Modulation Diversity in Fading Channels with a Quantized Receiver

Saif Khan Mohammed, *Member, IEEE,* Emanuele Viterbo, *Fellow, IEEE*
Yi Hong, *Senior Member, IEEE,* and Ananthanarayanan Chockalingam, *Senior Member, IEEE*

***Abstract*—In this paper, we address the design of codes which achieve modulation diversity in block fading single-input single-output (SISO) channels with signal quantization at the receiver. With an unquantized receiver, coding based on algebraic rotations is known to achieve maximum modulation coding diversity. On the other hand, with a quantized receiver, algebraic rotations may not guarantee gains in diversity. Through analysis, we propose specific rotations which result in the codewords having equidistant component-wise projections. We show that the proposed coding scheme achieves maximum modulation diversity with a low-complexity minimum distance decoder and perfect channel knowledge. Relaxing the perfect channel knowledge assumption we propose a novel channel training/estimation technique to estimate the channel. We show that our coding/training/estimation scheme and minimum distance decoding achieves an error probability performance similar to that achieved with perfect channel knowledge.**

***Index Terms*—Modulation diversity, fading, quantization.**

I. Introduction

In practical communication receivers, the analog received signal is quantized into a finite number of bits for further digital baseband processing. With increasing bandwidth requirements of modern communication systems, analog-to-digital converters (ADC) are required to operate at high frequencies. However, at high operating frequencies, the precision of ADC's is limited [1]. Limited precision generally leads to high quantization noise, which degrades performance. In case of fading channels, floors in the bit error performance have been reported, and it seems difficult to avoid this behavior [2], [3]. On the other hand, channel capacity results show that even with 2-bit quantizers, the capacity of a quantized output channel is not far from that of a channel with unquantized output [4], [5]. Therefore, there appears to be a gap between the theoretical limits of communication with quantized receivers, and the current state of art.

In communication systems with fading, an important performance metric is the diversity order of reception. For single antenna fading scenarios, modulation diversity is a well known signal space diversity technique to improve the reliability/diversity of reception [6], [7]. However, with a quantized receiver, this coding alone *does not* guarantee improvement in diversity.

In this paper, we propose 2-dimensional constellations rotated by an angle θ, which can achieve full modulation diversity with a quantized receiver. With a quantized receiver, the maximum likelihood (ML) decoder is not the usual minimum distance decoder and would be much more complex to implement. We therefore assume a minimum distance decoder operating on the quantized channel outputs. We observe that, with a quantized receiver, *i*) for a given rate of information transmission in bits per channel use, there is a minimum requirement on the number of quantization bits, without which floors[1] appear in the error probability performance, *ii*) there is only a small subset of *admissible* rotation angles which can guarantee diversity improvement and no error floors, and *iii*) for a quantized receiver with perfect channel knowledge and minimum distance decoding, we analytically show that, among all *admissible* rotation angles, a good choice is one in which the transmitted vectors have *equidistant projections* along both the transmitted components. We then show that the M^2-QAM constellation rotated by $\theta = \tan^{-1}(1/M)$ has equidistant projections.

Further, we relax the perfect channel knowledge assumption and propose novel training sequences and a channel estimation scheme, which achieves an error probability performance close to that achieved with perfect channel knowledge. Through Monte-Carlo simulations we show that even with coarse analog-to-digital conversion, and short training sequences, the error performance with the estimated channel is similar to that with perfect channel knowledge. The main *interesting* result is that, even when the channel estimate is not perfect, an error probability performance exactly the same as that with perfect channel estimate is achievable under some sufficiency conditions on the channel estimate and the number of quantization bits. These conditions are analytically derived and shown to be

[1]Error probability performance is said to *floor*, if and only if it converges to a non-zero positive constant as the signal-to-noise ratio tends to infinity.

Manuscript received April 7, 2011; revised July 27, 2011; accepted September 27, 2011. The associate editor coordinating the review of this paper and approving it for publication was A. Sezgin.

The work of S. K. Mohammed was supported by the Swedish Foundation for Strategic Research (SSF) and the Excellence Center at Linköping-Lund in Information Technology (ELLIIT). The work of E. Viterbo and Y. Hong was supported in part by the Monash Professorial Fellowship 2011-2013. The work of S. K. Mohammed and A. Chockalingam was supported in part by the DRDO-IISC Program on Advanced Research in Mathematical Engineering. The material in this paper was presented in part at the International Symposium on Information Theory, Russia, July 2011.

S. K. Mohammed is with the Communication Systems Division at the Dept. of Electrical Eng. (ISY), Linköping University, 581 83 Linköping, Sweden (e-mail: saif@isy.liu.se).

E. Viterbo and Y. Hong are with the Dept. of Electrical and Computer Systems Eng., Monash University at Clayton, Melbourne, Victoria 3800, Australia (e-mail: {emanuele.viterbo, yi.hong}@monash.edu.

A. Chockalingam is with the Dept. of Electrical Communication Eng. (ECE) Indian Institute of Science, Bangalore 560012, India (e-mail: achockal@ece.iisc.ernet.in).

Digital Object Identifier 10.1109/TWC.2011.111211.110630

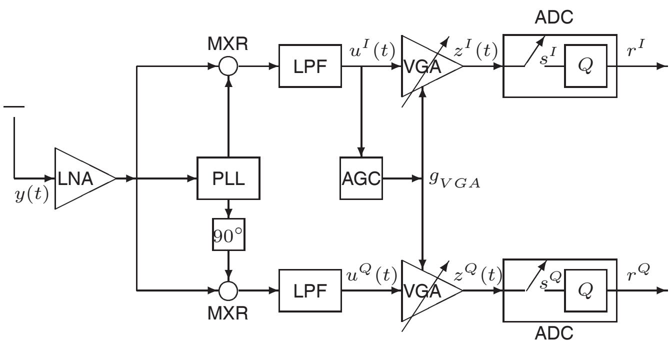

Fig. 1. Receiver analog front end (AFE) .

satisfied by the proposed training/estimation scheme for some specific scenarios. Another interesting result is that, with sufficient number of quantization bits, the error performance never floors *irrespective* of the quality of channel estimate. Also, to the best of our knowledge, the issue of achieving modulation diversity with quantized receivers having imperfect channel knowledge *has not been* addressed before in literature.

II. System Model and Quantized Receiver

We consider SISO block fading channels with a single transmit and a single receive antenna. The channel gains are assumed to be quasi-static for the coherence interval of the channel, and change to an independent realization in the next coherence interval. We assume frequency flat fading, with the complex channel frequency response given by

$$H(f) = |h|e^{-j2\pi\tau f} \;,\; |f - f_c| \leq \frac{W}{2}, \qquad (1)$$

and zero elsewhere, i.e., scaling by $|h|$, and a delay of τ seconds[2]. Figure 1 shows the signal path of a typical heterodyne receiver [8]. Due to space constraints we are unable to present the details of the continuous-time received signal model, for which we refer the reader to [11].

Prior to the transmission of information symbols, there is a training phase in which a known preamble sequence (carrier of constant amplitude A) is transmitted to enable carrier frequency synchronization in the receiver and also for tuning the receiver gain. Let the combined gain of the analog front end (AFE) (consisting of Low Noise Amplifier (LNA), Mixer (MXR) and Low Pass Filter (LPF)) be denoted by g_{AFE}. In the training phase, after the Phase Locked Loop (PLL) has locked, the LPF output is digitized using a Nyquist rate sample & hold type analog-to-digital converter (ADC), as shown in Fig. 1.

Let the input dynamic range of the ADC[3] be $-c_q/2$ to $c_q/2$. For optimum performance, it is desirable that the range of the input signal to the ADC matches with its dynamic range (ADC range matching). Due to fading, the input level at the ADC may vary, and therefore a variable gain amplifier (VGA) is generally used to ensure ADC range matching. The gain of the VGA is controlled by the automatic gain control (AGC) module [8]. During the training phase, AGC detects the peak of the LPF output signal using a conventional analog peak detector whose output is given by

$$V_{agc-pk} = A g_{AFE} |h|. \qquad (2)$$

Let X denote the peak value of the transmitted symbols (both real and imaginary component), during normal information transmission phase. During information transmission phase, ADC range matching (i.e., $\frac{c_q}{2} = g_{VGA} g_{AFE} |h| X$) requires the VGA gain to be

$$g_{VGA} = \frac{c_q}{2} \frac{A}{X} \frac{1}{V_{agc-pk}}. \qquad (3)$$

Since the ratio A/X and $c_q/2$ are known *a priori*, this computation can be performed in the AGC using simple analog circuits [9]. In the rest of the paper, we assume that this computation is perfect.

