IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 February 2025, accepted 17 February 2025, date of publication 25 February 2025, date of current version 4 March 2025.
Digital Object Identifier 10.1109/ACCESS.2025.3545673

RESEARCH ARTICLE

Codebook-Based Trellis-Coded Quantization Scheme Using K-Means Clustering for Massive MIMO Systems

SUNGSOO PARK [1], (Member, IEEE), AND GYUYEOL KONG [2], (Member, IEEE)
[1]Department of Train Control and Communications Research, Korea Railroad Research Institute, Uiwang-si 16105, Republic of Korea
[2]Division of Mechanical and Electronics Engineering, Hansung University, Seoul 02876, Republic of Korea

Corresponding author: Gyuyeol Kong (gykong@hansung.ac.kr)

The work of Sungsoo Park was supported in part by the Research and Development Program (Development of Core Technologies for Virtually-Coupled Train Set to Provide Demand-Oriented Train Operation Services) of the Korea Railroad Research Institute (KRRI) under Grant PK2501B1. The work of Gyuyeol Kong was supported by Hansung University.

ABSTRACT This paper introduces a codebook-based trellis-coded quantization (TCQ) approach utilizing K-means clustering, designed specifically for massive multiple-input multiple-output systems. The proposed TCQ scheme follows the structure of conventional TCQ, quantizing channels through a trellis and source constellation to minimize latency. The performance of channel quantization significantly depends on the chosen source constellation. The conventional TCQ scheme, using a fixed set of constellations as source constellations, imposes limitations on beamforming gain performance. In our proposed TCQ scheme, we employ a codebook to utilize a set of constellations tailored to the channel environments as the source constellation, aiming for enhanced beamforming gain performance. This approach leads to a slight increase in feedback overhead. The codebook's constellation sets are generated by determining centroids through K-means clustering of channel elements and mapping them to high-order quadrature amplitude modulation constellations. Simulation results demonstrate that our proposed TCQ scheme, with a comparable computational complexity, shows the improved average beamforming gains compared to the conventional TCQ scheme.

INDEX TERMS Limited feedback, massive MIMO system, trellis-coded quantization (TCQ), URLLC, K-means clustering.

I. INTRODUCTION

For new applications with strict reliability and latency requirements such as intelligent transportation, augmented/virtual reality, and industrial automation, the fifth-generation (5G) New Radio (NR) standard provides ultra-reliable low-latency communications (URLLC) service category [1]. The physical design of URLLC is a challenging problem because it requires simultaneous consideration of conflicting factors of reliability and latency. In this regard, the third generation partnership project (3GPP) Release 15 defines the requirements for the transmission 32-Byte payloads with a latency of 1 ms and a reliability target of 10^{-5} for URLLC design [2]. To support new industrial use cases requiring lower latency, Release 16 defines the further enhanced URLLC requirements with a latency of 0.5 ms and a reliability target of 10^{-6} [3].

Massive multiple-input multiple-output (MIMO) systems equipped with a large number of transmit antennas in a base station (BS) can improve spectral efficiency and energy efficiency [4], [5]. In particular, the use of large numbers of antennas with channel state information (CSI) in BS can acquire a significant array gain through beamforming [6]. In the frequency-division duplex (FDD) systems, the BS must be fed back downlink CSI from the user equipment (UE) to uplink channel. Since it is difficult to feedback the entire

The associate editor coordinating the review of this manuscript and approving it for publication was Nurul I. Sarkar.

massive CSI due to the limitation of uplink channel capacity, channel quantization with lower feedback overhead and higher beamforming gain plays an important role for FDD massive MIMO systems.

Various channel quantization schemes have been introduced for FDD massive MIMO systems [7], [8], [9], [10]. The random vector quantization (RVQ) is a codebook-based channel quantization scheme where the distance between randomly generated unit-norm vectors is maximized [7], [8]. In order to reduce quantization errors, the RVQ scheme increases the number of codewords as the number of transmit antennas increases, making the computational complexity of an appropriate codeword search infeasible. The trellis-coded quantization (TCQ) scheme has been proposed to perform channel quantization with low complexity [9], [10]. In the TCQ scheme, the channel is quantized into codewords on the trellis path with a source constellation, which can perform channel quantization with a linearly increasing computational complexity as the number of transmit antennas increases. Although this scheme can perform channel quantization with low complexity, it can not actively cope with channel situations by using a fixed source constellation, resulting in performance limitations. To flexibly cope with the channel situation, the differential TCQ schemes in [11] have been proposed for the temporally and spatially correlated channels by using the source constellation considering the correlation effect. However, this technique results in significant performance loss in environments with low channel correlation, such as fast time-varying channels. From a URLLC perspective, channel quantization is required to achieve high beamforming gain while minimizing computational complexity rather than lowering feedback overhead to achieve low latency and high reliability requirements.

