Non-uniform Directional Dictionary-Based Limited Feedback for Massive MIMO Systems

Panos N. Alevizos*, Xiao Fu†, Nicholas Sidiropoulos†, Ye Yang+, and Aggelos Bletsas*

*School of Electrical and Computer Engineering, Technical University of Crete
†Dept. of Electrical and Computer Engineering, University of Minnesota
+Physical Layer & RRM IC Algorithm Dept., WN Huawei Co., Ltd.

***Abstract*—This work proposes a new limited feedback channel estimation framework. The proposed approach exploits a sparse representation of the double directional wireless channel model involving an overcomplete dictionary that accounts for the antenna directivity patterns at both base station (BS) and user equipment (UE). Under this sparse representation, a computationally efficient limited feedback algorithm that is based on single-bit compressive sensing is proposed to effectively estimate the downlink channel. The algorithm is lightweight in terms of computation, and suitable for real-time implementation in practical systems. More importantly, under our design, using a small number of feedback bits, very satisfactory channel estimation accuracy is achieved even when the number of BS antennas is very large, which makes the proposed scheme ideal for massive MIMO 5G cellular networks. Judiciously designed simulations reveal that the proposed algorithm outperforms a number of popular feedback schemes in terms of beamforming gain for subsequent downlink transmission, and reduces feedback overhead substantially when the BS has a large number of antennas.**

I. Introduction

Future 5G wireless cellular systems are expected to use a very large number of antennas at the base station (BS). This type of *massive* multiple-input multiple-output (MIMO) multi-antenna architectures has attracted a lot of attention in recent years by both academia and industry. Massive MIMO can achieve very high throughputs [1], [2] through spatial multiplexing and array gains [3]. Massive MIMO systems also have the advantage of being energy-efficient since every antenna only operates at a low-energy level [4].

To realize high-throughput massive MIMO systems, acquiring accurate channel state information (CSI) at the BS in a timely fashion is crucial [5]. In particular, achieving high-accuracy downlink channel estimation at the BS using only few feedback bits from the user equipment (UE) is a challenge, especially in massive MIMO systems. In frequency-division-duplex (FDD) systems, where channel reciprocity does not hold and thus the BS cannot acquire downlink channel information from uplink training sequences, the feedback overhead may scale proportionally to the number of BS antennas [6]. Note that this feedback overhead problem does not only exist in FDD; even in the time-division-duplex (TDD) mode, relying only on channel reciprocity is not accurate enough, since reciprocity is itself an approximation, and the uplink measurements at the BS cannot capture the downlink interference from neighboring cells [7], [8]. Therefore, studying limited feedback based channel estimation is well-motivated for massive MIMO.

Many limited feedback mechanisms that aim at acquiring accurate CSI at the BS with reduced communication overhead can be found in the literature [9]. Many of these methods utilize a vector quantization (VQ) codebook that is known to both the BS and the UE. After estimating the instantaneous downlink CSI at the UE, the UE sends through a limited feedback channel the index of the codeword that best matches the estimated channel, in the sense of minimizing the outage probability [10], or maximizing the normalized average signal-to-noise-ratio (SNR) [11], [12]. VQ-based feedback schemes are usually combined with transmit beamforming.

Many limited feedback approaches in MIMO systems consider a Rayleigh fading channel model [6], [11], [12], [13], [14]. Under this channel model, the number of VQ feedback bits required to guarantee reasonable performance is linear in the number of transmit antennas at the BS [15] – which is costly in the case of massive MIMO systems. In addition, since most of the heavy computations are carried out at the UE, these approaches assume that the UE has adequate computational, memory, and power resources for codebook storage and search (at a minimum), which should be avoided if possible.

Contributions: In this work, an alternative approach to the existing limited feedback methods is considered. Specifically, we aim at designing a feedback scheme that meets the demands of massive MIMO. We employ the double directional MIMO channel model (DD model) [16] instead of the Rayleigh fading model. The DD channel model parameterizes each channel path using angle of departure (AoD) at BS, small- and large-scale propagation coefficients, and angle of arrival (AoA) at UE – a parametrization that is well-accepted and advocated by 3GPP [17], [18]. Our idea is to make use of a 'virtual sparse representation' of the downlink channel under the double directional MIMO model [16]. Specifically, we quantize the angular space of AoA and AoD so that overcomplete dictionaries are constructed which contain steering vectors approximating those of the true arrival and departure angles. Unlike previous work that directly applies compressive sensing techniques at the UE to recover the channel as in [16], we propose a scheme in which the UE judiciously compresses the received measurements and sends back only the signs of the compressed measurements to the BS. Upon receiving these

sign bits, the BS estimates the channel using a variation of single-bit compressed sensing (CS) [19], [20], [21], [22]. For the DD channel model, channel estimation with CS techniques is considered in [23], [24], [25]

Our new framework features several desirable properties. First, the computational burden is shifted to the BS side – the UE only carries out matrix-vector multiplications and takes signs. This is sharply different from most limited feedback schemes in the literature, where the UE does the 'heavy lifting' [5], [9], [14]. Second, by exploiting sparsity of the virtual channel model and leveraging sophisticated single-bit compressive sensing algorithms, the proposed approach significantly outperforms VQ-based approaches in terms of estimation accuracy for a given bit budget. Third, the single-bit compressive sensing technique that we advocate here has a simple closed-form solution, which can be easily implemented and performed in real time – thus relieving the computational burden on the BS as well.

In a addition to designing the new limited feedback scheme, a new angle dictionary construction methodology is presented to enhance performance, based on a companding quantization technique [26]. The idea is to create dictionaries that concentrate the angle density in a non-uniform manner along the angles where directivity patterns attain higher values, achieving better angle granularity over those regions, which eventually translates to better channel estimation performance. Baseline 3GPP and dipole antenna directivity patterns are considered and contrasted with uniform quantization to showcase this important point. Judicious simulations reveal that accounting for antenna directivity patterns in the channel model offers increased beamforming gain. It is also shown that the proposed method outperforms baseline least-squares (LS) scalar and vector quantization in terms of beamforming gain while requiring less feedback overhead.

Related Work: In [27], one-bit feedback was considered in the context of power spectrum sensing, where the feedback bits correspond to signs of randomly compressed power measurements. In [28], instantaneous SNR measurements are compressed to one bit and then fed back to the BS for beamformer design and channel covariance tracking. The approach in [28] cannot estimate the instantaneous channel, e.g., phase is already lost due to the quadratic measurement, even before taking the sign. The method reported in [29] models the temporal channel evolution using a vector auto-regressive process, and uses the sign of the Kalman filtering innovation sequence as feedback to track the channel. The method in [29] requires strong temporal channel correlation, and it does not exploit the powerful DD channel model. In addition, all the above involve computationally demanding iterative optimization algorithms – in contrast with the method proposed herein, which is closed-form and hence very lightweight in terms of computational resources. We also note that the proposed method aims at recovering the *instantaneous channel* without assuming any temporal correlation model.

II. System Model

We consider a FDD cellular system consisting of a BS serving K active UE terminals, where the downlink channel is estimated at the BS through feedback from each UE. For brevity of exposition, we focus on a single UE. Generalization to multiple users is straightforward, as the channel estimation process can be performed separately for each UE. The BS is equipped with M_{T} antennas, and the UE is equipped with M_{R} antennas.