During the information transmission phase, PLL tracking is turned off and VGA gain is frozen to the value given by (3). Subsequently, without loss of generality, we assume $c_q/2 = 1$. Assuming perfect timing synchronization, the k-th output of the sample & hold circuit, at time $t = \tau + kT$ is given by[4]

$$s_k^I = \frac{x_k^I}{X} + \frac{w_k^I}{|h|X} \quad , \quad s_k^Q = \frac{x_k^Q}{X} + \frac{w_k^Q}{|h|X}, \qquad (4)$$

where $x_k = x_k^I + j x_k^Q$ is the k-th transmitted information symbol. The additive noise components w_k^I and w_k^Q are i.i.d. Gaussian random variables with variance denoted by $\sigma^2/2$. Let the average transmit power be denoted by $P_T \triangleq \mathbb{E}[|x_k|^2]$. Then the instantaneous signal-to-noise ratio (SNR) at the output of the sample & hold circuit is given by $\gamma_{inst} \triangleq P_T |h|^2/\sigma^2$. Assuming a Rayleigh fading model with $h \sim \mathcal{CN}(0,1)$ (complex Gaussian with zero mean and unit variance), the average SNR is given by $\gamma \triangleq \mathbb{E}_h[\gamma_{inst}] = P_T/\sigma^2$. The output of the sample & hold circuit is then quantized by a b-bit uniform quantizer Q, as shown in Fig. 1. The quantizer is modeled by the function $Q_b(t), t \in \mathbb{R}$, which is given by

$$Q_b(t) = \begin{cases} +1, & \xi(t) \geq (2^{b-1} - 1) \\ -1, & \xi(t) \leq -(2^{b-1} - 1) \\ \frac{(2\xi(t)+1)}{2^b - 1}, & \text{otherwise} \end{cases} \qquad (5)$$

$$\xi(t) \triangleq \left\lfloor \frac{t\,(2^b - 1)}{2} \right\rfloor \qquad (6)$$

where $\lfloor x \rfloor$ denotes the largest integer not greater than x. For a n-dimensional complex vector $\mathbf{z} = (z_1, z_2, \cdots, z_n)$, let $\mathbf{Q}_b(\mathbf{z})$ denote the n-dimensional component-wise quantized version of $\mathbf{z}$. That is, $\tilde{\mathbf{z}} = (\tilde{z}_1, \tilde{z}_2, \cdots, \tilde{z}_n) = \mathbf{Q}_b(\mathbf{z})$ implies that

$$\tilde{z}_i^I = Q_b(z_i^I) \;,\; \tilde{z}_i^Q = Q_b(z_i^Q) \quad i = 1, 2, \ldots, n. \qquad (7)$$

The k-th quantized received symbol, $r_k = r_k^I + j r_k^Q$ is therefore given by

$$r_k^I = Q_b(s_k^I) \,,\, r_k^Q = Q_b(s_k^Q) \qquad (8)$$

where s_k^I and s_k^Q are the real and imaginary components of the k-th sample & hold output symbol.

Since achieving modulation coding diversity with quantized output is the main subject of this paper, we next give a brief

[2] W is the signaling bandwidth and $f_c \gg W$ is the carrier frequency.

[3] We also refer to $c_q/2$ as the clip level, since any input greater than $c_q/2$ would be limited to $c_q/2$.

[4] $1/T$ is the rate at which information symbols are transmitted.

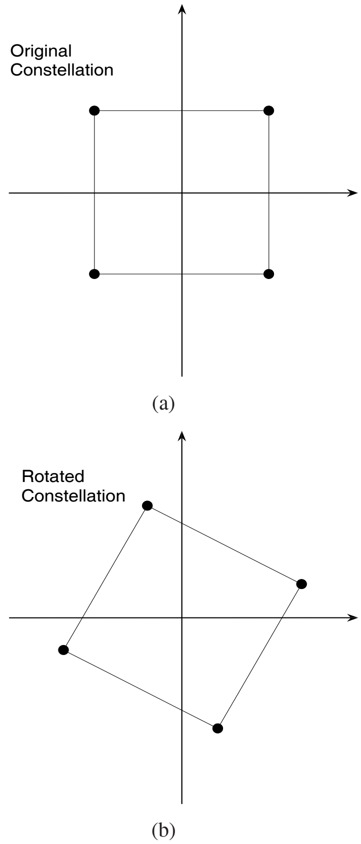

Fig. 2. Signal space of the transmitted constellation consisting of four codewords in 2 dimensions (horizontal and vertical). The average transmit power constraint is $P_T = 1/4$. For the rotated constellation, the rotation angle is $28.5°$. The codewords are represented by solid dots.

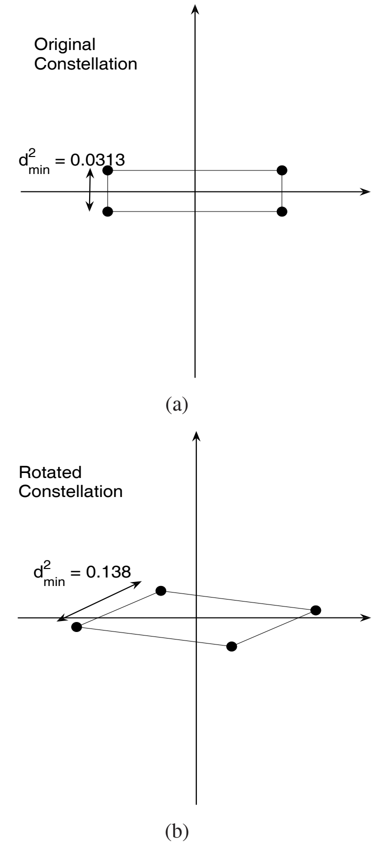

Fig. 3. Signal space of the received 2-D codewords. The gains along the horizontal and vertical signal component/dimension are $|h_1| = 1$ and $|h_2| = 1/4$ respectively. d^2_{min} is the minimum Euclidean distance between any two received codewords.

introduction to modulation coding/signal space diversity in SISO fading channels with an unquantized receiver [7]. We illustrate the signal space diversity technique through Fig. 2 and Fig. 3. We consider two possible transmission schemes. In the first scheme, the original 2-dimensional constellation (with four codewords) is transmitted as it is (Fig. 2(a)), whereas in the second scheme the original constellation is rotated before transmission (Fig. 2(b)). For a given 2-D codeword chosen for transmission, the two components of the chosen codeword are transmitted separately during two different coherence time intervals with channel gains $|h_1|$ and $|h_2|$ respectively. From Fig. 3, it is observed that with a total transmit power constraint of $P_T = 1/4$, the minimum distance between the received codewords is larger when the rotated constellation is transmitted. Under deep fading along a certain signal dimension/component (i.e., channel gain ≈ 0 along this component), the received codeword is essentially the projection of the transmitted codewords onto the other signal dimension. With no rotation, some of the transmitted codewords have the same projection onto a given signal component, and therefore the minimum distance would be approximately 0. However with a rotated constellation, since the projections of the codewords are distinct, the minimum distance between received codewords is strictly positive even when the channel gain along a certain component is zero. Therefore, when the rotated constellation is transmitted, an error event (i.e., very small minimum distance) occurs only when the channel gain along *both* the signal components are small, thereby implying second order diversity. In contrast, with no rotation, an error event occurs if the channel gain along *any one* of the signal components is small, thereby implying only single order diversity.

The example above was with $n = 2$ signal components/dimensions. For a general n-dimensional signal space, modulation diversity coding is illustrated in Fig. 4. Coding is performed across $n > 1$ information symbols resulting in n coded symbols/codeword. These n coded symbols are interleaved and then transmitted over n independent channel coherence intervals (realizations). At the receiver, the channel outputs during the n coherence intervals are buffered, followed by de-interleaving and detection. Suitable coding across n independent channel realizations results in an n-fold increase in the diversity of reception. In fading channels, codes designed using algebraic lattices can achieve modulation diversity, and are therefore employed to improve the diversity of reception [6]. With an unquantized receiver, it is known that lattice codes based on algebraic rotations can achieve full modulation diversity [7][10]. *However, with quantized receivers, this is no longer true.*

In this paper we consider the case of $n = 2$. Let the information symbol vector be denoted by $\mathbf{u} = (u_1, u_2)^T$, where u_1 and u_2 are restricted to M^2-QAM, though a generalization to non-square QAM is trivial. Let the set $\mathcal{S}_M = \{-(M-1), \ldots, -1, 1, \cdots, (M-1)\}$ denote the M-PAM signal set. Then, M^2-QAM is denoted by the set

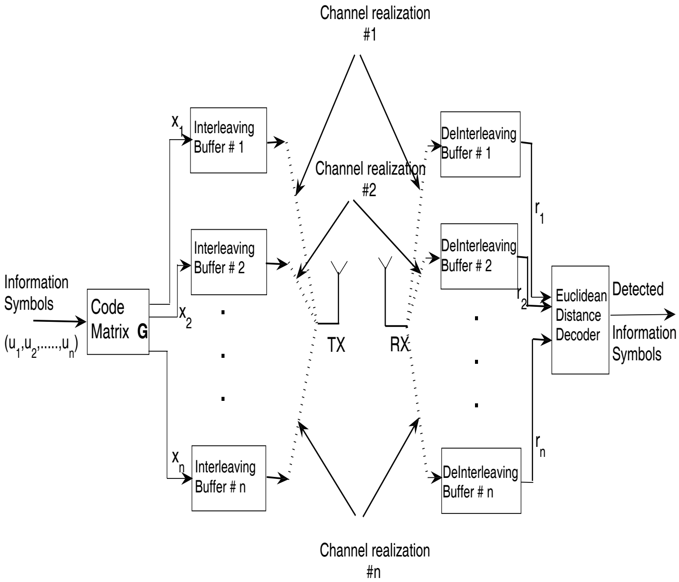

Fig. 4. Achieving modulation diversity by coding across n different channel realizations.