Motivated by the above, in this paper, a codebook-based TCQ scheme using K-means clustering is proposed for massive MIMO systems. The proposed TCQ scheme adopts the structure of the conventional TCQ scheme in [10] for low latency, except for the source constellation setting constituting the trellis to improve reliability. That is, the proposed TCQ scheme focuses on maximizing beamforming gain performance and minimizing additional computational complexity compared to conventional TCQ scheme in [10] by adding a small number of overheads. The proposed TCQ scheme selects the source constellation based on the codebook to utilize the various source constellations suitable for the channel environments, and additional feedback bits are required to deliver the selected codeword information. The codebook is designed in advance through a channel training process according to the channel environment, and this process is performed when the statistical characteristics of the channel change. Each codeword consists of a combination of high-order quadrature amplitude modulation (QAM) constellations close to the centroids generated through K-means clustering of massive channel elements. Among various generated codewords, a codebook is constructed in consideration of additional feedback overhead by grouping

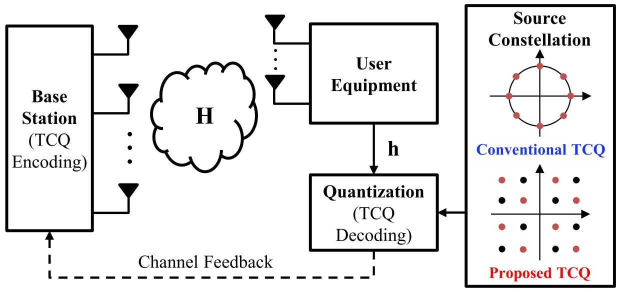


FIGURE 1. **Block diagram of the conventional and proposed TCQ schemes using 8 constellation points as the source constellation.**

codewords that occur frequently and have a large distance between constellations in the codeword. By operating TCQ based on codebook-based source constellations suitable for the channel environment, the proposed TCQ scheme significantly improves beamforming gain without increasing computational complexity compared to the conventional TCQ scheme, except for the training process for codebook generation in advance.

The remainder of this paper is organized as follows. In Section II, the system model is introduced. In Section III, the conventional TCQ scheme is briefly reviewed, and the proposed TCQ scheme is introduced in Section IV. The performance of proposed TCQ scheme is evaluated in Section V. Finally, Section VI summarizes the results of this study.

II. SYSTEM MODEL

We consider a MIMO channel with M_t transmit antennas at the BS and M_r receive antennas at the UE as shown in Fig. 1. The received signal $y \in \mathbb{C}^{M_r \times 1}$ can be modeled as

$$y = HFx + n, \tag{1}$$

where $H = \left[h_1^T, \ldots, h_{M_r}^T\right]^T$ is a $M_r \times M_t$ MIMO channel matrix, $\mathbf{h}_r \in \mathbb{C}^{1\times M_t}$ denotes the rth receive channel vector, $F = \left[f_1^T, \ldots, f_{M_t}^T\right]$ is a $M_t \times M_t$ precoding matrix, $\mathbf{f}_p \in \mathbb{C}^{1\times M_t}$ denotes the pth beamforming vector, $\mathbf{x} \in \mathbb{C}^{M_t\times 1}$ is the transmitted symbol vector, and $\mathbf{n} \in \mathbb{C}^{M_r\times 1}$ is the complex Gaussian noise. In this paper, for simplicity of expression, channel feedback is considered only for single multiple-input single-output (MISO) channel $\mathbf{h}$ that omit the index symbol r of $\mathbf{h}_r$. In addition, a vector $\mathbf{f}$ corresponding to the MISO channel $\mathbf{h}$ in F is considered as a beamforming vector by removing the index symbol p of $\mathbf{f}_p$. For the MISO channel h, three different channels are considered: an independent and identically distributed (i.i.d.) Rayleigh fading channel, a spatially correlated channel, and a temporally correlated channel.

The i.i.d. Rayleigh fading channel is modeled by

$$h = h_{iid}, \tag{2}$$

where h_{iid} is a $1 \times M_t$ vector with entries distributed according to $\mathcal{CN}(0, 1)$. The spatially correlated MISO

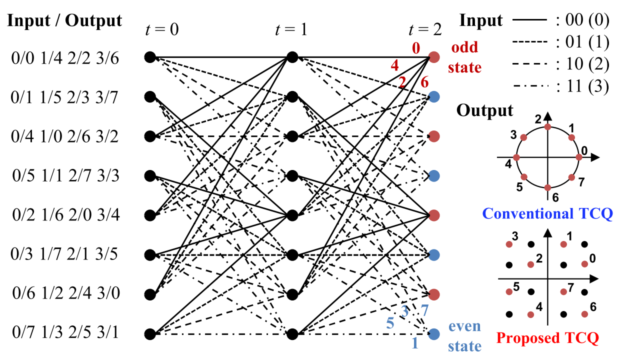


FIGURE 2. **Example of input/output state transition of Ungerboeck trellis with $B = 2$.**

channel is modeled by

$$h = h_{iid} R_T^{\frac{1}{2}}, \tag{3}$$

where R_T is a transmit spatial correlation matrix. For a spatially correlated channel, the exponential correlation model [13] is used as

$$[R_T]_{p_1, p_2} = \rho^{|p_1 - p_2|}, \tag{4}$$

where $[\cdot]_{p_1, p_2}$ denotes the matrix element at the p_1th row and the p_2th column and ρ is the correlation factor with $0 \leq \rho \leq 1$. As the spatial correlation increases, ρ approaches 1.

The temporally correlated channel in the kth fading block is modeled by a first-order Gauss-Markov process [14] written as

$$h[k] = \eta h[k-1] + \sqrt{1 - \eta^2} g[k], \tag{5}$$

where $g[k]$ is a $1 \times M_t$ vector with entries distributed according to $\mathcal{CN}(0, 1)$, and $h[0]$ is independent of $g[k]$ for all k. The temporal correlation coefficient is denoted by η with $0 \leq \eta \leq 1$. η approaches 1 as the channel has a higher temporal correlation and η approaches 0 as the channel has a lower temporal correlation.