The channel is assumed static for a coherence block of $U_{\mathrm{c}} \approx B_{\mathrm{c}}T_{\mathrm{c}}$ complex OFDM symbols, where B_{c} is the coherence bandwidth (in Hz), T_{c} is the coherence time (in seconds). In downlink transmission, the BS has to acquire CSI through feedback from the active UE terminals, and then design the transmit signals accordingly. At the training phase, the BS employs N_{tr} training symbols for channel estimation. The narrowband (over time-frequency) discrete baseband-equivalent model over a period of N_{tr} training symbols is given by

$$\mathbf{y}_n = \mathbf{H}\,\mathbf{s}_n + \mathbf{n}_n, \quad n = 1, 2, \ldots, N_{\mathrm{tr}}, \tag{1}$$

where $\mathbf{s}_n \in \mathbb{C}^{M_{\mathrm{T}}}$ is the nth transmitted training symbol, $\mathbf{y}_n \in \mathbb{C}^{M_{\mathrm{R}}}$ is the received vector, $\mathbf{H} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}}}$ denotes the complex baseband equivalent channel matrix, and $\mathbf{n}_n \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{M_{\mathrm{R}}})$ is additive white Gaussian noise at the receiver. All quantities in the right hand-side of (1) are independent of each other. A total power constraint per channel use is considered, i.e., $\mathbb{E}[\mathbf{s}_n\mathbf{s}_n^{\mathsf{H}}] = \frac{P_{\mathrm{T}}}{M_{\mathrm{T}}}\mathbf{I}_{M_{\mathrm{T}}}$, for all n, where P_{T} denotes the average total transmit power. The transmit signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} \triangleq \frac{P_{\mathrm{T}}}{\sigma^2}$.

To estimate $\mathbf{H}$, we can use linear least-squares (LS) [6], or, if the channel covariance is known, the linear minimum mean-squared error (LMMSE) approach [5]. These linear approaches need more than $M_{\mathrm{T}}M_{\mathrm{R}}$ training symbols to establish identifiability of the channel (to 'over-determine' the problem) – and this is rather costly in massive MIMO scenarios. In addition, if one estimates the channel at the UE, the UE needs to be equipped with significant computational resources, which is unpractical – also considering that channel estimation is a 'background' job, not the main function of the UE.

A more practical approach to the problem of downlink channel acquisition at the BS of massive MIMO systems would be to shift the computational burden to the BS, relying on relatively lightweight computations at the UE, and assuming that only low-rate feedback is available as well. The motivation for this is clear: the BS is connected to the communication backbone, plugged to the power grid, and may even have access to cloud computing – thus is far more capable of performing intensive computations. The challenge of course is how to control the feedback overhead – without a limitation on feedback rate, the UE can of course simply relay the signals it receives back to the BS, but such an approach is clearly wasteful and impractical. The name of the game is how to achieve accurate channel estimation with low feedback overhead: our goal is to estimate $\mathbf{H}$ using just a few feedback bits.

Towards this end, our starting idea is to employ a finite scatterer (multipath or double directional) channel model comprising L paths [16], which can be parameterized using a virtual sparse representation. This sparse representation will lead to a feedback scheme that is rather parsimonious in terms of both overhead and computational complexity.

A. Double Directional Model and Sparse Representation

The narrowband downlink channel matrix $\mathbf{H}$ can be parametrized as follows:

$$\mathbf{H} = \sqrt{\frac{M_{\mathrm{T}} M_{\mathrm{R}}}{L}} \sum_{l=1}^{L} \alpha_l \, \mathsf{c}_{\mathrm{T}}(\phi_l') \, \mathsf{c}_{\mathrm{R}}(\phi_l) \, \mathbf{a}_{\mathrm{R}}(\phi_l) \, \mathbf{a}_{\mathrm{T}}^{\mathsf{H}}(\phi_l'), \quad (2)$$

where α_l is the complex gain of the l-th path incorporating path-loss as well as small- and large-scale fading effects, and ϕ_l, ϕ_l' are the azimuth angle of arrival (AoA) and angle of departure (AoD) of the lth path, respectively. Functions $\mathsf{c}_{\mathrm{T}}(\cdot)$ and $\mathsf{c}_{\mathrm{R}}(\cdot)$ are the transmit and received antenna directivity patterns, respectively (all transmit antenna elements have the same directivity pattern, and the same holds for receive antenna elements). Examples of transmit and receive antenna patterns are the uniform directivity pattern over a sector $[\phi_{\mathrm{T}}^{\mathrm{l}}, \phi_{\mathrm{T}}^{\mathrm{u}}]$, given by

$$\mathsf{c}_{\mathrm{T}}(\phi) = \begin{cases} 1, & \phi \in [\phi_{\mathrm{T}}^{\mathrm{l}}, \phi_{\mathrm{T}}^{\mathrm{u}}] \\ 0, & \text{otherwise}, \end{cases} \quad (3)$$

and likewise for $\mathsf{c}_{\mathrm{R}}(\phi)$. The so-called dipole directivity pattern (on y-axis) [30] is also often considered, and has the following form:

$$\mathsf{c}_{\mathrm{T}}(\phi) = \begin{cases} \mathtt{G}_{\mathrm{A}} \cos^3(\phi), & \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \\ -\mathtt{G}_{\mathrm{A}} \cos^3(\phi), & \phi \in \left[-\pi, -\frac{\pi}{2}\right) \cup \left[\frac{\pi}{2}, \pi\right), \end{cases} \quad (4)$$

where $\mathtt{G}_{\mathrm{A}}$ is the maximum directional gain of the antenna. Another baseline directivity pattern is advocated by 3GPP [31], [32]:

$$\mathsf{c}_{\mathrm{T}}(\phi) = \mathtt{G}_{\mathrm{B}} \, 10^{\max\left\{-0.6\left(\frac{\phi}{\phi_{3\mathrm{dB}}}\right)^2, -\frac{\mathtt{A}_m}{20}\right\}}, \quad (5)$$

with $\phi \in [-\pi, \pi)$, where $\mathtt{G}_{\mathrm{B}}$ is the maximum directional gain of the radiation element, $\mathtt{A}_m$ is the front-to-back ratio, and $\phi_{3\mathrm{dB}}$ is the 3dB-beamwidth. In (2), $\mathbf{a}_{\mathrm{R}}(\phi_l)$ and $\mathbf{a}_{\mathrm{T}}^{\mathsf{H}}(\phi_l')$ denote receive and transmit array steering vectors of the lth path, respectively. The specific form of the steering vectors depends on the array geometry; e.g., when a uniform linear arrays (ULA) (w.r.t. y axis) is employed, the steering vector at the BS is given by

$$\mathbf{a}_{\mathrm{T}}(\phi) = \sqrt{\frac{1}{M_{\mathrm{T}}}} \left[1 \;\; \mathsf{e}^{-\mathsf{j}\frac{2\pi d_y}{\lambda}\sin(\phi)} \;\; \ldots \;\; \mathsf{e}^{-\mathsf{j}\frac{2\pi d_y (M_{\mathrm{T}}-1)}{\lambda}\sin(\phi)}\right]^{\top},$$

where λ is the propagation wavelength, d_y is the distance between the antenna elements in y axis, usually $d_y = \lambda/2$.