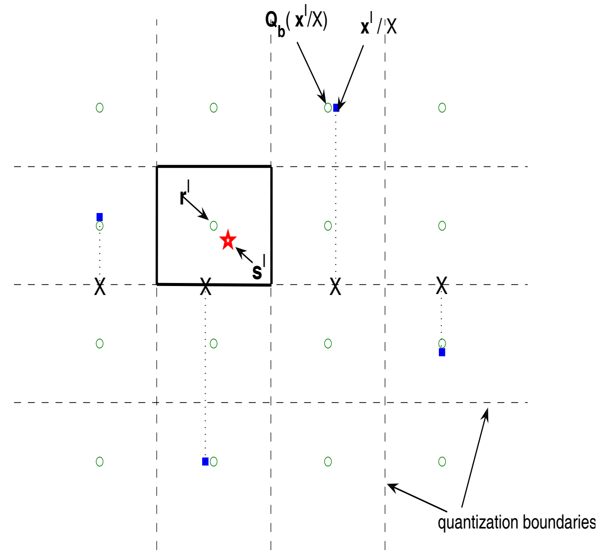

Fig. 5. Signal space at the quantizer input with $b = 2$ (real component). Rotated 4-QAM ($\theta = 20°$) depicted with dark filled squares.

$\mathcal{S}_M^2 \triangleq \{w + jv \mid w, v \in \mathcal{S}_M\}$. The information symbols are coded using a 2×2 rotation matrix $\mathbf{G}$, resulting in the transmit vector $\mathbf{x} = (x_1, x_2)^T = \mathbf{Gu}$, where

$$\mathbf{G} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{9}$$

Due to QAM symmetry, one can restrict the rotation angle in (9) to $[0, \pi/4)$. The set of transmitted vectors $\mathcal{X}$ and the peak component value X are given by

$$\begin{aligned} \mathcal{X} &= \left\{ \mathbf{x} \mid \mathbf{x} = \mathbf{Gu}, u_1, u_2 \in \mathcal{S}_M^2 \right\}, \\ X &= \max_{\mathbf{x}\in\mathcal{X}} \left\{ \max_{i=1,2} \left[\max(|x_i^I|, |x_i^Q|) \right] \right\} \end{aligned} \tag{10}$$

Also, let the channel gain during the transmission of x_1 and x_2 be denoted by $|h_1|$ and $|h_2|$, respectively. We assume h_1 and h_2 to be i.i.d. $\mathcal{CN}(0, 1)$. Let $\mathbf{r} = (r_1, r_2)^T$ denote the quantized received vector, where $r_1 = r_1^I + jr_1^Q$ and $r_2 = r_2^I + jr_2^Q$ are the ADC outputs during the transmission of x_1 and x_2, respectively. From (4) and (8) it follows that

$$r_i^I = Q_b\Big(\frac{x_i^I}{X} + \frac{w_i^I}{|h_i|X}\Big),\, r_i^Q = Q_b\Big(\frac{x_i^Q}{X} + \frac{w_i^Q}{|h_i|X}\Big). \tag{11}$$

With the above quantized receiver model, ML decoding is no more given by the minimum distance decoder, and is rather complex. Nevertheless, due to its lower decoding complexity, we shall assume a minimum distance decoder taking $\mathbf{r}$ as its input, and the output (detected information symbols) given by[5]

$$\begin{aligned} \hat{\mathbf{u}} &= \arg\min_{\mathbf{u}\in\mathcal{S}_M^2\times\mathcal{S}_M^2} \Big\| \mathrm{diag}(|h_1|, |h_2|)\Big(\mathbf{r} - \frac{\mathbf{Gu}}{X}\Big)\Big\|^2 \\ &= \arg\min_{\mathbf{u}\in\mathcal{S}_M^2\times\mathcal{S}_M^2} \left(\mathbf{r} - \Big(\frac{\mathbf{Gu}}{X}\Big)\right)^{\dagger} \mathbf{D}_\rho \left(\mathbf{r} - \Big(\frac{\mathbf{Gu}}{X}\Big)\right), \\ \mathbf{D}_\rho &\triangleq \mathrm{diag}(1, \rho^2) \end{aligned} \tag{12}$$

where $\rho \triangleq |h_2|/|h_1|$ is the channel gain ratio, and $\dagger$, $\|.\|$ denote Hermitian transpose and Euclidean norm respectively. In the subsequent sections III and IV, assuming perfect receiver knowledge of ρ, we show that even with the suboptimal minimum distance decoder in (12), we can avoid error floors and also achieve full modulation diversity.

III. Rotation Coding in Quantized Receiver

In this Section, we study the error performance of 2-dimensional rotated constellations (as defined in (10)) with a *quantized* receiver and minimum distance decoding, and derive the conditions under which full modulation diversity can be achieved. In case of a quantized receiver, the sample & hold outputs (4), are quantized to the appropriate *quantization box* containing it. As an example, Fig. 5 illustrates the rotated 4-QAM constellation with $\theta = 20°$. The dark filled squares represent the 4 possible values taken by the real component of the normalized transmit vector $\mathbf{x}^I/X = (x_1^I/X\,, x_2^I/X)^T$. The *projections* of the 4 possible vectors onto the first component (horizontal) are marked with a cross. A $b = 2$-bit quantizer is used along both codeword components. The dashed horizontal and vertical lines represent the quantization boundaries along the 2 components. As an example, in Fig. 5 the real component of the sample & hold output vector $\mathbf{s}^I = (s_1^I, s_2^I)^T$ (marked with a star), is therefore quantized to $\mathbf{r}^I = (r_1^I, r_2^I)^T$ (Note that there are totally 16 different quantized outputs marked with empty circles). The quantization box corresponding to the output $\mathbf{r}^I$ is shown in the figure as a square with solid lines.

As the noise variance $\sigma^2 \to 0$, the sample & hold output $\mathbf{s}$ is almost the same as the normalized transmitted vector $\mathbf{x}/X$. Therefore at sufficiently high SNR, if there exist two different

[5] The assumed minimum distance detector is essentially the ML detector for an unquantized receiver. With an unquantized receiver, the received signal model would be same as (11), but without the $Q(.)$ operator. For the unquantized receiver model, the conditional probability distribution of $\mathbf{r}$ given $\mathbf{x}$ is $p_{\text{un}}(\mathbf{r}|\mathbf{x}) = \prod_{i=1}^{2} \frac{|h_i|^2X^2}{\pi\sigma^2} e^{-\frac{|h_i|^2X^2|r_i - \frac{x_i}{X}|^2}{\sigma^2}}$. The minimum distance decoder that we use, is one which maximizes $p_{\text{un}}(\mathbf{r}|\mathbf{x})$ over all $\mathbf{x} \in \mathcal{X}$.

transmit vectors $\mathbf{x}$ and $\mathbf{y}$, such that $\mathbf{Q}_b(\mathbf{x}/X)$ and $\mathbf{Q}_b(\mathbf{y}/X)$ are identical, then it is obvious that the error probability performance would floor as SNR $\to \infty$. This is because, at high SNR the quantizer output would be the same irrespective of whether $\mathbf{x}$ or $\mathbf{y}$ was transmitted, which makes it impossible for the the receiver to distinguish between the two transmit vectors leading to erroneous detection. More formally, two transmit vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be *distinguishable* if and only if $\mathbf{Q}_b(\mathbf{x}/X) \neq \mathbf{Q}_b(\mathbf{y}/X)$. Therefore, in order to avoid floors in the error probability performance, we propose the first code design criterion.

Criterion I *: A necessary and sufficient condition to avoid error floors with a quantized receiver, is that any two transmit vectors must be distinguishable.*

To achieve full modulation diversity, it is required that even under deep fading conditions in one component, any two transmit vectors $\mathbf{x}$ and $\mathbf{y}$ must still be *distinguishable* in the other component. This, therefore, implies that the projections of all the transmit vectors onto any one component must be *distinguishable* by the quantizer in that component. Therefore, we have the second criterion.

Criterion II *: Given a b-bit quantized receiver, in order to achieve full modulation diversity, a necessary condition on the rotation angle θ is that, any two distinct transmit vectors $\mathbf{x}$ and $\mathbf{y}$ satisfy*

$$\mathbf{Q}_b(x_i/X) \neq \mathbf{Q}_b(y_i/X), \;\; i = 1, 2. \tag{13}$$

With a rotated M^2-QAM there are totally M^2 distinct projections onto any component, and therefore in order to achieve full modulation diversity, the minimum number of quantization bits required for the transmit vectors to be *distinguishable* along any component is at least $\lceil 2\log_2(M) \rceil$ [6] i.e.

$$b \geq \lceil 2\log_2(M) \rceil. \tag{14}$$

Subsequently, we assume that for a given M, b is fixed to the lower bound value in (14). We further note that, with a $b = \lceil 2\log_2(M) \rceil$-bit quantizer, Criterion II is not satisfied by all rotation angles[7].

With a $b = \lceil 2\log_2(M) \rceil$-bit quantizer, the set of angles (between 0 and $\pi/4$) which result in *distinguishable* projections along both the codeword components will be referred to as the *admissible* angles (i.e., angles which satisfy Criterion II). For example, with 4- and 16-QAM, the admissible angles lie in the range $(\tan^{-1}(1/5)\,,\,\pi/4)$ and $(11.3^\circ\,,\,16.9^\circ)$, respectively. With increasing M, the interval of admissible angles reduces. With 256-QAM, the range of admissible angles is only $(3.47^\circ\,,\,3.68^\circ)$. Another interesting fact is that, for M^2-QAM, $\theta = \tan^{-1}(1/M)$ is always in the set of admissible angles. Further, as M increases, $\tan^{-1}(1/M) \pm \epsilon$ are observed to be the only admissible angles.

