitter's spatial channel estimation problem discussed in Section VI-A. We use $M = 24$ for $N_{t,1D} = 8$ and $M = 30$ for $N_{t,1D} = 32$ by numerically evaluating the worst-case distortion of pairwise distances relevant for the channel estimation problem (in Fig. 3 we plot the worst-case distortion as a function of L for a random instance of \mathbf{A} and $N_{t,1D} = 32$). Using the receive energy preservation criterion given in Section VI-B, we choose the number of receive weights for the 4×4 receive array as $L = 6$. Using these values for M and L in (22), we obtain that the channel sounding bandwidth must satisfy

$$W_s \leq \begin{cases} 8.8124 \text{ MHz} & N_{t,1D} = 8 \\ 674.34 \text{ KHz} & N_{t,1D} = 32. \end{cases}$$

Our specification of the channel sounding protocol will be complete when we give f_B , the rate at which we sound the channel (see Fig. 7) which must satisfy (18). Assuming that the closest user is at a distance of $R = 20$ m and that the maximum speed of a user in the picocell v_{max} is 45 miles per hour (20 m/s), we have that: $f_B \geq 8$ Hz for $N_{t,1D} = 8$ and $f_B \geq 32$ Hz for $N_{t,1D} = 32$. Choosing the minimum value for f_B , we have that the overhead for our channel sounding protocol is $MLf_B/W_s = 0.0131\%$ for $N_{t,1D} = 8$ and 0.8542% for $N_{t,1D} = 32$.

D. Reuse Analysis for Channel Sounding

We investigate how a sequence of basestations employed in an urban canyon environment can share resources when estimating the spatial channel to users in their respective cells. The envisioned mm-wave system involves alternating between channel estimation and communication phases as shown in Fig. 7. We assume that channel sounding rounds across basestations are aligned in time. We now characterize how the 2 GHz spectrum is to be shared in space so as to limit the effect of interference from neighboring picocells

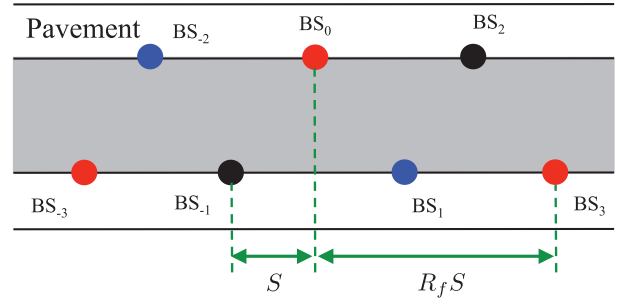


Fig. 8. Reuse of frequency resources for reuse factor $R_f = 3$.

on channel estimation performance. Such interference management is essential in the sounding phase; unlike the highly directive beams used in the communication phase, compressive sounding beacons are essentially omnidirectional. To see this, consider the average transmit power along any direction ω . This is given by $P_e |\mathbf{A}\mathbf{x}_t(\omega)|^2/L \approx P_e |\mathbf{x}_t(\omega)|^2 = P_e N_{t,1D}^2 = P$, the total transmit power. The approximation $|\mathbf{A}\mathbf{x}_t(\omega)|^2/L \approx |\mathbf{x}_t(\omega)|^2$ holds when the number of beacons L is large enough. Therefore, the average energy *per-measurement* received by an antenna at a distance r from a transmitter sending compressive beacons is given by $PG(r)$, where $G(r) = 10^{G_{dB}(r)/10} = \lambda^2/(16\pi^2 r^2) e^{-\nu r}$ ($\nu = (\mu/10) \ln 10$) is the omnidirectional power gain at range r . We assume that basestations are deployed regularly as shown in Fig. 8 and that the inter-basestation separation (along the street) is given by S . Suppose that the reuse factor is R (i.e., every R th basestation uses the same slice of the frequency spectrum to estimate downlink spatial channels). We assume that for narrow urban canyons, the distance between a user and all interfering basestations (those that are allocated the same sounding BW) are well approximated by $\{kR_f S, k \in \mathbb{Z} \setminus \{0\}\}$. Thus, the interference power seen by a single antenna is given by