For any antenna array structure and directivity pattern, let

$$\mathbf{A}_{\mathrm{R}} \triangleq [\mathsf{c}_{\mathrm{R}}(\phi_1)\mathbf{a}_{\mathrm{R}}(\phi_1) \;\; \ldots \;\; \mathsf{c}_{\mathrm{R}}(\phi_L)\mathbf{a}_{\mathrm{R}}(\phi_L)], \quad (6)$$

$$\mathbf{A}_{\mathrm{T}} \triangleq [\mathsf{c}_{\mathrm{T}}(\phi_1')\mathbf{a}_{\mathrm{T}}(\phi_1') \;\; \ldots \;\; \mathsf{c}_{\mathrm{T}}(\phi_L')\mathbf{a}_{\mathrm{T}}(\phi_L')], \quad (7)$$

$$\boldsymbol{\alpha} \triangleq \sqrt{\frac{M_{\mathrm{T}} M_{\mathrm{R}}}{L}} [\alpha_1 \;\; \alpha_2 \;\; \ldots \alpha_L]^{\top}, \quad (8)$$

then the channel in (2) can be written more compactly as

$$\mathbf{H} = \mathbf{A}_{\mathrm{R}} \, \mathsf{diag}(\boldsymbol{\alpha}) \, \mathbf{A}_{\mathrm{T}}^{\mathsf{H}}. \quad (9)$$

Starting from the model in (9), one can come up with a sparse representation of the channel [16]. First, the angle space of AoA and AoD is quantized by discretizing the angular space. Let us denote these dictionaries $\mathcal{P}_{\mathrm{T}}$ and $\mathcal{P}_{\mathrm{R}}$ for AoDs and AoAs, respectively. Dictionary $\mathcal{P}_{\mathrm{T}}$ contains G_{T} dictionary members, while $\mathcal{P}_{\mathrm{R}}$ contains G_{R} dictionary members. One simple way of constructing these dictionaries is to use a uniform grid of phases in an angular sector $[a, b) \subseteq [-\pi, \pi)$. In that case, $\mathcal{P}_{\mathrm{R}} = \left\{a + \frac{j\,(b-a)}{G_{\mathrm{R}}+1}\right\}_{j=1}^{G_{\mathrm{R}}}$ and $\mathcal{P}_{\mathrm{T}} = \left\{a + \frac{j\,(b-a)}{G_{\mathrm{T}}+1}\right\}_{j=1}^{G_{\mathrm{T}}}$.

Let $\widetilde{\mathbf{A}}_{\mathrm{R}} = \{\mathsf{c}_{\mathrm{R}}(\phi)\,\mathbf{a}_{\mathrm{R}}(\phi) : \phi \in \mathcal{P}_{\mathrm{R}}\} \in \mathbb{C}^{M_{\mathrm{R}} \times G_{\mathrm{R}}}$ and $\widetilde{\mathbf{A}}_{\mathrm{T}} = \{\mathsf{c}_{\mathrm{T}}(\phi)\,\mathbf{a}_{\mathrm{T}}(\phi) : \phi \in \mathcal{P}_{\mathrm{T}}\} \in \mathbb{C}^{M_{\mathrm{T}} \times G_{\mathrm{T}}}$ be dictionary matrices, associated with angle dictionaries $\mathcal{P}_{\mathrm{T}}$ and $\mathcal{P}_{\mathrm{R}}$. Up to some quantization errors, the channel matrix in (9) can be written in terms of basis matrices $\widetilde{\mathbf{A}}_{\mathrm{R}}$ and $\widetilde{\mathbf{A}}_{\mathrm{T}}$ as

$$\mathbf{H} \approx \widetilde{\mathbf{A}}_{\mathrm{R}} \mathbf{G} \, \widetilde{\mathbf{A}}_{\mathrm{T}}^{\mathsf{H}}, \quad (10)$$

where $\mathbf{G} \in \mathbb{C}^{G_{\mathrm{R}} \times G_{\mathrm{T}}}$ is the interaction matrix. Note that $\mathbf{G}$ *is not* diagonal, because in principle any AoA can be associated with any AoD. The (j,k)th element in $\mathbf{G}$ is associated with the jth and kth columns in $\widetilde{\mathbf{A}}_{\mathrm{R}}$ and $\widetilde{\mathbf{A}}_{\mathrm{T}}$, respectively – $[\mathbf{G}]_{j,k} \neq 0$ means that a propagation path associated with the kth angle in $\mathcal{P}_{\mathrm{T}}$ and the jth angle in $\mathcal{P}_{\mathrm{R}}$ is active. In practice, the number of (prominent) active paths is typically very small compared to the number of elements of $\mathbf{G}$ (i.e., $G_{\mathrm{T}} G_{\mathrm{R}}$). Thus, the matrix $\mathbf{G}$ is in most cases very sparse.

Setting $\mathbf{Y} = [\mathbf{y}_1 \; \mathbf{y}_2 \; \ldots \; \mathbf{y}_{N_{\mathrm{tr}}}]$, $\mathbf{S} = [\mathbf{s}_1 \; \mathbf{s}_2 \; \ldots \; \mathbf{s}_{N_{\mathrm{tr}}}]$, $\mathbf{N} = [\mathbf{n}_1 \; \mathbf{n}_2 \; \ldots \; \mathbf{n}_{N_{\mathrm{tr}}}]$, and using the channel approximation in (10), the baseband-equivalent model in (1) can be written in a compact matrix form as

$$\mathbf{Y} = \widetilde{\mathbf{A}}_{\mathrm{R}} \mathbf{G} \, \widetilde{\mathbf{A}}_{\mathrm{T}}^{\mathsf{H}} \, \mathbf{S} + \mathbf{N}. \quad (11)$$

Vectorizing Eq. (11) and using the property $\mathsf{vec}(\mathbf{ABC}) = (\mathbf{C}^{\top} \otimes \mathbf{A})\,\mathsf{vec}(\mathbf{B})$, where $\otimes$ denotes the Kronecker product, the baseband received signal is given by

$$\mathbf{y} = \left((\mathbf{S}^{\top} \widetilde{\mathbf{A}}_{\mathrm{T}}^{*}) \otimes \widetilde{\mathbf{A}}_{\mathrm{R}}\right) \mathbf{g} + \mathbf{n} = \mathbf{Q}\mathbf{g} + \mathbf{n}, \quad (12)$$

where $\mathbf{y} \triangleq \mathsf{vec}(\mathbf{Y}) \in \mathbb{C}^{M_{\mathrm{R}} N_{\mathrm{tr}}}$, $\mathbf{g} \triangleq \mathsf{vec}(\mathbf{G}) \in \mathbb{C}^{G_{\mathrm{T}} G_{\mathrm{R}}}$, $\mathbf{n} \triangleq \mathsf{vec}(\mathbf{N}) \in \mathbb{C}^{M_{\mathrm{R}} N_{\mathrm{tr}}}$, and $\mathbf{Q} \triangleq (\mathbf{S}^{\top} \widetilde{\mathbf{A}}_{\mathrm{T}}^{*}) \otimes \widetilde{\mathbf{A}}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{R}} N_{\mathrm{tr}} \times G_{\mathrm{T}} G_{\mathrm{R}}}$. We define $G \triangleq G_{\mathrm{T}} G_{\mathrm{R}}$, standing for the joint dictionary size. Now, our task amounts to properly compressing $\mathbf{y}$ and sending a few feedback bits to the BS which can enable simple and effective estimation of $\mathbf{g}$ – since once $\mathbf{g}$ is recovered, the channel $\mathbf{H}$ can be approximately recovered from (10).

III. Construction of Angle Dictionaries Accounting Antenna Directivity Patterns

Before introducing the feedback scheme and algorithm that recovers the downlink channel at the BS, let us consider the

practical problem of quantizing the angular space. Prior works that employ the sparse representation in (9) use uniformly discretized angles as dictionaries [16], [24]. However, a more appealing angle dictionary should take into consideration the antenna directivity patterns, since the channel itself naturally reflects these directivity patterns – e.g., if a receive antenna has a spatial null, then it is impossible to receive a path coming from the null direction, and likewise for transmit nulls. A more refined version of this idea is to pack more angles around the peaks of the antenna directivity pattern, because the dominant paths will likely fall in those regions, and this is where we need higher angular resolution. Denser discretization within high-antenna-power regions can reduce quantization errors more effectively compared to a naive uniform quantization that ignores the directivity pattern. In this section, we propose a simple and easily implementable angle quantization technique that is based on the above rationale. As will be seen in the simulations, considering the directivity patterns substantially enhances the performance of channel estimation – especially in the low SNR regime.