III. CONVENTIONAL TCQ SCHEME

As shown in Fig. 1, the conventional TCQ scheme [10] using a trellis with a source constellation quantizes the channel through the decoding process at the UE and the encoding process at the BS, enabling low complexity for massive MISO channel feedback compared to the codebook approaches. The source constellation is composed of 2^{B+1} constellation points, where B is the number of quantization bits per the feedback channel. In [10], quadrature phase shift keying (QPSK), 8-ary phase shift keying (8-PSK), and 16-QAM with $B = 1$, 2, and 3, respectively, are considered as the source constellation. For the trellis, the Ungerboeck rate $B/(B+1)$ convolutional codes in [12] are adopted. The example of the input/output state transition of the trellis corresponding to the convolutional codes when the source constellation is 8-PSK ($B = 2$) is shown in Fig. 2. For the detailed description of the trellis with input/output transitions, see [12].

At the UE, through the decoding process with M_t-stage trellis, a trellis path with minimum Euclidean distance to the normalized channel vector $\bar{h} = h/\|h\|$ is selected by employing Viterbi algorithm, and an input bit sequence b representing the selected path is generated. Let $\bar{h}_t$ be the truncated vector of $\bar{h}$ up to the first t entries ($1 \leq t \leq M_t$) and $out\,(p_t)$ be the sequence of output constellation points for a given path p_t, where p_t is the trellis path up to the stage t. Then, the path metric $m\,(\cdot)$ can be defined as

$$\begin{aligned} m\,(p_t, \theta) &= \left\| \bar{h}_t - e^{j\theta} out\,(p_t) \right\|^2 \\ &= m\,(p_{t-1}, \theta) + \left| \bar{h}_t - e^{j\theta} out\,([p_{t-1}\ p_t]) \right|^2, \end{aligned} \tag{6}$$

where $\theta \in [0, 2\pi)$, and $\bar{h}_t$ and p_t are the tth entries of $\bar{h}$ and p_t, respectively. The path metric is calculated in parallel for the quantized values of $\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_{2^T}\}$, where T is the number of quantization bits for θ. The best path $\mathbf{p}_{best} \in \mathbb{P}_{M_t}$ and the phase $\theta_{best} \in \Theta$ which minimize the path metric $m\left(\mathbf{p}_{M_t}, \theta\right)$ can be obtained through the Viterbi algorithm, where $\mathbb{P}_{M_t}$ denotes all possible paths up to stage M_t. Then, the input bit sequence b is generated by mapping binary input stream corresponding to the selected path p_{best}. Then, the total feedback overhead bit is defined as $B_{tot} = B \cdot M_t$ by fixing the first state at $t = 0$ for each feedback channel as in [10]. It is important to point out that, as explained in [10], minimizing over θ incurs no additional feedback overhead because the transmitter does not need to know the value of θ_{best} during the beamforming vector reconstruction process.

At the BS, through the encoding process, an output sequence corresponding to the input bit sequence is reconstructed, and quantized channel c_{opt} is obtained by mapping the generated output sequence to the source constellation by out $(\mathbf{p}_{best})$. Finally, the beamforming vector f is computed as $c_{opt}/\left\|c_{opt}\right\|_2$.

IV. PROPOSED TCQ SCHEME

In this section, we present the proposed TCQ scheme for URLLC in massive MISO systems. In order to achieve high reliability, the proposed TCQ scheme focuses on improving beamforming gain performance by adding a small number of overhead bits compared to the conventional TCQ scheme [9], [10]. Also, for low latency, the proposed TCQ scheme basically adopts the conventional TCQ scheme structure using θ for path metric calculation in parallel, but the source constellation used to construct trellis is somewhat different. Unlike the conventional TCQ scheme which utilizes whole source constellation points of QPSK ($B = 1$), 8-PSK ($B = 2$), or 16-QAM ($B = 3$), the proposed TCQ scheme constitutes source constellation by selecting 4 ($B = 1$), 8 ($B = 2$), or 16 ($B = 3$) constellation points out of the higher-order Q-QAM constellations depending on the channel environments.

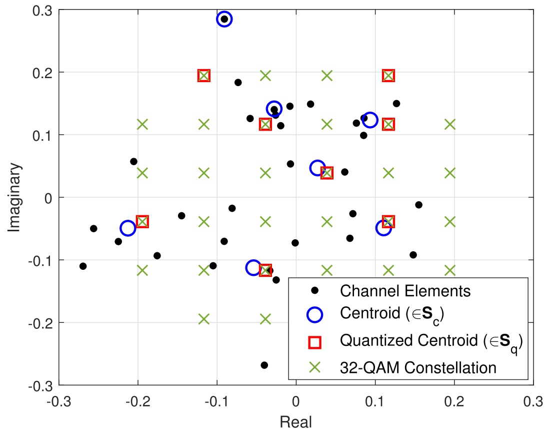


FIGURE 3. Example of quantization for the source constellation to 32-QAM constellations based on K-means clustering with $B = 2$ and $M_t = 32$.

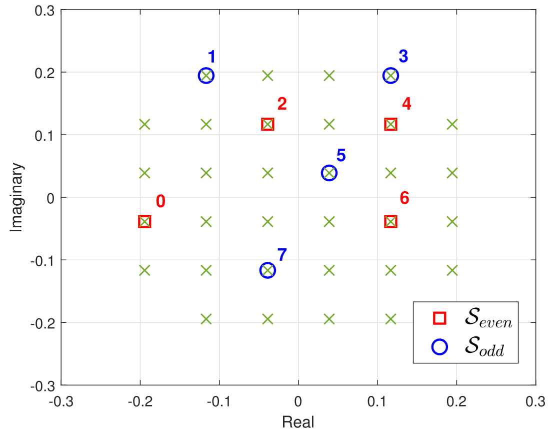


FIGURE 4. Example of set partitioning for quantized centroid S_q with $B = 2$ and $M_t = 32$.