$$I = 2 \times \sum_{k=1}^{\infty} \sum_{\text{paths}} PG(kR_f S) = 8P \sum_{k=1}^{\infty} G(kR_f S),$$

where we have assumed that there are 4 viable paths between the interfering basestation and user, each introducing the same amount of interference as the LoS path. This is a pessimistic assumption, since NLOS paths are attenuated by larger path lengths and reflection losses. Plugging in the expression for $G(r)$, we have that

$$\begin{aligned} I &= (P\lambda^2/2\pi^2 R_f^2 S^2) \sum_{k=1}^{\infty} e^{-\nu R_f S k} / k^2 \\ &= (P\lambda^2/2\pi^2 R_f^2 S^2) \text{Li}_2(e^{-\nu R_f S}), \end{aligned}$$

where $\text{Li}_2(z) = \sum_{k=1}^{\infty} z^k/k^2$ is the dilogarithm function. The interference seen per antenna adds to thermal noise to give an effective per-element noise level of $\sigma_e^2 = N_0 W_s + I$. Assuming a worst-case geometry for the user of interest (distance of S from the basestation) and proceeding as in Section VI-C, we see that effective Signal to Interference and Noise Ratio SINR_{eff} is given by

$$\text{SINR}_{\text{eff}} = MLPG(S)/\sigma_e^2.$$

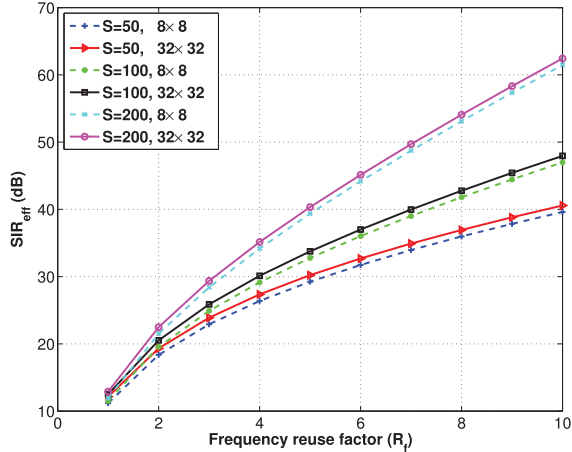


Fig. 9. Effective Signal to Interference Ratio SIR_{eff} for $ML = 24 \times 6$ (8×8 scenario) and $ML = 30 \times 6$ (32×32 scenario; dashed) as a function of reuse factor R_f .

This can be rewritten as

$$1/\text{SINR}_{\text{eff}} = 1/\text{SNR}_{\text{eff}} + 1/\text{SIR}_{\text{eff}},$$

where $\text{SNR}_{\text{eff}} = MLPG(S)/N_0W_s$ and the Signal to Interference Ratio $\text{SIR}_{\text{eff}} = MLPG(S)/I = MLR_f^2 e^{-\mu S} / 8Li_2(e^{-\mu R_f S})$. We need to ensure that SINR_{eff} exceeds the ZZB SNR threshold for successful estimation. We choose the reuse factor R_f so that we are in the noise-limited regime by setting

$$1/\text{SIR}_{\text{eff}} < 0.1 \times (1/\text{SNR}_{\text{th}}) \approx -10 - 16 \text{ dB}.$$

Assuming that protocol parameters are chosen so that SNR_{eff} exceeds SNR_{th} , we can ignore interference in SINR_{eff} calculations when

$$\text{SIR}_{\text{eff}} > \text{SNR}_{\text{th}} + 10 \approx 26 \text{ dB for } N_{t,1D} = 8, 32.$$