To explain our approach, let $\mathsf{q} : [a,b) \longrightarrow \mathbb{R}^+$ be a given antenna directivity pattern function, which is assumed continuous over $\phi \in [a,b)$ and suppose that we want to represent it using N quantization points; see Fig. 1 for the 3GPP directivity pattern. We define the cumulative function of q, given by $\mathsf{G}(\phi) \triangleq \int_a^{\phi} \mathsf{q}(x)\mathrm{d}x$. As the range space of function q takes positive values, its continuity implies that G is monotone increasing. Thus, the following set

$$\mathcal{C}_{\mathsf{q}} \triangleq \left\{ \mathsf{G}(a) + \frac{n(\mathsf{G}(b)-\mathsf{G}(a))}{N+1}, \right\}_{n=1}^{N}, \tag{13}$$

partitions the range of G in $N+1$ intervals of equal size. By the definition of G, the set in (13) partitions function q in $N+1$ equal area intervals. Having the elements of set $\mathcal{C}_{\mathsf{q}}$, we can find the phases at which $\mathsf{q}(\phi)$ is partitioned in $N+1$ equal area intervals – which means that we achieve our goal of putting denser grids in the angular region where the q function has higher intensity. These phases can be found as

$$\mathcal{F}_{\mathsf{q}} \triangleq \left\{ \mathsf{G}^{-1}(y) \right\}_{y \in \mathcal{C}_{\mathsf{q}}}, \tag{14}$$

where $\mathsf{G}^{-1} : [\mathsf{G}(a), \mathsf{G}(b)) \longrightarrow [a,b)$ is the inverse (with respect to composition) function of G. Observe that G^{-1} is continuous, monotone increasing function since G is as well continuous and monotone increasing. The discrete set $\mathcal{F}_{\mathsf{q}}$ is a subset of $[a,b)$ and concentrates more elements at points where function q has larger values.

Following the above principles, let us further exemplify the procedure of constructing the angle dictionaries using the function q of the 3GPP antenna directivity pattern that is given in Eq. (5). As the most general case, we assume $a \leq -\phi_{3\mathrm{dB}}\sqrt{\frac{\mathtt{A}_{\mathrm{m}}}{12}}$ and $b \geq \phi_{3\mathrm{dB}}\sqrt{\frac{\mathtt{A}_{\mathrm{m}}}{12}}$. The domain of q can be partitioned into 3 disjoint intervals as

$$[a,b) = [a,-\phi_0) \cup [-\phi_0,\phi_0) \cup [\phi_0,b), \tag{15}$$

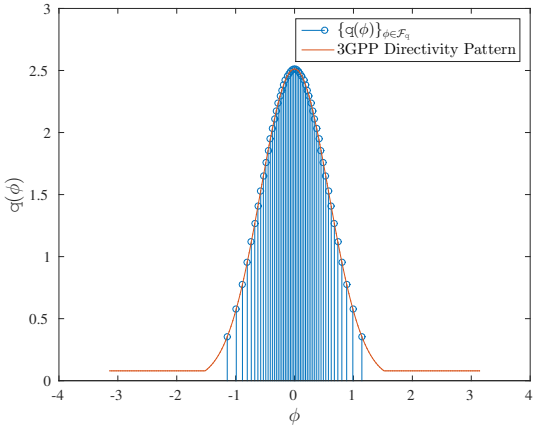

Fig. 1. 3GPP directivity pattern along with function q applied on the proposed dictionary using $a = -\pi$ and $b = \pi$. Note the proposed angle dictionaries contain more elements on the domain of q which obtains larger values.

with $\phi_0 \triangleq \phi_{3\mathrm{dB}}\sqrt{\frac{\mathtt{A}_{\mathrm{m}}}{12}}$. Applying the definition of cumulative function $\mathsf{G}(\phi) = \int_a^{\phi} \mathsf{q}(x)\mathrm{d}x$ and using its continuity, we obtain

$$\mathsf{G}(\phi) = \begin{cases} \mathtt{G}_{\mathrm{B}}(\phi - a)10^{-\frac{\mathtt{A}_m}{20}}, & \phi \in [a,-\phi_0), \\ \mathsf{G}(-\phi_0) + \mathtt{G}_{\mathrm{B}}\sqrt{\frac{\pi\,\phi_{3\mathrm{dB}}^2}{\ln(10)\,2.4}}\cdot \\ \left(\mathsf{erf}\left(\sqrt{\frac{\ln(10)\,0.6\,\mathtt{A}_m}{12}}\right) + \right. \\ \left. \mathsf{sign}(\phi)\mathsf{erf}\left(\sqrt{\frac{\ln(10)\,0.6}{\phi_{3\mathrm{dB}}^2}}|\phi|\right)\right), & \phi \in [-\phi_0,\phi_0), \\ \mathsf{G}(\phi_0) + \mathtt{G}_{\mathrm{B}}\left(\phi - \phi_0\right)10^{-\frac{\mathtt{A}_m}{20}}, & \phi \in [\phi_0,b), \end{cases} \tag{16}$$

where we exploit the fact that $\mathsf{erf}(x)\frac{\sqrt{\pi}}{2} = \int_0^x \mathrm{e}^{-t^2}\mathrm{d}t$ and the fact that $10^x = \mathrm{e}^{\ln(10)x}$. Upon defining $y^- \triangleq \mathsf{G}(-\phi_0)$, $y_0 \triangleq \mathsf{G}(0)$, and $y^+ \triangleq \mathsf{G}(\phi_0)$, the inverse of $\mathsf{G}(\cdot)$ can be calculated using Eq. (16) in closed form as

$$\mathsf{G}^{-1}(y) = \begin{cases} \frac{y}{\mathtt{G}_{\mathrm{B}}}10^{\frac{\mathtt{A}_m}{20}} + a, & y \in [0,y^-), \\ -\frac{\mathsf{erf}^{-1}\left(\frac{2\sqrt{\ln(10)\,0.6}}{\phi_{3\mathrm{dB}}\sqrt{\pi}}\frac{(y_0-y)}{\mathtt{G}_{\mathrm{B}}}\right)}{\frac{\sqrt{\ln(10)\,0.6}}{\phi_{3\mathrm{dB}}}}, & y \in [y^-,y_0), \\ \frac{\mathsf{erf}^{-1}\left(\frac{2\sqrt{\ln(10)\,0.6}}{\phi_{3\mathrm{dB}}\sqrt{\pi}}\frac{(y-y_0)}{\mathtt{G}_{\mathrm{B}}}\right)}{\frac{\sqrt{\ln(10)\,0.6}}{\phi_{3\mathrm{dB}}}}, & y \in [y_0,y^+), \\ \phi_0 + \frac{y-y^+}{\mathtt{G}_{\mathrm{B}}}10^{\frac{\mathtt{A}_m}{20}}, & y \in [y^+,\mathsf{G}(b)), \end{cases} \tag{17}$$

where $\mathsf{erf}^{-1}(\cdot)$ is the inverse (with respect to composition) function of $\mathsf{erf}(\cdot)$, and is well tabulated by several software packages, such as Matlab. The definition of inverse function in (17) for interval $[a,b)$, such that $[-\phi_0,\phi_0) \subseteq [a,b) \subseteq [-\pi,\pi)$, is the most general case. As one can see in Fig. 1, our quantization of the angular space is well-aligned with the directivity pattern.

We also provide the details of constructing the angle dictionaries for dipole antenna directivity patterns in Appendix A.