Apart from the fact that the chosen angle must have distinguishable projections, it can be analytically shown that for M^2-QAM, any rotation angle for which the rotated constellation satisfies

$$\mathbf{Q}_b(\mathbf{x}/X) = \mathbf{x}/X \;,\;\; \mathbf{x} \in \mathcal{X} \tag{15}$$

achieves a diversity order of 2 (i.e., full modulation diversity since $n = 2$), with a $b = \lceil 2\log_2(M) \rceil$-bit quantized receiver and minimum distance decoding given by (12) (See Appendix A in [11]). Subsequently, a rotated constellation which satisfies (15) shall be referred as being *matched* to the quantizer. It is easy to see that a rotated M^2-QAM constellation is matched to a $b = 2\lceil \log_2(M) \rceil$-bit uniform quantizer, if and only if, the projections of the transmit vectors are component-wise *equidistant* and distinguishable.

Even with a mismatched rotated constellation having distinguishable projections (i.e., when the projections are not equidistant), full modulation diversity may be achieved, but then the error probability would be higher, since some transmit vectors would be closer to the edge of their quantization boxes (making it easier for noise to move the transmitted vector to another quantization box when received) (illustrated through Fig. 10 in Appendix A [11]). This therefore leads us to the third code design criterion.

Criterion III *: In order to minimize the error probability of a rotated M^2-QAM constellation with a $b = \lceil 2\log_2(M) \rceil$-bit quantized receiver, the rotation angle must be such that the rotated M^2-QAM constellation is matched to the quantizer.*

IV. Rotated Constellation Design for Quantized Receiver

In this section, we construct rotated M^2-QAM constellations which satisfy Criterion III. We next show that a rotation by $\theta = \tan^{-1}(1/M)$, satisfies Criterion III[8]. Since $\mathbf{G}$ is real-valued, it suffices to prove the equidistant projections property for the real component only. The information symbols u_1^I and u_2^I take values from the M-PAM signal set $\mathcal{S}_M$. It is clear that the pair (u_1^I, u_2^I) can take any of the M^2 values from the ordered sequence of values $\mathcal{S}_M^1 = \{(-M+1, -M+1), (-M+1, -M+3), \ldots (-M+1, M-1), (-M+3, -M+1), (-M+3, -M+3), \ldots (-M+3, M-1), \ldots (M-1, -M+1), (M-1, -M+3), \ldots (M-1, M-1)\}$. With $\theta = \tan^{-1}(1/M)$, the two rows of $\mathbf{G}$ are $(M, 1)$ and $(-1, M)$ scaled by $1/\sqrt{M^2+1}$. From basic algebra, it now follows that the value of the first component of the transmit vector i.e., x_1^I, increases in steps of $2/\sqrt{M^2+1}$ as (u_1^I, u_2^I) takes values sequentially from the set $\mathcal{S}_M^1$. This then proves that the projections along the first component are indeed equidistant. Also, since the values in the M-PAM signal set are symmetric around 0, it follows that the set of all M^2 values which the first component x_1^I takes, is same as the set of the M^2 values taken by the second component x_2^I. Hence the projections along the second component are also equidistant.

For M^2-QAM with $\theta = \tan^{-1}(1/M)$, it can be shown that the minimum product distance of the code is $4M/(M^2+1)$ ($\approx 4/M$ for $M \gg 1$) [6]. On the other hand, a rotation angle of $\theta = 1/2\tan^{-1}(2)$ is known to have a minimum product distance of at least $4/\sqrt{5}$ *irrespective of the QAM size* [6], [10]. Also, for any rotation angle the error performance with a quantized receiver is inferior to that with an unquantized receiver. Hence, with increasing M, the error performance of

[6] $\lceil x \rceil$ denotes the smallest integer not smaller than x.

[7] For example, even though $\theta = 1/2\tan^{-1}(2)$ guarantees a rotation code having non-vanishing minimum product distance, with a $b = 4$-bit uniform quantizer and $M^2 = 16$-QAM it *does not* satisfy Criterion II.

[8] We would like to make a note of the fact that, in [12], for a totally different system setting, the optimal rotation angle for M^2-QAM has been mentioned to be $\theta = \tan^{-1}(1/M)$.

a *quantized* receiver with $\theta = \tan^{-1}(1/M)$ is expected to be increasingly less power efficient than that of an *unquantized* receiver with $\theta = 1/2 \tan^{-1}(2)$. With increasing M, the set of admissible angles appeared to be only $\tan^{-1}(1/M) \pm \epsilon$ and therefore, it can be argued that, the best possible error performance with a $b = \lceil 2\log_2(M) \rceil$-bit quantized receiver would have a loss in power efficiency when compared to an unquantized receiver. However, this appears to be the cost to achieve full modulation diversity in quantized receivers with limited precision.

V. Imperfect Receiver Knowledge of ρ

In the previous sections, in order to achieve full modulation diversity, minimum distance decoding at the receiver assumed perfect knowledge of ρ. In this section, we relax this assumption and present novel techniques to estimate ρ accurately. It is expected that the error performance would degrade with imperfect estimate of ρ. Interestingly, in section V-A we propose an optimality criterion, which if satisfied by the estimate of ρ, would guarantee *no loss* in the error probability performance of the minimum distance decoder with estimated ρ when compared to the error performance with perfect knowledge of ρ. Such an estimate would be referred to as an *optimal* estimate of ρ. We estimate ρ based on the quantized receiver outputs for a known transmitted sequence. We refer to this sequence as the ρ-training sequence. Any ρ-training sequence which results in an optimal estimate of ρ is referred to as an optimal ρ-training sequence. In section V-B we present receiver control techniques required to estimate ρ. ML estimation of ρ based on the quantized receiver outputs of the ρ-training sequence is discussed in section V-C. Finally, in section V-D, for $M = 2$ (rotated 4-QAM) we present an optimal ρ-training sequence which satisfies the optimality criterion introduced in section V-A.

For $M > 2$, the length of ρ-training sequences which satisfy the optimality criterion is expected to be large resulting in too much training overhead and hence loss in effective throughput. Therefore, a novel design of short ρ-training sequences (referred to as 'good' ρ-training sequences) is proposed, which can achieve an error probability performance close to that achieved with optimal ρ-training sequences. Also throughout this section, it is assumed that *i*) with rotated M^2-QAM, a $b = \lceil 2\log_2(M) \rceil$-bit uniform quantizer is employed, *ii*) a minimum distance decoder is used for detection, and *iii*) the rotated constellation satisfies Criterion III.

A. Criterion for the optimal ρ estimate

In this Section, we are interested in studying the *conditions* under which the error probability performance with $\hat{\rho}$ (i.e., estimated ρ) is exactly the same as the error probability performance assuming perfect receiver knowledge of ρ. Any estimate of ρ, which satisfies these conditions would be an *optimal estimate* in terms of achieving an error probability performance same as that achieved with perfect receiver knowledge of ρ. The minimum distance decoder with the estimated ρ, is also given by (12), but with ρ replaced by its estimate $\hat{\rho}$. Further, since $\mathbf{G}$ is real-valued, it suffices to consider the error probability for only the real component of the transmitted information symbols. To simplify notations, for any received vector $\mathbf{r}$, information symbol vectors $\mathbf{u}$ and $\mathbf{v}$ and any real $\zeta > 0$, we define

$$m(\zeta, \mathbf{r}^I, \mathbf{u}^I) \triangleq \left(\mathbf{r}^I - \left(\frac{\mathbf{G}\mathbf{u}^I}{X}\right)\right)^T \mathbf{D}_\zeta \left(\mathbf{r}^I - \left(\frac{\mathbf{G}\mathbf{u}^I}{X}\right)\right) \quad (16)$$

$$D_E(\zeta, \mathbf{r}^I, \mathbf{u}^I, \mathbf{v}^I) \triangleq \left(m(\zeta, \mathbf{r}^I, \mathbf{u}^I) - m(\zeta, \mathbf{r}^I, \mathbf{v}^I)\right). \quad (17)$$

The detected information symbols with perfect knowledge of ρ can therefore be stated in terms of $m(.)$ as

$$\widehat{\mathbf{u}}^I = \arg \min_{\mathbf{u}^I \in \mathcal{S}_M^2} m(\rho, \mathbf{r}^I, \mathbf{u}^I). \quad (18)$$

This then implies that, for any information symbol vector $\mathbf{v}$

$$D_E(\rho, \mathbf{r}^I, \widehat{\mathbf{u}}^I, \mathbf{v}^I) \leq 0 \quad (19)$$

With an estimated $\hat{\rho}$, if for all information symbol vectors $\mathbf{v} \in \mathcal{S}_M^2$

$$D_E(\hat{\rho}, \mathbf{r}^I, \widehat{\mathbf{u}}^I, \mathbf{v}^I) \leq 0 \quad (20)$$

then it is obvious that the output of the minimum distance decoder with estimated ρ is the *same* as the output of the minimum distance decoder with perfect knowledge of ρ. If (20) holds for all information symbol vectors $\mathbf{v} \in \mathcal{S}_M^2$, then along with (19), it follows that

$$D_E(\rho, \mathbf{r}^I, \widehat{\mathbf{u}}^I, \mathbf{v}^I) D_E(\hat{\rho}, \mathbf{r}^I, \widehat{\mathbf{u}}^I, \mathbf{v}^I) \geq 0 \quad (21)$$

for all information symbol vectors $\mathbf{v} \in \mathcal{S}_M^2$. Since $\widehat{\mathbf{u}}^I$ could be any information symbol vector in $\mathcal{S}_M^2$, it is easy to see that the output of the minimum distance decoder with estimated $\hat{\rho}$ would be the same as that with perfect knowledge of ρ if

$$D_E(\rho, \mathbf{r}^I, \mathbf{u}^I, \mathbf{v}^I) D_E(\hat{\rho}, \mathbf{r}^I, \mathbf{u}^I, \mathbf{v}^I) \geq 0 \quad (22)$$

for all possible received vector $\mathbf{r}$ (finitely many due to receiver quantization) and all possible information symbol vectors $\mathbf{u}$ and $\mathbf{v}$. We formally prove this observation in the following theorem.