In Fig. 9, we plot achievable effective SIRs as a function of frequency reuse factor R_f for two example systems in Section VII-C: i.e., 8×8 and 32×32 arrays with total number of measurements given by $ML = 24 \times 6$ and $ML = 30 \times 6$ respectively. As the picocell size S grows, exponential attenuation due to oxygen absorption (the $e^{-\nu S}$ term in the expression for SIR_{eff}) helps in attenuating interference and improving SIR for same reuse factor R_f . To illustrate this we plot SIR as a function of R_f for three cell sizes $S = 50, 100, 200$ m in Fig. 9. We observe that, in order to ensure $\text{SIR}_{\text{eff}} > 26$ dB, a reuse factor of $R_f = 4$ is needed for $S = 50$ m, while $R_f = 3$ suffices for $S = 200$ m. Plugging in the per-basestation sounding bandwidth W_s calculations in Section VII-C, we see that the overall system-level channel sounding bandwidth $W_s \times R_f$ for a $S = 50$ m picocell is as small as $8.8124 \text{ MHz} \times 4 = 35.2 \text{ mHz}$ and $674.34 \text{ KHz} \times 4 = 2.7 \text{ mHz}$ for $N_{t,1D} = 8, 32$, respectively, which is dwarfed by the total available bandwidth (2 GHz).

VIII. SIMULATION RESULTS

We perform simulations for the two example systems considered in Section VII (8×8 and 32×32 transmit arrays). We

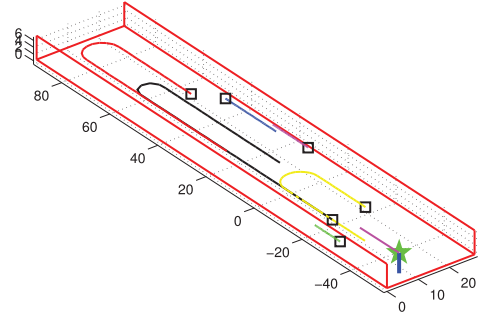


Fig. 10. Six users in the urban canyon moving over the duration of the 7 second simulation interval. Their positions at time $t = 0$ is marked using a \square -symbol.

report results for the algorithm proposed in Section V and two feedback strategies: (i) ‘full’: users feedback the measured virtual channel matrix \mathbf{Y} ($M \times L$ matrix; $L = 6$ for both systems) and (ii) ‘svd’: users feedback the 2 dominant left singular vectors of \mathbf{Y} , scaled by their corresponding singular values ($M \times 2$ matrix; one-third feedback overhead).

We consider 6 mobile users moving in the urban canyon at speeds of 20, 3, 15, 1.5, 2.1 and 10 meters per second (covering both vehicular and pedestrian settings). The height at which each mobile device is held is in the 1.3 – 1.4 m range. The basestation is mounted on a lamppost on the pavement (7 meters from a canyon wall), at a height of 6 meters. The basestation antenna array is tilted by about 7.5° in both the azimuth and elevation directions so that the boresight of the array points towards middle of the corresponding cell. This helps in more accurate spatial frequency estimation: since a change in direction near the boresight of the array results in larger changes in spatial frequencies than far away from the boresight, resolving paths is easier when the array points towards a direction in which we are likely to see more paths. We do not model blockage in these simulations, assuming that the LoS path and the three first order reflections are all available. Our goal is to estimate and track the $K = 4$ paths to all 6 users.

1) *Estimation Error*: Let $T = \{\omega_m : m = 1, \dots, K\}$ denote the true spatial frequencies and $P = \{\hat{\omega}_n : n = 1, \dots, \hat{K}\}$ denote the set of estimated spatial frequencies. When the base station uses one of the estimates in P , say $\hat{\omega}$, to form a beam, we do not realize the full $20 \log N_{t,1D}$ dB beamforming gain. A measure of the sub-optimality is the estimation error $|\omega - \hat{\omega}|_2$, which we normalize by the DFT spacing of $2\pi/N_{t,1D}$ to define the following error metric:

$$\Delta\omega(m) = \min_n |\omega_m - \hat{\omega}_n|_2 / (2\pi/N_{t,1D}). \quad (23)$$

When no true spatial frequency exists near an estimate $\hat{\omega}$, i.e., when $\hat{\omega}$ is a ‘phantom estimate’, we will quickly be able to discard it when we beamform in the direction of $\hat{\omega}$ and find that the mobile does not receive power commensurate to what it expects with the $20 \log N_{t,1D}$ beamforming gain.