IV. Recovering the Channel From A Few Bits: A Closed-Form Solution

In order to reduce the feedback overhead without irrevocably sacrificing our ability to recover accurate CSI at the BS,

we propose to use a *pseudo-random* dimensionality-reducing linear operator $\mathbf{P}^{\mathsf{H}}$ to $\mathbf{y}$. The outcome is quantized with a very simple sign quantizer and is communicated to BS through a low-rate limited feedback channel. More specifically, the BS receives

$$\mathbf{b}_{\Re} + \mathsf{j}\,\mathbf{b}_{\Im} = \mathsf{sign}\big(\Re\big(\mathbf{P}^{\mathsf{H}}\mathbf{y}\big)\big) + \mathsf{j}\,\mathsf{sign}\big(\Im\big(\mathbf{P}^{\mathsf{H}}\mathbf{y}\big)\big)\,, \tag{18}$$

where $\mathbf{P} \in \mathbb{C}^{M_{\mathrm{R}}N_{\mathrm{tr}} \times N_{\mathrm{fb}}}$, with $N_{\mathrm{fb}} \leq M_{\mathrm{R}}N_{\mathrm{tr}}$. To facilitate operating in the more convenient real domain, consider the following definitions

$$\mathbf{C}_{\Re} \triangleq \left[\Re(\mathbf{Q}^{\mathsf{H}}\mathbf{P})^{\top}\ \Im(\mathbf{Q}^{\mathsf{H}}\mathbf{P})^{\top}\right]^{\top}, \tag{19a}$$

$$\mathbf{C}_{\Im} \triangleq \left[-\Im(\mathbf{Q}^{\mathsf{H}}\mathbf{P})^{\top}\ \Re(\mathbf{Q}^{\mathsf{H}}\mathbf{P})^{\top}\right]^{\top}, \tag{19b}$$

$$\mathbf{C} \triangleq [\mathbf{C}_{\Re}\ \ \mathbf{C}_{\Im}] = [\mathbf{c}_1\ \mathbf{c}_2 \ldots\ \mathbf{c}_{2N_{\mathrm{fb}}}] \in \mathbb{R}^{2G \times 2N_{\mathrm{fb}}}, \tag{19c}$$

$$\mathbf{x} \triangleq \left[\Re(\mathbf{g})^{\top}\ \Im(\mathbf{g})^{\top}\right]^{\top} \in \mathbb{R}^{2G}, \tag{19d}$$

$$\mathbf{b} \triangleq \left[\mathbf{b}_{\Re}^{\top}\ \mathbf{b}_{\Im}^{\top}\right]^{\top} = [b_1\ b_2\ \ldots\ b_{2N_{\mathrm{fb}}}]^{\top} \in \mathbb{R}^{2N_{\mathrm{fb}}}, \tag{19e}$$

$$\mathbf{z} \triangleq \left[\mathbf{z}_{\Re}^{\top}\ \mathbf{z}_{\Im}^{\top}\right]^{\top} = [z_1\ z_2\ \ldots\ z_{2N_{\mathrm{fb}}}]^{\top} \in \mathbb{R}^{2N_{\mathrm{fb}}}, \tag{19f}$$

with $\Re\big(\mathbf{P}^{\mathsf{H}}\mathbf{Q}\mathbf{g}\big) = \mathbf{C}_{\Re}^{\top}\mathbf{x}$, $\Im\big(\mathbf{P}^{\mathsf{H}}\mathbf{Q}\mathbf{g}\big) = \mathbf{C}_{\Im}^{\top}\mathbf{x}$, $\mathbf{z}_{\Re} = \Re\big(\mathbf{P}^{\mathsf{H}}\mathbf{n}\big)$, and $\mathbf{z}_{\Im} = \Im\big(\mathbf{P}^{\mathsf{H}}\mathbf{n}\big)$. Using the above, in conjunction with (18), the received feedback bits at the BS are given by

$$b_i = \mathsf{sign}\big(\mathbf{c}_i^{\top}\mathbf{x} + z_i\big)\,, \quad i = 1, 2, \ldots, 2N_{\mathrm{fb}}. \tag{20}$$

The objective at the BS is to estimate $\mathbf{x}$ from the $2N_{\mathrm{fb}}$ sign observations $\mathbf{b}$ and the compound dictionary-training matrix $\mathbf{C}$. Since the sparse complex channel vector $\mathbf{g}$ has L non-zero elements, the real vector $\mathbf{x}$ has $2L$ non-zero elements.

A. Single-Bit Compressed Sensing

We make use of the fact that $\mathbf{x}$ is sparse, (cf. $\mathbf{G}$ being sparse). We seek a sparse $\mathbf{x}$ that yields maximal agreement between the observed and the reconstructed signs. Ideally, one would use a formulation that minimizes

$$-\sum_{i=1}^{2N_{\mathrm{fb}}} \mathsf{sign}(\mathbf{c}_i^{\top}\mathbf{x})b_i + \zeta\|\mathbf{x}\|_0, \tag{21}$$

where $\zeta > 0$ is a regularization parameter, but this is difficult in terms of optimization. As an optimization surrogate, we consider the following convex estimator

$$\widehat{\mathbf{x}} = \arg \min_{\|\mathbf{x}\|_2 \leq R_2} -\mathbf{x}^{\top}\mathbf{C}\,\mathbf{b} + \zeta\,\|\mathbf{x}\|_1, \tag{22}$$

where R_2 is an upper bound on the norm of $\mathbf{x}$ and the constraint prevents elements of $\mathbf{x}$ from becoming unbounded when ζ is small. Note that the above formulation in (22) has been shown effective for being an optimization surrogate of the objective function (21), both in practice and in theory. In fact, our reformulated problem is an instance of single-bit compressive sensing (CS), which has attracted significant attention in recent years [19], [20], [21], [22], [33], [34]. We note that the specific estimator in (22) has been proposed in [22]. The merit of this approach is that it has a closed-form solution

$$\widehat{\mathbf{x}} = \begin{cases} \mathbf{0}, & \|\mathbf{C}\,\mathbf{b}\|_{\infty} \leq \zeta, \\ \frac{R_2\,\mathsf{T}(\zeta;\,\mathbf{C}\,\mathbf{b})}{\|\mathsf{T}(\zeta;\,\mathbf{C}\,\mathbf{b})\|_2}, & \text{otherwise}, \end{cases} \tag{23}$$

where $\mathsf{T}(\zeta;\mathbf{x})$ denotes the shrinkage thresholding operator defined as $[\mathsf{T}(\zeta;\mathbf{x})]_i = (|x_i| - \zeta)_+\,\mathsf{sign}(x_i),\ \ i = 1, 2, \ldots, 2G$. This closed-form solution is ideal for real-time implementation. If the elements of $\mathbf{C}$ are drawn from a Gaussian distribution and $R_2 = 1$, the above simple closed-form solution can recover approximately $2L$-sparse vectors on the unit hypersphere with sample complexity $\mathcal{O}(\frac{2L\log G}{\epsilon^4})$, where sample complexity is the number of single-bit measurements needed to recover the desired vector within precision at most ϵ. Note that this identifiability result also justifies our dropping of the coupling between the supports of the real and imaginary parts of $\mathbf{g}$ in the conversion of the complex model to real form. The overall computational cost of computing Eq. (23) is $\mathcal{O}(N_{\mathrm{fb}}G)$.

B. Remarks

1) Channel Reconstruction: For a given estimate $\widehat{\mathbf{x}}$, using (19d) we set $\widehat{g}_i = \widehat{x}_i + \mathsf{j}\,\widehat{x}_{G+i},\ i = 1, 2, \ldots, G$. Then, having $\widehat{\mathbf{g}}$, the procedure to estimate $\mathbf{H}$ is straightforward. Namely, an estimate $\widehat{\mathbf{G}}$ from $\widehat{\mathbf{g}}$ is formed, and through Eq. (10) an estimate $\widehat{\mathbf{H}}$ is obtained.