Theorem 5.1: For a given realization of ρ, and estimated $\hat{\rho}$, if (22) is satisfied for all possible received vector $\mathbf{r}$ and all possible information symbol vectors $\mathbf{u}$ and $\mathbf{v}$, then $\hat{\rho}$ is an *optimal* estimate of ρ.

Proof: See Appendix C in [11]. ■

We now analyze the condition set-forth in Theorem 5.1 regarding the *optimal* estimate of ρ. With each information symbol belonging to M^2-QAM, and $b = \lceil 2\log_2(M) \rceil$ we make the following definitions

$$\begin{aligned} \mathcal{D}_M &\triangleq \left\{ \frac{(a_1 - a_2)}{2^b - 1} \,|\, a_1, a_2 \in \mathcal{S}_{M^2} \right\}, \\ \mathcal{D}_M^2 &\triangleq \left\{ a^2 \,|\, a \in \mathcal{D}_M \right\}, \\ \mathcal{Q}_M &\triangleq \left\{ \frac{(a_1 - a_2)}{(a_3 - a_4)} \,|\, a_1, a_2, a_3, a_4 \in \mathcal{D}_M^2 \,,\, a_3 \neq a_4 \right\}, \\ \mathcal{Q}_M^+ &\triangleq \left\{ a \,|\, a \in \mathcal{Q}_M \,,\, a \geq 0 \right\} \end{aligned} \quad (23)$$

where $\mathcal{S}_{M^2}$ is the M^2-PAM signal set (see Section II for PAM set definition). It is also noted that $\mathcal{S}_{M^2}$ is not the same as $\mathcal{S}_M^2$. As an example, with $M = 2$ and $b =$

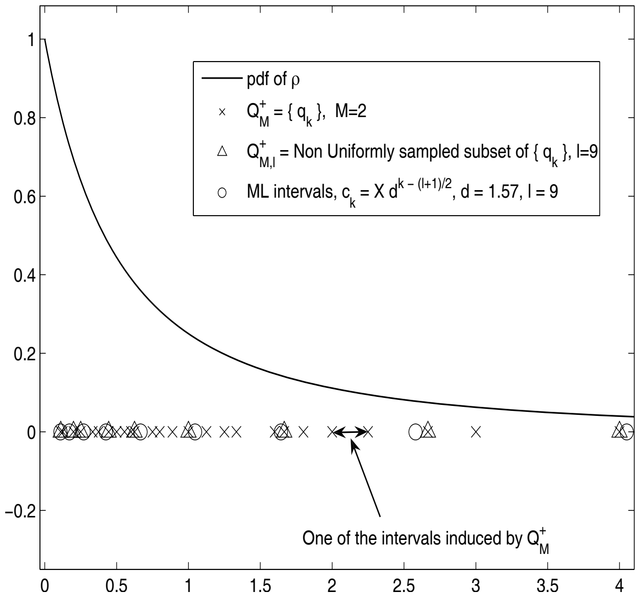

Fig. 6. The intervals induced by the set $\mathcal{Q}_M^+$ for $M = 2$. ML intervals for the proposed ρ-training and estimation schemes.

$\lceil 2\log_2(M)\rceil = 2$, $\mathcal{S}_{M^2} = \{-3,-1,1,3\}$ and $\mathcal{D}_M = \{-2,-4/3,-2/3,0,2/3,4/3,2\}$.

Essentially, $\mathcal{D}_M$ represents the difference set for the M^2-PAM signal set $\mathcal{S}_{M^2}$, normalized by the maximum amplitude M^2-1. $\mathcal{D}_M^2$ contains the squared elements of $\mathcal{D}_M$. Due to squaring, the elements in $\mathcal{D}_M^2$ are no more equally spaced, and are more densely distributed near the origin than farther away. We then consider the difference set for $\mathcal{D}_M^2$. The quotient set of this difference set is denoted by $\mathcal{Q}_M$. $\mathcal{Q}_M^+$ is simply the set of non-negative elements of $\mathcal{Q}_M$. Since, the distribution of elements in $\mathcal{D}_M^2$ is denser near the origin, the quotient set $\mathcal{Q}_M^+$ is also denser near the origin. As we shall see later, this property of the set $\mathcal{Q}_M^+$ will be exploited to design short length ρ-training sequences, which result in near-optimal estimates of ρ. An example of this set is shown in Fig. 6 for $M = 2$. Based on Theorem 5.1, the next theorem gives a useful sufficiency condition for the optimality of an estimate of ρ.

Theorem 5.2: Consider a rotated M^2-QAM constellation matched to the quantizer. Let $\hat{\rho}$ be an estimate of ρ satisfying the following condition

$$\begin{aligned} \forall\, l \in \mathcal{Q}_M^+ &: \rho^2 \leq l \Rightarrow \hat{\rho}^2 \leq l \\ \forall\, l \in \mathcal{Q}_M^+ &: \rho^2 \geq l \Rightarrow \hat{\rho}^2 \geq l \end{aligned} \tag{24}$$

Then, $\hat{\rho}$ is an *optimal* estimate of ρ.

Proof : See Appendix D in [11]. ■

Let $\mathcal{Q}_M^+ = \{0, q_1, q_2, \cdots, q_{L_M}\}$, with $0 < q_1 < q_2 < \cdots < q_{L_M}$. It is clear that the elements of the set $\mathcal{Q}_M^+$ *partition* the positive real line $[0,\infty)$ into (L_M+1) intervals with $[0, q_1)$ and $[q_{L_M},\infty)$ being the first and the last interval respectively. The k-th intermediate interval is given by $[q_{k-1}, q_k)$, $k = 2, 3, \cdots, L_M$. For any finite set $\mathcal{S} = \{s_1, s_2, \cdots, s_n\}$ with $0 \leq s_1 < s_2 < \cdots < s_n$, let $\mathcal{I}(\mathcal{S})$ be the set of intervals *induced* by the set $\mathcal{S}$. That is

$$\mathcal{I}(\mathcal{S}) = \Big\{[s_1, s_2)\,,\,[s_2, s_3)\,,\,\cdots\,,\,[s_n,\infty)\Big\} \tag{25}$$

The sufficiency condition in (24) can now be understood in terms of the (L_M+1) intervals of the positive real line induced by the set $\mathcal{Q}_M^+$. The sufficiency condition basically states that, for an estimate of ρ to be *optimal*, it *must* belong to the same interval of $\mathcal{I}(\mathcal{Q}_M^+)$ in which ρ lies. This therefore also implies that, for an estimate $\hat{\rho}$ to be *optimal* it is *not* necessary that $\hat{\rho}$ be exactly equal to ρ.

B. Receiver control for estimating ρ

In this section, we discuss receiver control techniques required for estimating ρ. In the proposed rotation coding scheme, coding is performed across $n = 2$ channel realizations. Using a 2-dimensional rotation matrix $\mathbf{G}$, a pair of information symbols is transformed into a pair of coded output symbols. The first coded symbol in the pair is transmitted during channel realization 1 (with channel gain $|h_1|$), whereas the second coded symbol is transmitted during channel realization 2 (with channel gain $|h_2|$). For both channel realizations, the preamble sequence used for tuning the VGA gain is the same as discussed in Section II. However, for channel realization 2, even before transmitting this preamble sequence, a known ρ-training sequence is transmitted for estimating ρ. During the transmission of the ρ-training sequence the analog gains g_{AFE} and g_{VGA} are set to the values programmed during the transmission of coded information symbols in channel realization 1, and therefore

$$g_{VGA}\,g_{AFE} = \frac{1}{|h_1|X}. \tag{26}$$

The ρ-training sequence is a sequence of l distinct *positive* valued symbols with each symbol being repeated multiple times to average out the effect of receiver noise [9]. Subsequently, we shall denote an arbitrary ρ-training sequence by $\mathcal{T}$. Let the k-th training symbol be given by c_k, $k = 1, 2, \cdots, l$. The l corresponding inputs to the sample and hold circuit are given by

$$s_k = g_{AFE}\,g_{VGA}|h_2|c_k\,,\; k = 1, 2, \cdots l \tag{27}$$

Using (26) in (27), the sample and hold, and quantizer outputs during the transmission of the ρ-training sequence in channel realization 2 are given by

$$\begin{aligned} s_k &= \frac{|h_2|c_k}{|h_1|X} = \rho\frac{c_k}{X}\,, \\ r_k &= Q_b\Big(s_k\Big) = Q_b\Big(\rho\frac{c_k}{X}\Big)\,,\; k = 1, 2, \ldots l. \end{aligned} \tag{28}$$

In channel realization 2, after all the ρ-training symbols are transmitted, the receiver estimates ρ based on the l observations $\{r_k, k = 1, 2 \cdots, l\}$.