$(Q > 2^{B+1})$ To select 2^{B+1} source constellation points suitable for the channel environment out of the higher-order Q-QAM constellations, the K-means clustering algorithm [13], [14] is used. To enhance beamforming gain performance while avoiding additional feedback overhead required to deliver the selected source constellation points, we propose to employ a codebook consisting of the selected source constellation points tailored to the channel environments. For effective source constellation generation and utilization, the proposed TCQ scheme includes three detailed techniques: quantization for source constellation to high-order QAM constellations, set partitioning and source constellation codebook design.

A. QUANTIZATION FOR SOURCE CONSTELLATION TO HIGH-ORDER QAM CONSTELLATION

The proposed TCQ scheme performs channel feedback through the source constellation selected in high-order Q-QAM constellations $(Q > 2^{B+1})$ rather than the fixed source constellation. The use of high Q enables more accurate representation of the channel, but it increases feedback overhead. One possible way to reduce the feedback overhead is to represent each channel element of $\bar{h}$ using 2^{B+1} constellation points selected from Q-QAM constellations. To select 2^{B+1} constellation points tailored to channel environments among Q constellations, we apply a clustering approach that separates channel elements into 2^{B+1} clusters. The K-means algorithm is an iterative algorithm that divides a data set into K non-overlapping clusters in which each data point belongs to only one group [13], [14]. Fig. 3 shows an example of quantization for the source constellation to 32-QAM constellations based on K-means clustering with $B = 2$ and $M_t = 32$. The purpose of quantization for the source constellation is to select 2^{B+1} constellation points from Q-QAM constellations representing M_t channel elements constituting $\bar{h}$. First, by using K-means algorithm, a set S_c containing 2^{B+1} centroids is produced by dividing M_t channel elements of $\bar{h}$ into 2^{B+1} clusters as shown in Fig. 3. Then, after Q-QAM demodulation of S_c, a set S_d containing the demodulated centroids with Q-QAM constellations can be obtained for the source constellation. If S_d contains 2^{B+1} constellation points, the quantized centroids S_q is set as S_d. However, fewer than 2^{B+1} constellation points may be selected for S_d if demodulation process quantizes more than one centroids in S_c to a single constellation point in Q-QAM constellations. In this case, S_q is constructed to have 2^{B+1} constellation points by adding S_d to $2^{B+1} - |S_d|$ constellation points with the closest distance among Q-QAM constellations from S_d. As shown in the example of Fig. 3, S_q with 8 constellation points is created by adding one additional constellation point to S_d containing 7 constellation points.

B. SET PARTITIONING

To map the quantized constellations S_q onto trellis, we present the process of set partitioning to maximize the free distance d_{free} for the rate $B/(B+1)$ convolutional codes in [12] as much as possible. Deriving optimal constellation points mapping onto trellis, which maximizes d_{free} under a combination of S_q in all cases, results in a computationally large complexity. To map constellation points onto trellis with low complexity, we consider a symbol mapping method that can enlarge d_{free} by only considering the distance between path outputs to each state rather than the distance between the entire paths. As shown in Fig. 2, we can observe the feature on trellis that the path outputs to the odd states are all even numbers $(0, 2, 4, 6)$, and the path outputs to the even states are all odd numbers $(1, 3, 5, 7)$. Based on this fact, we consider set partitioning of all path outputs into two sets, $\mathcal{S}_{even} = \{0, 2, 4, 6\}$ and $\mathcal{S}_{odd} = \{1, 3, 5, 7\}$. We map the symbols belonging to S_q to the constellation points by dividing the symbols into $\mathcal{S}_{even}$ and $\mathcal{S}_{odd}$, and set the minimum distance δ between symbols belonging to $\mathcal{S}_{even}$ and $\mathcal{S}_{odd}$ to exceed $d_{\min}$, where $d_{\min}$ is the minimum distance of Q-QAM constellation. Fig. 4 shows the example of set partitioning for S_q with

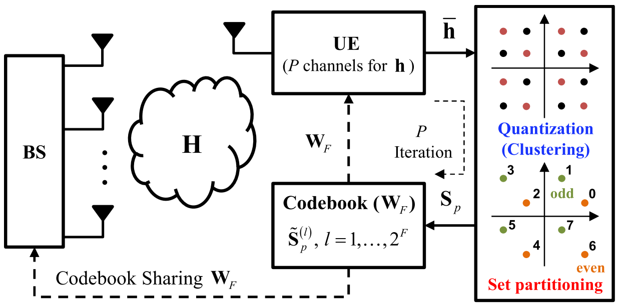


FIGURE 5. Source constellation codebook design process for channel training.

$B = 2$ and $M_t = 32$. Through set partitioning, a set S_p obtained by mapping the constellations in S_q into sets $\mathcal{S}_{even}$ and $\mathcal{S}_{odd}$ maximizes the Euclidean distance of each set, resulting in enlarging d_{free}.