2) Number of Feedback Bits: The proposed method requires from the UE to feed back $2N_{\mathrm{fb}}$ bits. By assumption $N_{\mathrm{fb}} \leq N_{\mathrm{tr}}M_{\mathrm{R}}$, so for binary quantization the number of feedback bits is upper bounded by $2N_{\mathrm{tr}}M_{\mathrm{R}}$. For massive MIMO $M_{\mathrm{T}} \gg N_{\mathrm{tr}}$ and $M_{\mathrm{T}} \gg M_{\mathrm{R}}$, hence the proposed algorithm is very-well suited for massive MIMO systems.

3) Feedback Delays: Compared to feedback schemes that need heavy optimization procedures at the UE [5], [9], [14], the proposed scheme introduces negligible feedback delays since the UE simply applies a linear transformation $\mathbf{P}^{\mathsf{H}}$ on received vector $\mathbf{y}$ and then takes signs.

V. Numerical Results

In this section, all directivity patterns are defined over the sector $[-\pi/2, \pi/2)$. Single-antenna UE devices are considered, i.e., $M_{\mathrm{R}} = 1$, while the BS is equipped with a ULA. The beamforming gain is employed as a performance metric [11], [12]

$$\text{Beamforming Gain} \triangleq \mathbb{E}\left[\frac{P_{\mathrm{T}}}{\|\widehat{\mathbf{h}}\|_2^2}\left|\mathbf{h}^{\mathsf{H}}\,\widehat{\mathbf{h}}\right|_2^2\right]. \tag{24}$$

This metric measures the similarity between the actual channel $\mathbf{h}$ and the normalized channel estimate $\widehat{\mathbf{h}}$ and is proportional to average received SNR. The following algorithms are used as baselines and compared with the proposed single-bit CS limited feedback sparse channel estimation algorithm:

- LS channel estimation at the UE, given by $\widehat{\mathbf{h}}_{\mathrm{LS}}^{\mathsf{H}} = \mathbf{Y}\mathbf{S}^{\dagger}$, and quantization of $\widehat{\mathbf{h}}_{\mathrm{LS}}^{\mathsf{H}}$'s elements using Lloyd-Max scalar quantizer of Q bits per real number. This feedback scheme requires exactly $2QM_{\mathrm{T}}M_{\mathrm{R}}$ feedback bits and is abbreviated as LS-SQ.
- LS channel estimation at UE, vector quantization of $\widehat{\mathbf{h}}_{\mathrm{LS}}$, and then feedback. The vector quantization strategy of [11] based on 2^Q-PSK is utilized due to very small

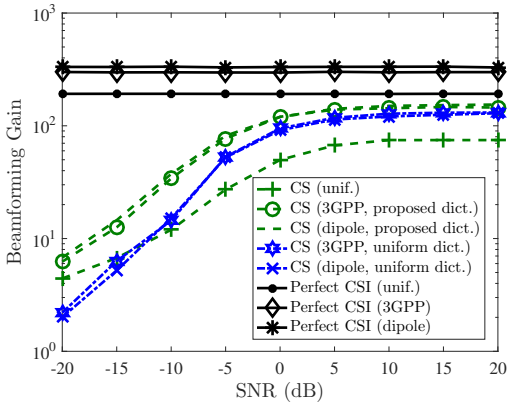

Fig. 2. The proposed angle dictionaries offer higher beamforming gain compared to uniform angle dictionaries that do not account directivity patterns.

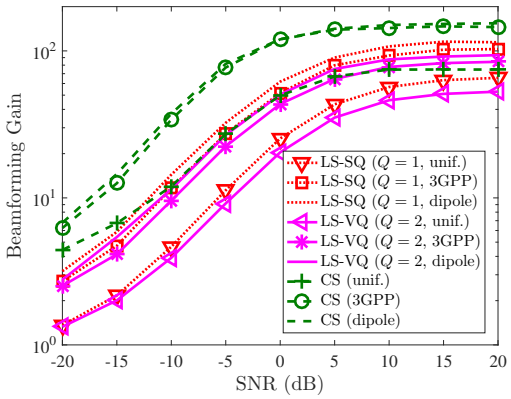

Fig. 3. The proposed algorithm outperforms LS-SQ and LS-VQ for every value of SNR and any studied antenna directivity pattern.

number of required feedback bits, equal to $Q(M_{\mathrm{T}}-1)$. This scheme is abbreviated as LS-VQ.

For CS, the AoD dictionary for 3GPP and dipole antenna directivity patterns is constructed following the procedures in Sec. III and Appendix A, respectively. The 3GPP directivity pattern employs $\mathtt{A}_m = 30$ dB and $\phi_{3\mathrm{dB}} = 55$ degrees. For fairness across different antenna directivity patterns their maximum directional gain is modified so as all of them have the same area over AoD interval $[-\pi/2, \pi/2)$, equal to π. Thus, $\mathtt{G}_{\mathrm{B}} = 2.184$ and $\mathtt{G}_{\mathrm{A}} = 2.35618$ are utilized.

Figs. 2 and 3 show the impact of transmit SNR on beamforming gain. We consider $P_{\mathrm{T}} = 1$ Watt transmission power, $M_{\mathrm{T}} = 192$ transmit antennas, and $N_{\mathrm{tr}} = 80$ training period; single-bit CS algorithm uses $N_{\mathrm{fb}} = 80$ (corresponding to 160 feedback bits overhead) and a random-column-permuted DFT matrix of size N_{fb} is used as dimensionality reducing matrix $\mathbf{P}$, while the dictionary sizes were set to $G_{\mathrm{T}} = G_{\mathrm{R}} = 240$. For the channel model in (2), we assume that L follows a discrete uniform distribution over $[5, 6, \ldots, 10]$. This choice for modeling the number of paths applies an implicit averaging with respect to paths to obtain a more robust average performance. AoD and AoA are uniformly distributed over $[-\pi/2, \pi/2)$, i.e., $\phi_l, \phi_l' \sim \mathcal{U}[-\pi/2, \pi/2]$, and Rician fading is considered, i.e., $\alpha_l \sim \mathcal{CN}\Big(\mathrm{e}^{\mathrm{j}\varphi_l}\sqrt{\frac{\kappa_l}{\kappa_l+1}}, \frac{1}{\kappa_l+1}\Big)$, with $\varphi_l \sim \mathcal{U}[0, 2\pi]$ and $\kappa_l \sim \mathcal{U}[0, 40)$.

Fig. 2 compares the proposed dictionaries with the uniform dictionary for different antenna directivity patterns. As a benchmark, the beamforming gain assuming perfect CSI is also included. It is observed that for 3GPP and dipole directivity patterns, the proposed angle dictionaries offer higher performance compared to uniform dictionaries, which do not account for the directivity patterns at the BS. The maximum performance gap among proposed and uniform dictionaries is 3 dB in the low SNR regime, which is an evident improvement. Furthermore, as the SNR increases the performance gap between the proposed CS method and perfect CSI becomes 3 dB.

Fig. 3 shows a comparison of the proposed algorithm with LS scalar and vector quantizations schemes. Using $Q = 1$ for LS-SQ and $Q = 2$ for LS-VQ result in 384 and 382 feedback bits, respectively. Thus, all schemes utilize roughly the same number of feedback bits. It can be seen that the proposed algorithm outperforms LS schemes in terms of beamforming gain for all directivity patterns considered. This is rather encouraging: although the proposed algorithm uses somewhat fewer feedback bits, it achieves much better performance due to its judicious design.