C. Estimation of ρ

The l discrete outputs of the ADC ($\{r_1, r_2, \cdots r_l\}$) can be used to estimate ρ as follows. Given the l discrete outputs, the maximum likelihood estimate (MLE) of ρ is given by

$$\rho_{ML} = \arg\max_{\rho>0} P(r_1, r_2, \cdots r_l | \rho, \{c_k\}, l) \tag{29}$$

[9]At the receiver, depending upon the repetition factor of the ρ-training symbols, the cut-off frequency of the LPF is appropriately reduced, which helps in noise reduction.

where $P(r_1, r_2, \cdots r_l|\rho, \{c_k\}, l)$ is the probability that the l outputs take the values $\{r_1, r_2, \cdots r_l\}$ for a given channel gain ratio ρ, and the ρ-training sequence $\{c_k\}$. For the k-th training symbol c_k, since $r_k = Q_b(\rho c_k/X)$, from (5) it must be true that

$$\frac{r_k - \frac{1}{2^b-1}}{c_k/X} \leq \rho < \frac{r_k + \frac{1}{2^b-1}}{c_k/X} \ , \text{ if } r_k < 1 \qquad (30)$$
$$\frac{1 - \frac{1}{2^b-1}}{c_k/X} \leq \rho < \infty \ , \text{ if } r_k = 1$$

The inequality in (30) defines an interval of the positive real line, which we shall denote by $\mathcal{L}_k \ , \ k = 1, 2, \ldots, l$. Therefore the l outputs would be $\{r_1, r_2, \cdots r_l\}$ if and only if $\rho \in \mathcal{L}$, where $\mathcal{L} \stackrel{\Delta}{=} \cap_{k=1}^{l} \mathcal{L}_k$. Further, we would call $\mathcal{L}$ as the "ML interval" corresponding to the training sequence $\{c_k\}$ and the l outputs $\{r_1, r_2, \cdots r_l\}$. Also, given the l outputs, all the values of ρ in the interval $\mathcal{L}$ are equally probable. Let $\mathcal{L}_{sup} \stackrel{\Delta}{=} \sup \mathcal{L}$, and $\mathcal{L}_{inf} \stackrel{\Delta}{=} \inf \mathcal{L}$, denote the supremum and infimum of the interval $\mathcal{L}$. One possible ML estimate of ρ, that we propose, is then given by

$$\widehat{\rho} = \begin{cases} \frac{\mathcal{L}_{sup} + \mathcal{L}_{inf}}{2} & \mathcal{L}_{sup} < \infty \\ \mathcal{L}_{inf} & \text{otherwise.} \end{cases} \qquad (31)$$

For a given ρ-training sequence $\mathcal{T} = \{c_1, c_2, \cdots, c_l\}$, and a set of corresponding outputs $\mathcal{R} = \{r_1, r_2, \cdots r_l\}$, let $\mathcal{L}(\mathcal{T}, \mathcal{R}) \subset \mathbb{R}^+$ denote the ML interval.

As an example, let us consider a b=2-bit quantizer, and a training sequence $\{c_k\} = \{X/4, X/2, X, 2X, 4X\}$. Let the $l = 5$ corresponding output symbols of the quantizer be $\{r_1 = 1/3, r_2 = 1/3, r_3 = 1, r_4 = 1, r_5 = 1\}$. The intervals $\mathcal{L}_k$ corresponding to these 5 outputs are

$$\mathcal{L}_1 : 0 \leq \frac{\rho}{4} < \frac{2}{3} \, , \, \mathcal{L}_2 : 0 \leq \frac{\rho}{2} < \frac{2}{3} \, , \, \mathcal{L}_3 : \frac{2}{3} \leq \rho < \infty \, ,$$
$$\mathcal{L}_4 : \frac{2}{3} \leq 2\rho < \infty \, , \, \mathcal{L}_5 : \frac{2}{3} \leq 4\rho < \infty. \qquad (32)$$

The ML interval $\mathcal{L}(\{X/4, X/2, X, 2X, 4X\}, \{1/3, 1/3, 1, 1, 1\})$ = $[2/3\ , \ 4/3)$ and hence $\widehat{\rho} = 1$.

For a b-bit quantizer and some fixed ρ-training sequence $\mathcal{T} = \{c_1, c_2, \cdots, c_l\}$ of l training symbols, it is clear that for each value of $\rho \in [0, \infty)$, there is a corresponding output sequence $\mathcal{R}(\rho, \mathcal{T}) = \{r_k = Q_b(\rho \frac{c_k}{X}) \, , \, k = 1, 2, \cdots, l\}$. We shall refer to each such possible output sequence as a *feasible* output sequence for the given training sequence. Note that even though the range of values of ρ is infinite, the number of distinct feasible output sequences is finite due to the finite length of the ρ-training sequence and the finite number (2^b) of quantizer levels for a b-bit uniform quantizer. Further, for each feasible output sequence $\mathcal{R}'$, there exists an ML interval $\mathcal{L}(\mathcal{T}, \mathcal{R}')$ [10].

Given a fixed ρ-training sequence $\mathcal{T}$, since $\hat{\rho}$ is always between the supremum and the infimum of the interval $\mathcal{L}(\mathcal{T}, \mathcal{R}(\rho, \mathcal{T}))$, it follows that both ρ and the proposed ML estimate $\hat{\rho}$ lie in the ML interval $\mathcal{L}(\mathcal{T}, \mathcal{R}(\rho, \mathcal{T}))$. In addition to this, if the ML interval corresponding to each feasible output sequence is a subset of some interval induced by the set $\mathcal{Q}_M^+$, then for any $l \in \mathcal{Q}_M^+$ it follows that if l is greater than ρ then it is also greater than $\hat{\rho}$, and similarly if l is smaller than ρ then it is also smaller than $\hat{\rho}$. However this is precisely the sufficiency condition in Theorem 5.2. We can therefore conclude that with the proposed ρ estimation technique (see (31)), a training sequence $\mathcal{T}$ results in an optimal estimate of ρ, if $\mathcal{T}$ satisfies the following conditions.

$$\forall \ 0 \leq \rho < \infty \quad : \quad \mathcal{L}(\mathcal{T}, \mathcal{R}(\rho, \mathcal{T})) \subseteq I \, , \text{for some } I \in \mathcal{I}(\mathcal{Q}_M^+). \qquad (33)$$

We next design optimal and near-optimal ρ-training sequences based on the criterion in (33).

D. Design of ρ-training sequence for estimating ρ

We first show that with $M = 2$ and a $b = 2$-bit uniform quantizer it is possible to design a ρ-training sequence which satisfies (33) and is therefore optimal.

Theorem 5.3: Consider a rotated constellation matched to the quantizer. Let $M = 2$, $b = 2$ and $\mathcal{Q}_M^+ = \{0, q_1, q_2, \cdots q_{L_M}\}, 0 < q_1 < q_2 < \cdots < q_{L_M}$. The following ρ-training sequence $\{c_k\}$ of length L_M with the proposed ML estimator (Section V-C) results in an *optimal* estimate of ρ.

$$c_k = \frac{2}{3} \frac{X}{q_{L_{M} - k+1}} \quad , \quad k = 1, 2, \cdots, L_M. \qquad (34)$$

Proof: See Appendix E in [11]. ∎

For a general $M > 2$, it is challenging to design an optimal ρ-training sequence which satisfies the sufficiency condition in Theorem 5.2. Further, we conjecture that, just as with $M = 2$, for any $M > 2$ also, the length of optimal ρ-training sequences based on Theorem 5.2 would be proportional to $L_M = |\mathcal{Q}_M^+|$. However, the cardinality of $\mathcal{Q}_M^+$ is a rapidly increasing function of M (e.g., $|\mathcal{Q}_2^+| = 29$, and $|\mathcal{Q}_4^+| = 3939$), which then implies that with increasing M a significant amount of communication bandwidth would be used up in the transmission of the ρ-training sequence, resulting in reduced overall throughput. Therefore, it is of *practical* interest to design ρ-training sequences which are short in length and which can still achieve an error performance comparable to that achieved with optimal ρ-training sequences.

Towards designing such *practical* sequences, we observe that the average error performance would be sensitive to the amount of overlap between the ML intervals induced[11] by the ρ-training sequence and the intervals induced by $\mathcal{Q}_M^+$. With short ρ-training sequences, the ML intervals induced by the ρ-training sequence would not coincide exactly with the intervals induced by $\mathcal{Q}_M^+$. Nevertheless, it may be possible to design short ρ-training sequences for which some of the ML intervals belong to $\mathcal{I}(\mathcal{Q}_M^+)$. With short ρ-training sequences, any interval in $\mathcal{I}(\mathcal{Q}_M^+)$, which is exactly the same as some ML interval induced by the ρ-training sequence, shall be referred to as "covered" by that ρ-training sequence.

From Fig. 6, we try to gain more insights into the problem of designing shorter length ρ-training sequences for $M = 2$. We observe that the density of the intervals induced by $\mathcal{Q}_M^+$ (depicted with cross 'X' marks on the horizontal axis) is

[10]For more interesting properties of ML intervals please refer to Section V in [11].