We plot the Complementary Cumulative Distribution Function (CCDF) of estimation errors (23) for the two systems ($N_{t,1D} = 8, 32$) in Fig. 10 and the Probability Distribution Function (PDF) of the number of paths estimated \hat{K} (correct value is $K = 4$) in Fig. 10. We have set $P_{fa} = 10^{-3}$ to arrive at

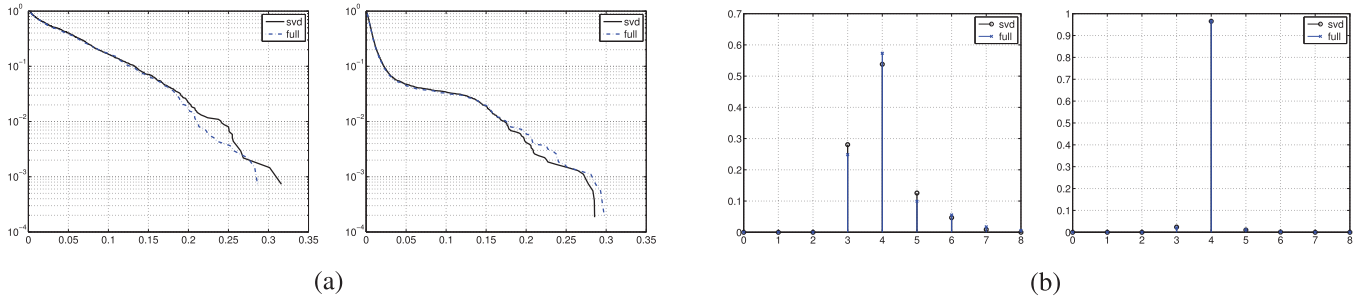


Fig. 11. Two feedback strategies considered: (i) $M \times 6$ matrix \mathbf{Y} ('full') and (ii) top two dominant singular vectors (one-third overhead) (a) CCDF of frequency estimation errors with new stopping criterion for $N_{t,1D} = 8$ (left) and $N_{t,1D} = 32$ (right) (b) PDF of # of estimated paths \bar{K} ($K = 4$) with new stopping criterion for $N_{t,1D} = 8$ (left) and $N_{t,1D} = 32$ (right).

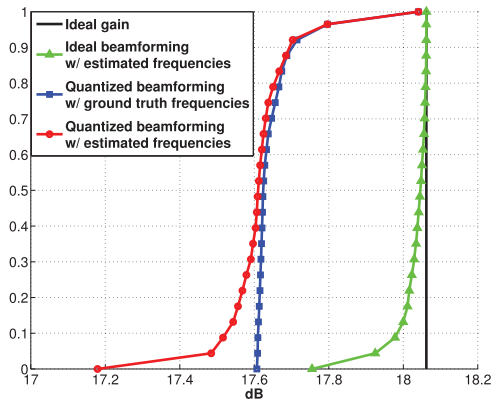


Fig. 12. CDF of beamforming gain achieved by an 8×8 antenna array for ideal and quantized beamforming techniques.

τ used in the stopping criterion (13). From Fig. 11, we see that feedback of dominant singular vectors is an efficient feedback strategy which performs just as well as feeding back the entire matrix \mathbf{Y} , while using only a third of uplink resources.