C. SOURCE CONSTELLATION CODEBOOK DESIGN

The proposed TCQ scheme needs to deliver the source constellation based on S_p according to the channel environments to the BS, which requires additional $2^{B+1} \cdot \log_2 Q$ bits transmission for every channel feedback. When $B = 2$ and $Q = 32$, additional 40 bits are required for the source constellation delivery, which requires a method of delivering source constellations with less overhead. One possible way to avoid delivering the source constellation every channel feedback is to utilize a codebook constituted by the most frequently observed source constellations. As a result of observing the pattern of variously generated source constellations S_p, the sets of the constellation points $S_p^{(l)}$ for $l = 1, \cdots, L$ that occur frequently depending on the channel environments were observed, where L is the number of the candidate sets of the constellation points. In order to efficiently deliver the source constellation to the BS, the proposed TCQ scheme constructs a source constellation codebook with 2^F frequently occurring constellation patterns among $S_p^{(l)}$ for $l = 1, \cdots, L$, and transmits the selected source constellation to the BS using an additional F bits for every channel feedback. This overhead can be further reduced by transmitting the source constellation to the BS only when UE initially selects a source constellation or a new source constellation is selected due to a change of wireless channel characteristics such as spatial/temporal correlations. This additional F bits are not transmitted for every channel feedback, but could only when a codebook suitable for the channel is initially selected or an appropriate codebook is selected as the channel characteristics change. The source constellation codebook design process is performed only during the channel training, and the codebook $W_F \ni \tilde{S}_p^{(l)}$ for $l = 1, \cdots, 2^F$ is shared with the BS and UE after the codebook is created as shown in Fig. 5, where $\tilde{S}_p^{(l)}$ is the l-th constellation set arranged in descending order of occurrence of $S_p^{(l)}$ during the channel training. Arranging in descending order of occurrence of $S_p^{(l)}$ is to construct frequently occurring constellation sets for $S_p^{(l)}$ into the codebook W_F. With W_F, the total numbers of bits for channel feedback in the proposed TCQ scheme are $B_{tot}^{prop} = B \cdot M_t + F$ for the codebook updates and $B_{tot}^{prop} = B \cdot M_t$ except the codebook updates. Therefore, in most cases where no change in the characteristics of the channel is observed and no codebook update is required, the proposed TCQ scheme has the same number of $B \cdot M_t$ feedback bits as the conventional TCQ scheme. The steps of codebook design are presented as follows:

- Step 1: For generating L candidate constellation sets, P channels for $\bar{h}$ are generated by simulation to perform quantization for source constellation (Section IV-A), set partitioning (Section IV-B), and the l-th candidate constellation set $S_p^{(l)}$, the corresponding count value $c^{(l)}$ and the minimum distance $\delta^{(l)}$ for $l = 1, \cdots, L$ are generated.
- Step 2: After setting $c^{(l)}$ to 0 where $\delta^{(l)}$ is less than a specific threshold value α for $l = 1, \cdots, L$, $\tilde{S}_p^{(l)}$ sorted by $S_p^{(l)}$ in descending order based on $c^{(l)}$ are created. In general, α is set to $\sqrt{2}d_{\min}$, but it is recommended to set α as large as possible.
- Step 3: Considering the additionally required number of bits F, a codebook W_F including the first to 2^F-th of $\tilde{S}_p^{(l)}$ is generated.

By exploiting W_F, it is possible not only to reduce the overhead of transmitting the source constellation, but also to drastically reduce the computational complexity. The proposed TCQ scheme can be operated by searching only the constellation set in W_F rather than performing K-means clustering on each channel. Therefore, except for the channel training, the proposed TCQ scheme has the similar computational complexity to the conventional TCQ scheme.

D. OVERALL OPERATION PROCESS

The proposed TCQ scheme first performs quantization for source constellation (Section IV-A), set partitioning (Section IV-B), and source constellation codebook design (Section IV-C) during the channel training phase to generate codebook W_F, and the generated codebook W_F is shared by BS and UE in advance as shown in Fig. 5. Then, UE performs channel feedback with an initially selected source constellation in the codebook W_F. At this time, the number of feedback bits is $B_{tot}^{prop} = B \cdot M_t + F$. While the channel environment does not change, the channel feedback is performed with a fixed source constellation as shown in Fig. 1, so the number of feedback bits is $B_{tot}^{prop} = B \cdot M_t$. If it is necessary to change to another source constellation due to a change in the channel environment, etc., the channel feedback is performed with the updated source constellation, so the number of feedback bits is $B_{tot}^{prop} = B \cdot M_t + F$. UE can select and update the optimal source constellation among 2^F source constellations, and operations for 2^F source constellations can be performed in parallel.

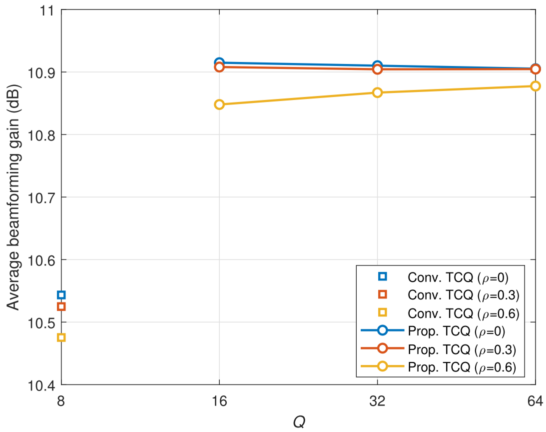


FIGURE 6. **Average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$, $B = 2$, $F = 6$ and $T = 3$ according to Q.**