[11]These are basically the ML intervals corresponding to all feasible output sequences for the given ρ-training sequence.

much higher near the origin than farther away. Furthermore, from the p.d.f. of $\rho = |h_2|/|h_1|$, we observe that most of the probability mass is distributed near the origin[12]. Based on these observations and the sufficiency conditions in Theorem 5.2, it can be argued that, to have an error performance comparable to that of an optimal ρ-training sequence, any short ρ-training sequence should aim to "cover" the intervals of $\mathcal{I}(\mathcal{Q}_M^+)$ which are closer to the origin. This reasoning is supported by two facts. Firstly, with i.i.d. channels gains $|h_1|$ and $|h_2|$, the probability of ρ taking large values is small, and hence the estimation error $(\hat{\rho} - \rho)$ for large values of ρ is expected to have less contribution to the average error probability than smaller values of ρ. Secondly, when $\rho \gg 1$, any error in the estimation of ρ is likely to have a lesser impact on the error performance compared to when $\rho < 1$. To see this, we note that for $\rho \gg 1$, the ML estimate for any ρ-training sequence would be the infimum value of the ML interval corresponding to the all ones output sequence which would also be large i.e., $\hat{\rho} \gg 1$. Therefore for any two transmit vectors $\mathbf{x} = (x_1, x_2)^T = \mathbf{Gu}$ and $\mathbf{y} = (y_1, y_2)^T = \mathbf{Gv}$, $D_E(\rho, \mathbf{r}^I, \mathbf{u}^I, \mathbf{v}^I) D_E(\hat{\rho}, \mathbf{r}^I, \mathbf{u}^I, \mathbf{v}^I) \approx \rho^2 \hat{\rho}^2 d_2^2 > 0$, where $d_2 = (r_2^I - x_2^I/X)^2 - (r_2^I - y_2^I/X)^2$. Using Theorem 5.1, this then implies that, with high probability, the output of the minimum distance decoder with estimated ρ is the same as its output with perfect knowledge of ρ.

Therefore, any short ρ-training sequence should aim to "cover" the intervals of $\mathcal{I}(\mathcal{Q}_M^+)$ which are closer to the origin. With $M = 2$, in Fig. 6 a short ρ-training sequence of length $l < L_M$ is designed in a way to "cover" only the intervals of $\mathcal{I}(\mathcal{Q}_M^+)$ which are closer to origin. This is done by *non-uniformly* sampling[13] out l distinct elements of the set $\mathcal{Q}_M^+$, such that the intervals induced by these l elements coincide with *most* of the intervals of $\mathcal{I}(\mathcal{Q}_M^+)$ which are near to origin. Let us denote this set of l elements as $\mathcal{Q}_{(M,l)}^+$. The corresponding short ρ-training sequence which has ML intervals coinciding exactly with the intervals induced by $\mathcal{Q}_{(M,l)}^+$ is then given by (34), where the set $\mathcal{Q}_M^+$ is replaced by the set $\mathcal{Q}_{(M,l)}^+$ and L_M is replaced by l. In Fig. 6, for $l = 9$, the elements of *one such* $\mathcal{Q}_{(M,l)}^+$ are depicted through 'triangles'. Short ρ-training sequences which achieve an error performance close to that achieved with optimal ρ-training sequences, would be subsequently referred to as 'good' ρ-training sequences.

Even though 'non-uniform' sampling of $\mathcal{Q}_M^+$ is one possible method for designing 'good' ρ-training sequences, with increasing M, the number of ways in which 'non-uniform' sampling can be done, would also increase rapidly, thereby increasing the complexity of finding 'good' ρ-training sequences. Therefore for large M, a simpler strategy is required to search for 'good' ρ-training sequences. We next present a very simple and parameterizable short ρ-training sequence design, which has been observed to result in 'good' ρ-training sequences. The k-th symbol of the proposed ρ-training sequence is given by

$$c_k = X d^{(k - \frac{l+1}{2})} \quad , \quad k = 1, 2, \cdots, l \tag{35}$$

where l is the length of the training sequence, and $d > 1$ is the ratio between the consecutive ρ-training symbols. Let us denote this ρ-training sequence by $\mathcal{T}_d$. This design is based on the observation that for many 'non-uniformly' sampled subsets $\mathcal{Q}_{(M,l)}^+ \subset \mathcal{Q}_M^+$, it is possible to find a value of d, such that a ρ-training sequence designed using (35), would have ML intervals "almost" same as the ML intervals of the ρ-training sequence designed using the 'non-uniformly' sampled subset. That is, for many non-uniformly sampled subsets $\mathcal{Q}_{(M,l)}^+$, for any $\rho > 0$, there exists some $d > 1$ and some ϵ close to zero, such that

$$|\mathcal{L}(\mathcal{T}_d, \mathcal{R}(\rho, \mathcal{T}_d)) \cap \mathcal{L}(\mathcal{T}', \mathcal{R}(\rho, \mathcal{T}'))^c| \quad \leq \quad \epsilon |\mathcal{L}(\mathcal{T}', \mathcal{R}(\rho, \mathcal{T}'))| \tag{36}$$

where $\mathcal{T}'$ refers to the ρ-training sequence designed using $\mathcal{Q}_{(M,l)}^+ \subset \mathcal{Q}_M^+$.[14,15]

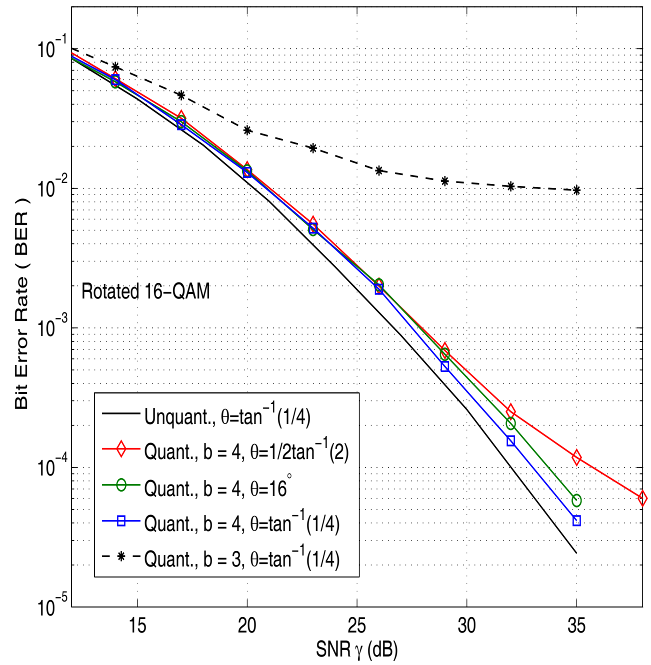

Fig. 7. BER vs. SNR for a quantized receiver. $b = 4$, 16-QAM and perfect channel state information at receiver.

VI. Simulation Results

All error probabilities reported in this section have been averaged over the Rayleigh flat fading statistics of the channel.

[12]In fact, for any other fading distribution also, it can be shown that $P(\rho < 1) = P(\rho > 1) = 1/2$.

[13]We use the word "non-uniform" since the sampling is biased towards choosing more elements which are closer to the origin.

[14]For any real interval I, $|I| \triangleq (\sup I - \inf I)$ refers to the length of the interval and I^c refers to the complementary set $\mathbb{R} - I$ (i.e., all real numbers which do not belong to I).

[15] For $M = 2$, this fact is illustrated through Fig. 6, where for the given 'non-uniformly' sampled subset of $\mathcal{Q}_M^+$, i.e., $\mathcal{Q}_{(M,l)}^+$ (shown with triangles), a ρ-training sequence designed using (35) has ML intervals (shown with circles) almost coinciding with the intervals induced by $\mathcal{Q}_{(M,l)}^+$. For a given M and l, the optimal d can be found at reasonable complexity, by minimizing the average error probability (as a function of d) using Monte-Carlo techniques.

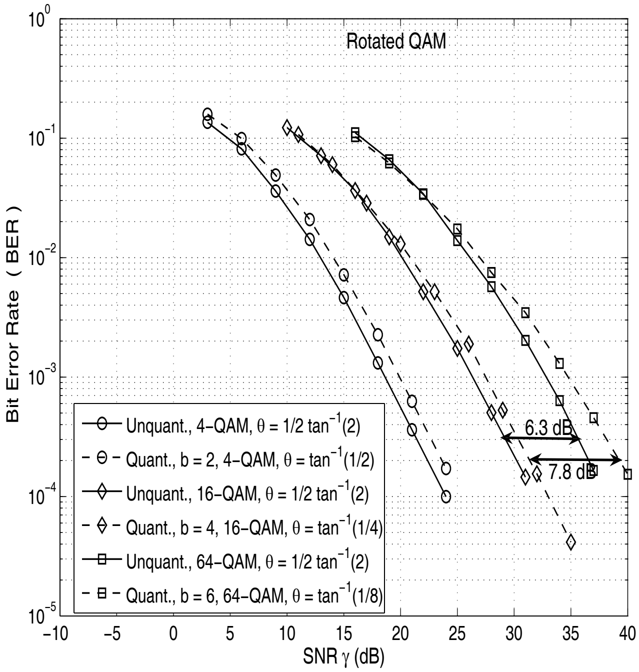

Fig. 8. BER comparison between quantized and unquantized receivers. 4-,16-,64-QAM. Perfect channel state information at receiver.