Next, in order to evaluate the effect of errors in spatial frequency estimation on beamforming performance, we simulate a simple scenario in which the transmitter beamforms towards the strongest estimated path. Fig. 12 shows the CDF of the achievable beamforming gain for an 8×8 array. While ideal beamforming requires adjustment of both gains and phases, suboptimal approaches for RF beamsteering with severely quantized phase-only control (four phases) have been studied in our earlier conference paper [25]. We see from Fig. 12 that if ideal beamforming were performed with our estimates, then the SNR loss is less than 0.3 dB. If four-phase control is used based on our estimates, then the SNR loss is less than 1 dB. The results for 32×32 arrays are entirely similar, and are therefore not plotted here.

Thus far, we have not said anything about channel frequency selectivity. Our proposed algorithm uses a small segment of the band to estimate the spatial channel, and the problem of channel dispersion is not addressed. However, we note that beamforming using a large array should reduce the effect of undesired paths, which simplifies the task of equalization. Fig. 13 shows the channel impulse responses for the 32×32 and 8×8 antenna arrays for a typical snapshot, when the transmitter beamforms towards the strongest estimated path. In our

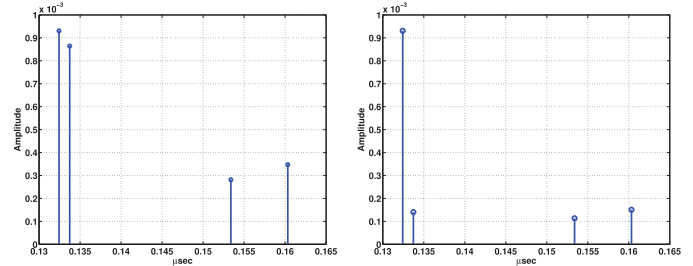


Fig. 13. Channel impulse response with quantized beamforming towards estimated strongest path for the 8×8 (left) and 32×32 (right) scenarios.

simulated setting, the LoS and ground reflection are close to each other in terms of both delays and angles of departure. We see that 8×8 array fails to resolve them, with both paths falling into the antenna's main lobe, while the 32×32 antenna array, which has smaller beamwidth (4° half power beamwidth), attenuates the undesired tap down to one-ninth of the desired path. Of course, it is possible to utilize the channel estimates far more intelligently, potentially with nulls directed both at strong undesired paths for the mobile of interest, and at the dominant paths for other nearby mobiles. The latter can be particularly useful for combating intra-cell interference when a base station face has multiple antenna arrays, each communicating with a different mobile.

IX. CONCLUSIONS

We have shown that it is possible to super-resolve mm wave spatial channels with a relatively small number of compressive measurements, in a manner that is compatible with coarse phase-only control and RF beamforming. This allows scaling to a very large number of antenna elements without relying on channel reciprocity. While our discussion of system design issues such as link budget and inter-cell beacon interference is tailored to outdoor 60 GHz picocellular networks, the basic approach is broadly applicable (e.g., to other bands, and to indoor environments). An important topic for future work is comprehensive experimental validation of our compressive approach, which is based on a simple channel model including only the dominant rays. At the network level, there are a host of design issues (e.g., see discussion in [3]). The compressive approach allows each base station to build up an inventory of viable paths to nearby mobiles, but there is a huge design space

to be explored on how base stations coordinate using this information to alleviate the effects of blockage (mobiles in urban environments can be routinely blocked by pedestrians, automobiles, trees and other obstacles), and to manage inter- and intra-cell interference. Optimization of arrays of subarrays in base station “faces” for communicating with multiple users, as well as for handling channel dispersion, presents interesting design challenges.

APPENDIX

A. Stopping Criterion

In this section we derive the CFAR-based stopping criterion (13). In essence, we wish to determine when the residual measurements, after subtracting out the contribution due to the paths estimated thus far, can be explained well enough by noise, up to a nominal false alarm probability of P_{fa} .