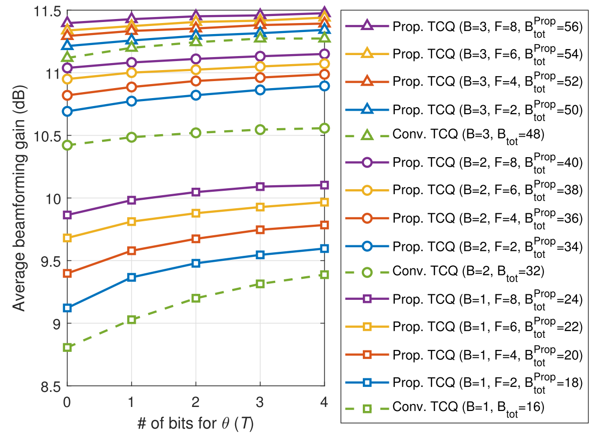


FIGURE 7. **Average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$ and $Q = 32$ in i.i.d. Rayleigh fading channels.**

V. SIMULATION RESULTS

In this section, the simulation results are presented to verify the benefits of the proposed TCQ scheme. We compare the performance of the proposed TCQ scheme with the conventional TCQ scheme [10], RVQ scheme [7], and unquantized beamforming. The average beamforming gain is used as the performance metric and is expressed in a decibel scale $10\log_{10}\left(E\left[\mathbf{hf}^H\right]\right)$. For the RVQ scheme, the analytic beamforming gain described as $M_t\left(1-2^{-\frac{B_{tot}}{M_t-1}}\right)$ is used, since it is computationally infeasible to implement as the number of feedback bits increases [7]. In the unquantized beamforming, the beamforming vector $\mathbf{f}$ is calculated as $\mathbf{h}/\|\mathbf{h}\|_2$. For the simulation parameters, the codebook W_F is generated in i.i.d. Rayleigh fading channel regardless of temporal and spatial corrrelation by considering $P = 10,000$, and the Ungerboeck rate $B/(B+1)$ convolutional codes in [12] are considered for the trellis. The Jakes' model [15] is considered in temporally correlated channels only for Fig. 10 to generate the temporal corrrelation coefficient $\eta = J_0\left(2\pi f_D \tau\right)$, where $J_0\left(\cdot\right)$ is the zeroth-order Bessel function of the first kind, f_D is the maximum Doppler frequency, and τ is the channel instantiation interval. A carrier frequency of 2.5 GHz and $\tau = 0.5ms$ are considered in this paper.

In the figures, we refer the proposed and conventional TCQ scheme as "Prop. TCQ" and "Conv. TCQ", respectively. Similarly, for the differential scheme, we refer to the proposed and conventional differential TCQ scheme as "Prop. Diff. TCQ" and "Conv. Diff. TCQ", respectively.

Fig. 6 shows the average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$, $B = 2$, $F = 6$ and $T = 3$ according to Q. When B is 2, the proposed TCQ scheme ($Q = 16, 32, 64$) shows the improved average beamforming gains compared to the conventional TCQ scheme ($Q = 8$) by using Q of 16 or more. In the proposed TCQ scheme, it can be seen that the average beamforming gains gradually decrease as Q increases when ρ is 0, while the average beamforming gains gradually improve as Q increases when ρ is 0.6. We consider 32 among 16, 32, and 64 as Q of the proposed TCQ scheme with the moderate average beamforming gains in environments with $\rho = 0$ and $\rho = 0.6$.

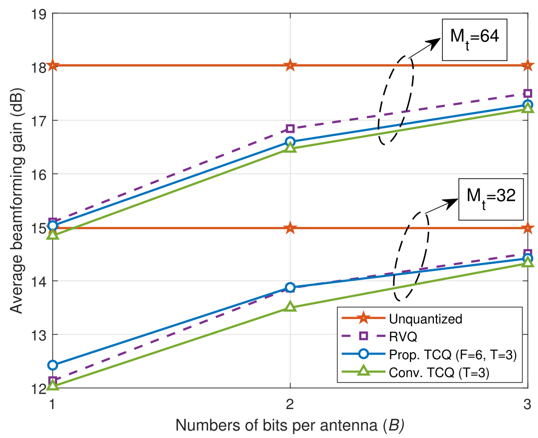


FIGURE 8. **Average beamforming gains of various channel feedback schemes according to B and M_t with $\rho = 0.6$.**

Fig. 7 shows the average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$ and $Q = 32$ in i.i.d. Rayleigh fading channels ($\rho = 0$). The average beamforming gains of the proposed and conventional TCQ schemes improve as B and T increase, and in particular, the average beamforming gains of the proposed TCQ scheme significantly improves as F increases for a given B. However, as B increases, the improvement of the average beamforming gains gradually decreases with the increase of T or F. With a small number of additional bits with $F = 8$, the proposed TCQ scheme with $T = 4$ provides larger beamforming gains of 0.7dB, 0.6dB, and 0.2dB over conventional TCQ scheme with $T = 4$ for 1, 2, and 3 of B, respectively. While the proposed TCQ scheme has the improved average beamforming gains compared to the conventional TCQ scheme, it may cause higher computational complexity by searching for more source constellation sets. For selecting the source constellation set of TCQ, the proposed TCQ scheme searches for $F \cdot 2^T$ constellation sets, while the

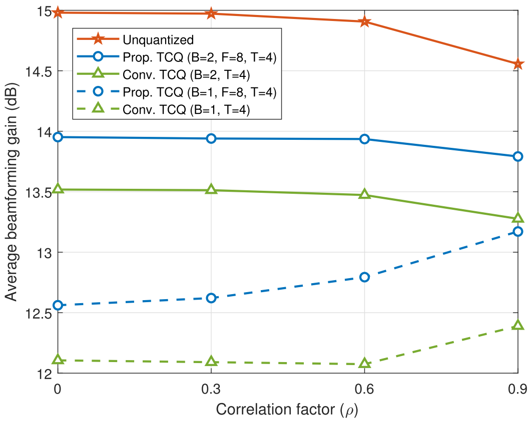


FIGURE 9. Average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 32$ according to ρ.