TABLE I
SIMULATION PARAMETERS.

Parameters	Values
Carrier Frequency (F_{c})	2 GHz
Carrier Wavelength (λ)	$\lambda = \mathtt{c}/F_{\mathrm{c}}$
Noise power (σ^2)	10^{-10} Watts
Transmit power (P_{T})	0.5 Watts
Path number (L)	Discrete uniform in $[5, 6, \ldots, 19, 20]$
AoA and AoD (ϕ_l and ϕ_l')	$\phi_l, \phi_l' \sim \mathcal{U}[-\pi/2, \pi/2]$
Distance of path l (d_l)	$d_l \sim \mathcal{U}[80, 120]$
Path loss exponent (η)	$\eta \sim \mathcal{N}(2.8, 0.1^2)$
Inverse path loss of path l (ρ_l)	$\left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{1}{d_l}\right)^{\eta}$
Shadowing (v_l)	$10\log_{10}(v_l) \sim \mathcal{N}(10\log_{10}(\rho_l), 4^2)$
Rician parameter (κ_l)	$\kappa_l \sim \mathcal{U}[0, 50]$
Multipath gain (α_l)	$\alpha_l \sim \mathcal{CN}\Big(\mathrm{e}^{\mathrm{j}\varphi_l}\sqrt{\frac{\kappa_l}{\kappa_l+1}v_l}, \frac{1}{\kappa_l+1}v_l\Big)$
φ_l	$\varphi_l \sim \mathcal{U}[0, 2\pi]$

Next, we consider a more realistic scenario where path-loss and shadowing are added to the channel model according to Table I. Please note that using the parameters of Table I the average received SNR (incorporating path-loss and small- and large-scale fading effects) changes per realization, so we apply an implicit averaging with respect to average received SNR. Under such settings, the simulations cover many more different cases and can better reflect the real-world performance of the various algorithms. Also, $N_{\mathrm{tr}} = 64$ training symbols are used, while single-bit CS employees $N_{\mathrm{fb}} = 64$ with dictionary sizes $G_{\mathrm{T}} = G_{\mathrm{R}} = 180$.

Fig. 4 examines a massive MIMO scenario where M_{T} becomes very large. Note that the proposed single-bit CS algorithm achieves the largest beamforming gain for all values

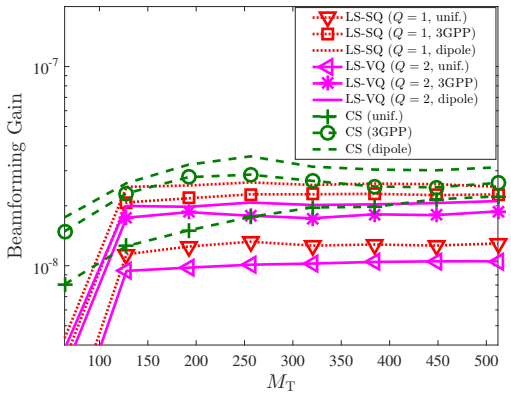

Fig. 4. The proposed algorithm outperforms LS-SQ and LS-VQ for any value of M_{T} and any studied antenna directivity pattern.

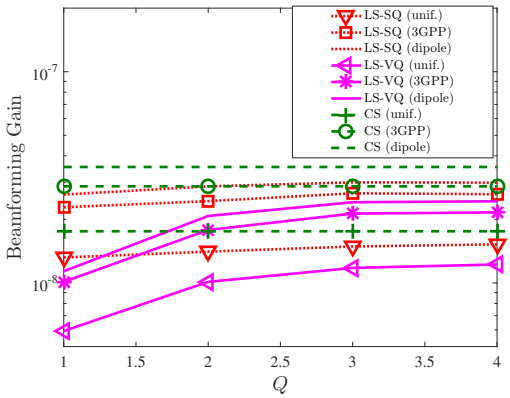

Fig. 5. Beamforming gain as a function of Q for 256 BS antennas. The proposed scheme offers an upper performance bound for LS schemes.

TABLE II
NUMBER OF FEEDBACK BITS USED UNDER VARIOUS M_{T}'S.

	M_{T}							
Algorithms	64	128	192	256	320	384	448	512
CS	128	128	128	128	128	128	128	128
LS-SQ	128	256	384	512	640	768	896	1024
LS-VQ	126	254	382	510	638	766	894	1022

of M_{T} and all directivity patterns. This is consistent with our previous simulation. LS-VQ has the worst performance in this case where the receive-SNR is varying. We observe that under this realistic scenario the beamforming gain takes values of the order of 10^{-8}. This is not surprising since on top of small-scale fading this scenario further incorporates path-loss and shadowing effects. The noise level is in the order of 10^{-10} Watts, and thus, such beamforming gain makes a lot of sense in practice. Table II illustrates the number of feedback bits used under different M_T's. Note that the proposed algorithm consistently outperforms the LS schemes in terms of beamforming gain, even though the latter schemes use many more feedback bits when M_T increases. Particularly, when $M_{\mathrm{T}} = 1024$, the number of feedback bits used by the LS schemes is almost 10 times higher than the number of feedback bits used by the proposed scheme.

Fig. 5 shows the beamforming gain as a function of parameter Q for different antenna directivity patterns using $M_{\mathrm{T}} = 256$. We recall that Q parametrizes the accuracy as well as the corresponding number of feedback bits for LS schemes. For all values of parameter Q and all studied antenna directivity patterns the proposed single-bit CS algorithm outperforms LS schemes.

VI. CONCLUSION

This work provides a new limited feedback sparse channel estimation framework using a single-bit compressed sensing algorithm that (a) uses overcomplete angle dictionaries accounting for the antenna directivity patterns; (b) entails very small computational complexity at both ends of the link, as it admits closed-form solution; and (c) attains a desired level of channel estimation accuracy at a feedback rate that is independent of the number of BS antennas, making it ideal for massive MIMO applications. A new angle dictionary construction methodology was presented that further boosts the performance of the proposed algorithm. Simulations reveal that the proposed algorithm outperforms least-squares scalar and vector quantization techniques in terms of beamforming gain.

APPENDIX A
DIPOLE ANTENNA DIRECTIVITY PATTERN

Dipole directivity pattern is given in (4). To study the most general case for interval $[a, b)$, it is assumed that $a \leq \frac{\pi}{2}$ and $b \geq \frac{\pi}{2}$. For notational convenience an auxiliary function $\mathsf{D}(\phi)$ is defined as

$$\mathsf{D}(\phi) \triangleq \frac{\mathsf{G}_{\mathrm{A}}}{12}\left(12\sin(\phi) - 4\sin^3(\phi) + 8\right). \tag{25}$$

The cumulative function for dipole directivity pattern in (4) is given by

$$\mathsf{G}(\phi) = \begin{cases} \mathsf{D}(a) - \mathsf{D}(\phi), & \phi \in \left[a, -\frac{\pi}{2}\right), \\ \mathsf{D}(a) + \mathsf{D}(\phi), & \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \\ \mathsf{D}(a) + 2\mathsf{D}\left(\frac{\pi}{2}\right) - \mathsf{D}(\phi), & \phi \in \left[\frac{\pi}{2}, b\right). \end{cases} \tag{26}$$