Under the noise-only hypothesis (i.e., assuming all existing paths have been detected and subtracted), we have

$$\mathbf{v}_l(\mathcal{P}_q) \approx \mathbf{z}_l \quad \mathbf{z}_l \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbb{I}_M) \quad l = 1, \dots, L$$

Recall that we use (9), maximized over an oversampled DFT grid, to detect a new frequency. For our stopping criterion (13), we consider the maximum over a DFT grid (this simplifies analysis and design), and stop if this is lower than a threshold, as follows:

$$\max_{\omega \in DFT} \frac{1}{\|\mathbf{x}(\omega)\|^2} \sum_{l=1}^L |\langle \mathbf{x}(\omega), \mathbf{v}_l(\mathcal{P}_q) \rangle|^2 \leq \tau \quad (24)$$

The false alarm rate is the probability that the threshold is exceeded under the noise-only hypothesis:

$$P_{fa} = \Pr \left\{ \max_{\omega \in DFT} \frac{1}{\|\mathbf{x}(\omega)\|^2} \sum_{l=1}^L |\langle \mathbf{x}(\omega), \mathbf{z}_l \rangle|^2 > \tau \right\} \quad (25)$$

The noise projections are i.i.d. $\mathcal{CN}(0, 1)$ across ω (due to the orthogonality of the DFT basis) and i.i.d. across measurement time indices l . Thus,

$$\nu_{l,\omega} := \langle \mathbf{x}(\omega), \mathbf{z}_l \rangle \sigma / \|\mathbf{x}(\omega)\| \sim \mathcal{CN}(0, 1)$$

and the sum of the squared magnitudes across projections is a Gamma random variable:

$$\mu_\omega := \sum_{l=1}^L |\nu_{l,\omega}|^2 \sim \text{Gamma}(L, 1)$$

We therefore obtain

$$\begin{aligned} P_{fa} &= P \left\{ \max_{\omega \in DFT} \mu_\omega > \tau \sigma^2 \right\} = 1 - P \{ \mu_{\omega_0} < \tau \sigma^2 \}^{N_{1D,t}^2} \\ &= 1 - (\gamma(L, \tau / \sigma^2) / \Gamma(L))^{N_{1D,t}^2} \end{aligned}$$

where $\gamma(L, x) = \int_0^x t^{L-1} e^{-t} dt$ is the incomplete gamma function, and $\Gamma(L) = \gamma(L, \infty)$ is the gamma function. For a given P_{fa} , the threshold τ is now given by

$$\tau = \sigma^2 \gamma^{-1} \left(L, \Gamma(L) (1 - P_{fa})^{1/N_{1D,t}^2} \right) \quad (26)$$

where $\gamma^{-1}(L, y)$ denotes the inverse of the incomplete gamma function with respect to the integral limit x .

B. Complexity Analysis

We go over the computations that are needed to maintain the estimate of one user’s channel over time and point out portions of the computation phase that can be reused across users.

The first step involves computing responses to each of the $R^2 N$ spatial frequency grids ($N = N_{1D,t}^2$ as the number of transmitter antennas):

$$\mathcal{X} = \{ \mathbf{x}(\omega) = \mathbf{A} \mathbf{x}_t(\omega) : \omega \in \Phi \}.$$

This can be efficiently computed using the Fast Fourier Transform in $O(MNR^2 \log(NR))$ time and can be shared across users. Furthermore, the set of responses \mathcal{X} can be stored at the basestation and reused across sounding rounds. Thus, the cost of computing \mathcal{X} is amortized over users in the network and across sounding rounds.

The next step involves finding the strongest path using (9) and the corresponding gains. This operation takes $O(MLNR^2)$ time. We then proceed to the refinement stage which takes $O(M(L+N))$ time ($O(MN)$ to compute $\partial \mathbf{x}(\omega) / \partial \omega_i$ and $O(ML)$ for gradient and Hessian computation).

When multiple paths are present, each path is refined again for every subsequently detected path. Therefore, the overall complexity when a new user joins the network is $O(K^2) \times$ refinement complexity $+ K \times$ detection complexity $= O(MK^2(L+N) + KMLNR^2)$. In arriving at this scaling, we use the fact that our algorithm is residue centric and therefore we keep track of $\{ \mathbf{v}_l(\mathcal{P}_q) \}$. As a result we can compute $\{ \mathbf{v}_l(\mathcal{P}_q \setminus \{ \hat{\omega}_k, \hat{h}_{k,l} \}) \}$ in $O(LM)$ time as opposed to the $O(qLM)$ time that would have been needed if this computation were to be done directly from $\{ \mathbf{y}_l \}$.