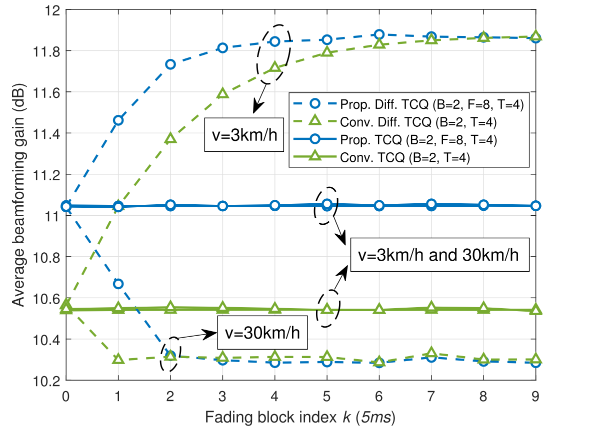


FIGURE 10. Average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$ according to fading block index k in temporally correlated channels with $v = 3km/h$ and $30km/h$.

conventional TCQ scheme searches for 2^T constellation sets. Thus, when T is fixed, the proposed TCQ scheme has approximately F times more decoding computational complexity than the conventional TCQ scheme, except for the codebook generation process for channel training. When comparing the proposed TCQ scheme $F = 8$ and $T = 1$ and the conventional TCQ scheme with $T = 4$ with the same complexity, the proposed TCQ scheme provides larger beamforming gains of 0.6dB, 0.5dB, and 0.15dB over conventional TCQ scheme for 1, 2, and 3 of B, respectively. Therefore, it can be seen that it is more efficient to perform TCQ using various constellation patterns suitable for the channel environments than to perform TCQ through a fixed source constellation such as 8-PSK and 16-QAM.

Fig. 8 shows the average beamforming gains of various channel feedback schemes according to B and M_t with $\rho = 0.6$. We set the proposed TCQ and RVQ schemes to have $B \cdot M_t + F$ feedback bits and the conventional TCQ scheme to have $B \cdot M_t$ feedback bits when F is fixed to 6. The TCQ and RVQ schemes approach the optimal unquantized beamforming performance as B increases, and when B is 3, the TCQ and RVQ scheme have a beamforming gain loss within 1dB compared to the unquantized beamforming. The proposed TCQ scheme has the better average beamforming gain performance than the conventional TCQ scheme for all B. However, the improvement of the average beamforming gains between the proposed and conventional TCQ schemes gradually decreases with the increase of B and M_t. With $M_t = 32$, the proposed TCQ scheme gives larger beamforming gains of 0.4dB and 0.1dB over conventional TCQ scheme for 2 and 3 of B, respectively. Therefore, the proposed TCQ scheme can be seen as more effective in an environment where the number of feedback bits is limited compared to the conventional TCQ scheme. When the number of feedback bits is small, such as $M_t \leq 32$ and $B \leq 2$, the proposed TCQ scheme shows the similar or better average beamforming gain performance than the RVQ technique.

Fig. 9 shows the average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 32$ according to ρ. The proposed TCQ scheme has more robust beamforming gain performance than the conventional TCQ scheme in terms of ρ. When B is 2, the TCQ schemes experience some beamforming gain degradation with $\rho > 0.6$, while the TCQ techniques do not experience significant beamforming gain degradation with $\rho \leq 0.6$. The proposed TCQ scheme shows the beamforming gain loss compared to the unquantized beamforming within 1dB when ρ is more than 0.6. On the other hand, when B is 1, the TCQ techniques improve performance as ρ increases. Since the channel elements are concentrated in a specific part as the ρ increases, the TCQ schemes with $B = 1$ could improve beamforming performance gains even with a small number of constellations. The conventional TCQ scheme has a slight improvement in beamforming gain performance only when $\rho > 0.6$, while the proposed TCQ scheme shows a significant performance improvement as ρ increases, and even the performance of the proposed TCQ scheme is close to that of the conventional TCQ scheme with $B = 2$. Therefore, it can be seen that the proposed TCQ scheme is suitable for channel environments with strong spatial correlation compared to the conventional TCQ scheme.