Note that $\mathsf{G}(\phi)$ is a continuous increasing function since $\mathsf{q}(\phi) > 0$ for all $\phi \in [a, b)$ except a set of measure zero [35]. Setting $\mathsf{G}(\phi) = y$, the calculation of the inverse function in (26) requires to solve a composite polynomial of the following form

$$\sin^3(\phi) - 3\sin(\phi) - c = 0, \tag{27}$$

where constant c changes according to which interval ϕ belongs in Eq. (26). To find the roots of the polynomial in (27), we first ignore sine function by solving

$$0 = \begin{cases} \phi^3 - 3\phi - \left(2 + \frac{3(y - \mathsf{D}(a))}{\mathsf{G}}\right), & \phi \in \left[a, -\frac{\pi}{2}\right), \\ \phi^3 - 3\phi - \left(2 - \frac{3(y - \mathsf{D}(a))}{\mathsf{G}}\right), & \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \\ \phi^3 - 3\phi - \left(2 + \frac{3\left(y - \mathsf{D}(a) - 2\mathsf{D}\left(\frac{\pi}{2}\right)\right)}{\mathsf{G}}\right), & \phi \in \left[\frac{\pi}{2}, b\right). \end{cases} \tag{28}$$

The polynomials above have real coefficients, thus, there is at least one real root, and the other two are either complex conjugates or both real. Using same methodology as in [36], we find: $\mathsf{D}_0 = 9,\ \iota_3 = \frac{-1-\sqrt{3}}{2}$,

$$\mathsf{D}_1(y) = \begin{cases} -27\left(2 + \frac{3(y-\mathsf{D}(a))}{\mathsf{G}_\mathrm{A}}\right), & y \in \left[0, \mathsf{G}\left(-\frac{\pi}{2}\right)\right), \\ -27\left(2 - \frac{3(y-\mathsf{D}(a))}{\mathsf{G}_\mathrm{A}}\right), & y \in \left[\mathsf{G}\left(-\frac{\pi}{2}\right), \mathsf{G}\left(\frac{\pi}{2}\right)\right), \\ -27\left(2 + \frac{3(y-\mathsf{D}\left(\frac{\pi}{2}\right)-\mathsf{D}(a))}{\mathsf{G}_\mathrm{A}}\right), & y \in \left[\mathsf{G}\left(\frac{\pi}{2}\right), \mathsf{G}(b)\right), \end{cases} \tag{29}$$

and

$$\mathsf{V}(y) = \sqrt[3]{\frac{\mathsf{D}_1(y) + \sqrt{\mathsf{D}_1^2(y) - 4\mathsf{D}_0^3}}{2}}. \tag{30}$$

Without getting to algebraic details, after finding the roots in (28), we incorporate sine function and solve (27), which offers the inverse of (26)

$$\mathsf{G}^{-1}(y) = \begin{cases} \mathsf{asin}\left(\frac{\iota_3 \mathsf{V}(y) + \frac{\iota_3 \mathsf{D}_0}{\mathsf{V}(y)}}{3}\right) - \pi, & y \in \left[0, \mathsf{G}\left(-\frac{\pi}{2}\right)\right), \\ -\mathsf{asin}\left(\frac{\iota_3 \mathsf{V}(y) + \frac{\iota_3 \mathsf{D}_0}{\mathsf{V}(y)}}{3}\right), & y \in \left[\mathsf{G}\left(-\frac{\pi}{2}\right), \mathsf{G}\left(\frac{\pi}{2}\right)\right), \\ \mathsf{asin}\left(\frac{\iota_3 \mathsf{V}(y) + \frac{\iota_3 \mathsf{D}_0}{\mathsf{V}(y)}}{3}\right) + \pi, & y \in \left[\mathsf{G}\left(\frac{\pi}{2}\right), \mathsf{G}(b)\right). \end{cases} \tag{31}$$

REFERENCES

[1] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless. Comm.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[3] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[5] G. Caire, N. Fal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theor.*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[6] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems." *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.

[7] H. Ji *et al.*, "Overview of full-dimension MIMO in LTE-Advanced Pro." *CoRR*, 2016.

[8] E. Dahlman, S. Parkvall, and J. Skold, *4G, LTE-Advanced Pro and The Road to 5G*. Elsevier Science, 2016.

[9] D. J. Love, R. W. Heath, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[10] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theor.*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[11] D. J. Ryan, I. V. L. Clarkson, I. B. Collings, D. Guo, and M. L. Honig, "QAM and PSK codebooks for limited feedback MIMO beamforming," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1184–1196, Apr. 2009.

[12] J. Choi, Z. Chance, D. J. Love, and U. Madhow, "Noncoherent trellis coded quantization: A practical limited feedback technique for massive MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5016–5029, Dec. 2013.

[13] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless. Comm.*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[14] B. Lee, J. Choi, J.-Y. Seol, D. J. Love, and B. Shim, "Antenna grouping based feedback compression for FDD-based massive MIMO systems," *IEEE Trans. Comm.*, vol. 63, no. 9, pp. 3261–3274, Sep. 2015.

[15] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theor.*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[16] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[17] 3GPP TS 36.101 V13.2.1, "Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) radio transmission and reception, Release 13," May 2016.

[18] A. Kammoun, H. Khanfir, Z. Altman, M. Debbah, and M. Kamoun, "Preliminary results on 3D channel modeling: From theory to standardization," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1219–1229, Jun. 2014.

[19] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in *Proc. IEEE Information Sciences and Systems (CISS)*, 2008, pp. 16–21.

[20] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Inf. Theor.*, vol. 59, no. 4, pp. 2082–2102, Apr. 2013.

[21] Y. Plan and R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach," *IEEE Trans. Inf. Theor.*, vol. 59, no. 1, pp. 482–494, Jan. 2013.

[22] L. Zhang, J. Yi, and R. Jin, "Efficient algorithms for robust one-bit compressive sensing," in *Proc. International Conference on Machine Learning (ICML)*, Beijing, China, Jun. 2014, pp. 820–828.

[23] C. Rusu, R. Méndez-Rial, N. González-Prelcic, and J. R. W. Heath, "Adaptive one-bit compressive sensing with application to low-precision receivers at mmWave," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, San Diego, CA, Dec. 2015.

[24] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.

[25] N. Gonzalez-Prelcic, K. T. Truong, C. Rusu, and R. W. Heath, "Compressive channel estimation in FDD multi-cell massive MIMO systems with arbitrary arrays," in *IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–5.

[26] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theor.*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[27] O. Mehanna and N. D. Sidiropoulos, "Maximum likelihood passive and active sensing of wideband power spectra from few bits," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1391–1403, Mar. 2015.

[28] B. Gopalakrishnan and N. D. Sidiropoulos, "Cognitive transmit beamforming from binary CSIT," *IEEE Trans. Wireless. Comm.*, vol. 14, no. 2, pp. 895–906, Feb. 2015.

[29] O. Mehanna and N. D. Sidiropoulos, "Channel tracking and transmit beamforming with frugal feedback," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6402–6413, Dec. 2014.

[30] C. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. Wiley, 2005.

[31] 3GPP TR 25.996 V9.0.0, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Spatial channel model for Multiple Input Multiple Output (MIMO) simulations, Release 9," Dec. 2009.

[32] 3GPP TR 37.840 V12.1.0, "Technical Specification Group Radio Access Network; Study of Radio Frequency (RF) and Electromagnetic Compatibility (EMC) requirements for Active Antenna Array System (AAS) base station, Release 12," Dec. 2013.

[33] A. Gupta, R. D. Nowak, and B. Recht, "Sample complexity for 1-bit compressed sensing and sparse classification," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Austin, TX, Jun. 2010, pp. 1553–1557.

[34] Y. Plan and R. Vershynin, "One-bit compressed sensing by linear programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1275–1297, 2013.

[35] G. B. Folland, *Real analysis: Modern techniques and their applications*, 2nd ed. John Wiley & Sons, Inc., New York, 1999.

[36] K. Huang and N. D. Sidiropoulos, "Consensus-ADMM for general quadratically constrained quadratic programming," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5297–5310, Oct. 2016.