When we track the channel using estimates from prior rounds, we compute the gains $\{ \hat{h}_{k,l} \}$ of the existing spatial frequency estimates from prior round given by $\{ \hat{\omega}_k \}$ as $(X^H X)^{-1} X^H Y$ and this takes $O(KM(L+N))$ time. We then refine these estimates using our refinement algorithm in $O(KM(L+N))$ time.

Paths that were viable in previous sounding rounds may be blocked in the current sounding round. To prune out such paths, we test whether the residual measurements after the deletion of the under question path satisfy our stopping criterion. This check takes $O(KM(L+N) + MLN)$ time for each path and path pruning complexity scales as $O(K^2 M(L+N) + KMLN)$. Checking whether a new path exists (stopping criterion) also costs $O(MLN)$ time. Therefore, the overall tracking complexity scales as $O(KM(KL + KN + LN)) = O(KMN(L+K))$ (since $N > K$) time.

REFERENCES

- [1] “Cisco visual networking index: Forecast, and methodology, 2011–2016,” *CISCO White paper*, pp. 2011–2016, 2012.
- [2] D. Ramasamy, S. Venkateswaran, and U. Madhow, “Compressive tracking with 1000-element arrays: A framework for multi-Gbps mm wave cellular downlinks,” *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2012.
- [3] Y. Zhu, Z. Zhang, Z. Marzi, C. Nelson, U. Madhow, B. Y. Zhao, and H. Zheng, “Demystifying 60ghz outdoor picocells,” *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, ser. MobiCom ’14.