Fig. 10 shows the average beamforming gains of the conventional and proposed TCQ schemes with $M_t = 16$ according to fading block index k with $\tau = 0.5ms$ in temporally correlated channels with $v = 3km/h$ and $30km/h$. For the comparison, the differential TCQ scheme in [11] is considered. The differential scheme can be applied not only to the conventional TCQ scheme but also to the proposed TCQ scheme, and the average beamforming gain performance of these differential TCQ schemes is compared in temporally correlated channels. The same source constellations are considered for the TCQ and differential TCQ schemes in the conventional and proposed cases, respectively. The differential TCQ schemes improve the average beamforming gains over time in slow time-varying channels with $v = 3km/h$, while the differential TCQ schemes cause the serious beamforming gains degradation in fast time-varying channels with $30km/h$. It can be seen that the differential TCQ schemes, which feeds back only the difference between time-varying channels, shows the limitation in the average beamforming gains in fast time-varying channels. The proposed differential TCQ scheme converges faster on the saturated beamforming gain performance than the conventional differential TCQ scheme in slow time-varying channels with $v = 3km/h$ and deteriorates the beamforming gain performance more slowly than the conventional differential TCQ scheme in slow time-varying channels with $v = 30km/h$. Regardless of the influence of the time-varying channel, the TCQ schemes without the differential scheme have the similar beamforming gain performance over time. In addition, the proposed TCQ scheme has the beamforming gain of about 0.5dB compared to the conventional TCQ scheme in time-varying channels with $v = 3km/h$ and $v = 30km/h$. Therefore, it can be seen that the differential TCQ scheme is not appropriate in fast time-varying channels, and the proposed TCQ scheme is more suitable in time-varying channels than the conventional TCQ scheme.

VI. CONCLUSION

In this paper, a codebook-based TCQ scheme using K-means clustering was proposed for massive MIMO systems. In the proposed TCQ scheme, a codebook was considered to use a subset of constellation points suitable for the channel environments as the source constellation to achieve high beamforming gain performance, at the cost of a slight increase in feedback overhead. By operating codebook-based TCQ, the proposed TCQ scheme significantly improved beamforming gain without increasing computational complexity compared to the conventional TCQ scheme, except for the training process for codebook generation in preparation phase. The benefit of proposed TCQ scheme was verified using simulations, and it was observed that the proposed TCQ scheme with a similar computational complexity shows the improved average beamforming gains compared to the conventional TCQ scheme.

REFERENCES

[1] Y. Kim, Y. Kim, J. Oh, H. Ji, J. Yeo, S. Choi, H. Ryu, H. Noh, T. Kim, F. Sun, Y. Wang, Y. Qi, and J. Lee, "New radio (NR) and its evolution toward 5G-advanced," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 2–7, Jun. 2019.

[2] *Study on Scenarios and Requirements for Next Generation Access Technologies*, document TR 38.913 v15.0.0, 3GPP, Sep. 2018.

[3] *New SID on Physical Layer Enhancements for NR URLLC*, document RP-182089, TSG-RAN-81, Huawei, HiSilicon, Nokia, Nokia Shanghai Bell, Gold Coast, QLD, Australia, 2018.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[6] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels Techniques and Standards for Multi-Antenna Multi-User and Multi-Cell Systems*. San Diego, CA, USA: Academic, 2013.

[7] D. Love, R. Heath, V. N. Lau, D. Gesbert, B. 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.

[8] C. Au-Yeung and D. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.

[9] C. K. Au-Yeung, D. J. Love, and S. Sanayei, "Trellis coded line packing: Large dimensional beamforming vector quantization and feedback transmission," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1844–1853, Jun. 2011.

[10] 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.

[11] J. Mirza, M. Shafi, P. J. Smith, and P. A. Dmochowski, "Limited-feedback massive MISO systems with trellis-coded quantization for correlated channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8240–8254, Oct. 2016.

[12] S. Park, H. Noh, and C. Lee, "A differential trellis-coded quantization for FDD massive MIMO systems in a spatially correlated channel," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 106–110, Jan. 2017.

[13] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

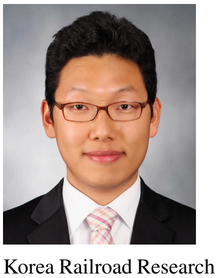
[14] R. H. Etkin and D. N. C. Tse, "Degrees of freedom in some underspread MIMO fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1576–1608, Apr. 2006.

[15] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Signal Process. Mag.*, vol. SPM-30, no. 1, pp. 40–60, Jan. 1982.

[16] T. Niknam, E. T. Fard, N. Pourjafarian, and A. Rousta, "An efficient hybrid algorithm based on modified imperialist competitive algorithm and K-means for data clustering," *Eng. Appl. Artif. Intell.*, vol. 24, no. 2, pp. 306–317, Mar. 2011.

[17] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[18] J. G. Proakis, *Digital Communication*, 4th ed., New York, NY, USA: McGraw-Hill, 2000.



SUNGSOO PARK (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Republic of Korea, in 2006, 2008, and 2012, respectively. From 2010 to 2011, he was a Visiting Scholar at Purdue University, West Lafayette, IN, USA. From 2012 to 2013, he was a Postdoctoral Researcher at Information and Telecommunication Laboratory (ITL), Yonsei University. Since 2013, he has been working with Korea Railroad Research Institute, where he is currently working as a Senior Researcher. His research interests include wireless communication systems for autonomous train control, future railway mobile communication systems, mission-critical communications and services, and vehicle-to-everything communications.

GYUYEOL KONG (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from the School of Electrical and Electronic Engineering, Yonsei University, Seoul, South Korea, in 2007 and 2014, respectively. From 2014 to 2016, he was a Senior Engineer at Samsung Electronics Company Ltd., Hwaseong-si, South Korea, where he was involved in the development of signal processing algorithms and error correction codes for NAND Flash memory. From 2016 to 2019, he was a Research Professor at the School of Electrical and Electronic Engineering, Yonsei University, where his research interests include the 5G and 6G wireless communication technologies based on artificial intelligence. Since 2019, he has been working as a Postdoctoral Researcher with the Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland. He is currently working as an Assistant Professor with the Division of Mechanical and Electronics Engineering, Hansung University, Seoul. His research interests include signal processing, wireless communications, and machine learning.