- [4] A. Valdes-Garcia, S. Nicolson, J.-W. Lai, A. Natarajan, P.-Y. Chen, S. Reynolds, J.-H. Zhan, D. Kam, D. Liu, and B. Floyd, "A fully integrated 16-element phased-array transmitter in sige BiCMOS for 60-ghz communications," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2757–2773, Dec. 2010.
- [5] E. Cohen, M. Ruberto, M. Cohen, O. Degani, S. Ravid, and D. Ritter, "A CMOS bidirectional 32-element phased-array transceiver at 60ghz with LTCC antenna," *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2012, pp. 439–442.
- [6] S. Rajagopal, S. Abu-Surra, Z. Pi, and F. Khan, "Antenna array design for multi-Gbps mmwave mobile broadband communication," *Proc. IEEE Global Telecommunications Conf. (GLOBECOM'11)*, Dec. 2011.
- [7] S. Akoum, O. El Ayach, and R. Heath, "Coverage and capacity in mmwave cellular systems," *Conf. Rec. 46th Asilomar Conf. Signals, Syst., Comput. (ASILOMAR)*, Nov. 2012.
- [8] T. Rappaport, Y. Qiao, J. Tamir, J. Murdock, and E. Ben-Dor, "Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems (invited paper)," *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2012, pp. 151–154.
- [9] J. Murdock, E. Ben-Dor, Y. Qiao, J. Tamir, and T. Rappaport, "A 38 GHz cellular outage study for an urban outdoor campus environment," *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 3085–3090.
- [10] A. Ghosh, T. Thomas, M. Cudak, R. Ratasuk, P. Moorut, F. Vook, T. Rappaport, G. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014.
- [11] D. Berraki, S. Armour, and A. Nix, "Application of compressive sensing in sparse spatial channel recovery for beamforming in mmwave outdoor systems," *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014.
- [12] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.
- [13] S. Hur, T. Kim, D. Love, J. Krogmeier, T. Thomas, and A. Ghosh, "Multilevel millimeter wave beamforming for wireless backhaul," *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011, pp. 253–257.
- [14] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.
- [15] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [16] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, "Single-sided adaptive estimation of multi-path millimeter wave channels," *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 125–129.
- [17] T. Bogale and L. B. Le, "Beamforming for multiuser massive MIMO systems: Digital versus hybrid analog-digital," *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 4066–4071.
- [18] C. Sheldon, M. Seo, E. Torkildson, M. Rodwell, and U. Madhow, "Four-channel spatial multiplexing over a millimeter-wave line-of-sight link," *IEEE MTT-S Int. Microw. Symp. Dig. (MTT'09)*, 2009, pp. 389–392.
- [19] G. Wang and G. Ascheid, "Joint pre/post-processing design for large millimeter wave hybrid spatial processing systems," *Proc. 20th Eur. Wireless Conf.: Proc. Eur. Wireless'14*, May 2014.
- [20] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014.
- [21] O. Ayach, R. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012.
- [22] O. El Ayach, R. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 100–104.
- [23] H. Zhang and U. Madhow, "Statistical modeling of fading and diversity for outdoor 60 GHz channels," *Proc. ACM Int. Workshop mmWave Commun.: Circuits Netw.*, 2010.
- [24] H. Zhang, S. Venkateswaran, and U. Madhow, "Channel modeling and MIMO capacity for outdoor millimeter wave links," *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*.
- [25] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive adaptation of large steerable arrays," *Proc. Inf. Theory Applications Workshop (ITA)*, Feb. 2012.
- [26] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive estimation in AWGN: General observations and a case study," *Conf. Rec. 46th Asilomar Conf. Signals, Syst., Comput. (ASILOMAR)*, Nov. 2012.
- [27] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive parameter estimation in AWGN," *IEEE Trans. Signal Processing*, vol. 62, no. 8, pp. 2012–2027, Apr. 2014.
- [28] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Frequency estimation for a mixture of sinusoids: A near-optimal sequential approach," *Proc. 3rd IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2015.
- [29] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," *arXiv preprint arXiv:1509.01942*, 2015.
- [30] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Construct. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.
- [31] D. Achlioptas, *Database-Friendly Random Projections*, in *Proc. 20th ACM SIGMOD-SIGACT-SIGART symp. Principles of database syst.*, ACM, 2001, pp. 274–281.
- [32] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New York, NY, USA: Wiley, 2007.
- [33] S. Basu and Y. Bresler, "A global lower bound on parameter estimation error with periodic distortion functions," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 1145–1150, May 2000.



Zhinus Marzi is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering at the University of California Santa Barbara (UCSB). She received her B.Sc. and M.Sc. in electrical engineering from the Iran University of Science and Technology and Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. She is interested in wireless communication and signal processing and her current focus is on millimeter wave communication. She has been a summer intern at Qualcomm Inc., San Diego, in 2015.



Dinesh Ramasamy received his Bachelor's degree in electrical engineering from the Indian Institute of Technology, Madras, in 2009 and his Ph.D. degree from the University of California, Santa Barbara (UCSB) in 2014. He is interested in machine learning, millimeter wave communication and inference from large scale networks.



Upamanyu Madhow received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1985, and his Ph.D. degree in electrical engineering from the University of Illinois, Urbana-Champaign in 1990. He is Professor of Electrical and Computer Engineering at the University of California, Santa Barbara. His research interests broadly span communications, signal processing and networking, with current emphasis on millimeter wave communication, and on distributed and bio-inspired approaches to networking and inference. He has worked as a research scientist at Bell Communications Research, Morristown, NJ, and as a faculty at the University of Illinois, Urbana-Champaign. Dr. Madhow is a recipient of the 1996 NSF CAREER award, and co-recipient of the 2012 IEEE Marconi prize paper award in wireless communications. He has served as Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON INFORMATION THEORY, and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He is the author of two textbooks published by Cambridge University Press, *Fundamentals of Digital Communication* (2008) and *Introduction to Communication Systems* (2014).