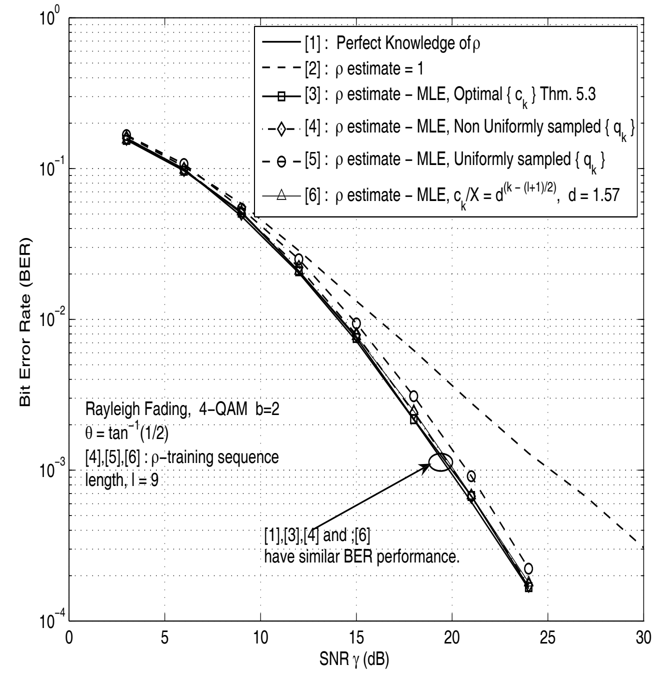

Fig. 9. BER performance with a quantized receiver ($b = 2$) and imperfect receiver knowledge of ρ. Rotated 4-QAM ($M = 2$).

In Fig. 7, we plot the average bit error rate/probability (BER), for rotated 16-QAM constellation ($M = 4$) and a $b = 4$-bit quantized receiver with perfect channel state information. The following four important observations can be made in Fig. 7: *i*) with $\theta = 1/2\tan^{-1}(2)$ (which is known to achieve full modulation diversity in an unquantized receiver, but does not satisfy Criterion II), the BER performance with a quantized receiver fails to achieve full diversity (note the difference in slope at high SNR), which validates Criterion II, *ii*) with $\theta = \tan^{-1}(1/4)$, which results in equidistant projections, the quantized receiver achieves full modulation diversity with $b = 4$. Further, the quantized receiver performs only 1 dB away from an ideal *unquantized* receiver at a BER of 10^{-4}, *iii*) with a quantized receiver a rotation angle of $\theta = 16^\circ$ also appears to achieve full modulation diversity, but performs poor when compared to a matched rotated constellation with $\theta = \tan^{-1}(1/4)$. This supports Criterion III, and *iv*) it is also observed that with 16-QAM rotated constellation ($\theta = \tan^{-1}(1/4)$), the error performance floors with $b = 3 < 4$ quantization bits, which validates Criterion I.

It was discussed in Section IV, that with increasing QAM size, a quantized receiver would be increasingly less power efficient when compared to an unquantized receiver. This fact is illustrated in Fig. 8, where the BER performance of both unquantized receiver with $\theta = 1/2\tan^{-1}(2)$ and quantized receiver with $\theta = \tan^{-1}(1/M)$ are plotted for $M^2 = 4-, 16-$ and 64-QAM, $b = \lceil 2\log_2(M)\rceil$ and perfect channel knowledge at the receiver. To achieve a fixed BER of 2×10^{-4}, with increasing QAM size, the increase in signal power required by a quantized receiver is more than that for an unquantized receiver.

In Fig. 9, the BER performance with minimum distance decoding and imperfect receiver knowledge of ρ, is plotted as a function of γ for $M = 2$, $\theta = \tan^{-1}(1/2)$ and $b = 2$. Firstly we note that, for a matched rotated constellation, the error probability performance with $\hat{\rho}$ being any arbitrary positive valued estimate of ρ, *does not have error floors*. This is because, in the absence of noise (i.e., $\gamma = \infty$), when a certain information symbol vector $\mathbf{v}$ is transmitted and $\mathbf{r}$ is the quantized output, the detection metric of some information symbol vector $\mathbf{u}$ (i.e., $m(\hat{\rho}, \mathbf{r}^I, \mathbf{u}^I)$ and $m(\hat{\rho}, \mathbf{r}^Q, \mathbf{u}^Q)$) is equal to zero only for $\mathbf{u} = \mathbf{v}$, and is positive for all other possible information symbol vectors. Therefore, the detector output $\widehat{\mathbf{u}}$ is the same as the transmitted vector $\mathbf{v}$, resulting in zero probability of error. This argument is supported by the fact that in Fig. 9, a fixed estimate of $\hat{\rho} = 1$, has no floors in its BER performance.

In Fig. 9 we also observe that the BER performance of the *optimal* ρ-training sequence designed using Theorem 5.3 is same as the BER achieved with perfect knowledge of ρ. Further, a short ($l = 9$) 'non-uniform' sampling based ρ-training sequence designed with $\mathcal{Q}^+_{(M,l)} = \{1/9, 1/5, 1/4, 4/9, 5/8, 1, 5/3, 8/3, 4\} \subset \mathcal{Q}^+_M$ achieves a BER close to that achieved by the optimal ρ-training sequence designed using Theorem 5.3 (compare curves 3 and 4). Also, the BER performance of a 'uniform' sampling based ρ-training sequence design with $\mathcal{Q}^+_{(M,l)} = \{1/9, 8/9, 8/5, 9/4, 3, 4, 5, 8, 9\}$ (the induced intervals are almost uniformly distributed) is inferior to the BER performance achieved by the 'non-uniform' sampling based design (compare curves 4 and 5). Finally, it is observed that, the BER achieved with the ρ-training sequence designed using (35) (with $d = 1.57, l = 9$) is similar to the BER achieved with perfect knowledge of ρ (compare curves 1 and 6).

For higher order rotated M^2-QAM, we proposed 'good' ρ-training sequences which are short and have near-optimal performance. We support this fact through Fig. 10, where we plot the BER performance for a rotated 16-QAM constellation

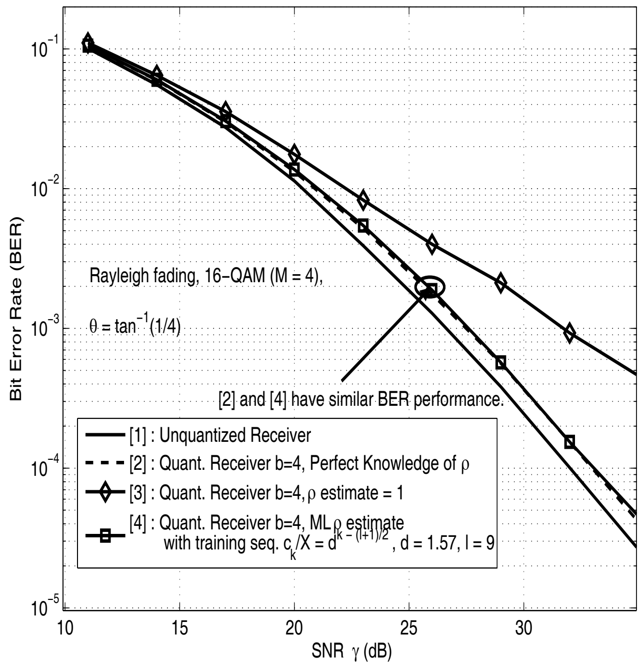

Fig. 10. BER performance with a quantized receiver ($b = 4$) and imperfect receiver knowledge of ρ. Rotated 16-QAM ($M = 4$).

($\theta = \tan^{-1}(1/4)$), with a quantized receiver ($b = 4$) and minimum distance decoding at the receiver with imperfect knowledge of ρ. An estimate of ρ is computed based on the proposed ML estimation scheme discussed in Section V-C. The ρ-training sequence used for estimation is the same ρ-training sequence used in simulation curve 6 of Fig. 9. From curve 4 in Fig. 10, it is observed that with a short ρ-training sequence of only 9 symbols, it is possible to achieve a BER performance comparable to the BER performance achieved with perfect knowledge of ρ (curve 2). This is interesting since the same ρ-training sequence was also observed to be near-optimal with $M = 2, b = 2$ in Fig. 9. It therefore appears that the length of near-optimal/'good' ρ-training sequences does not increase significantly with increasing QAM size. One possible reason for this could be that with increasing QAM size, the quantizer resolution b also increases, which makes the estimate of ρ more reliable.

VII. Conclusions

In this paper, we addressed the problem of achieving full modulation diversity in fading channels with quantized receiver. For 2-dimensional modulation coding, through analysis we showed that in quantized receivers with perfect channel knowledge, algebraic rotations with *equidistant* projections can achieve full modulation diversity even with a low complexity minimum distance decoder. We then relaxed the perfect channel knowledge assumption, and proposed novel channel training/estimation, which were shown to achieve an error probability performance similar to that achieved with perfect channel knowledge.

References

[1] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[2] A. Mezghani, M. S. Khoufi, and J. A. Nossek, "Maximum likelihood detection for quantized MIMO systems," in *Proc. 2008 International ITG Workshop Smart Antennas*, pp. 278–284.

[3] M. Gareth and A. Sabharwal, "On the impact of finite receiver resolution in fading channels," *2006 Allerton Conf. Commun., Control Comput.*,.

[4] M. T. Ivrlac and J. A. Nossek, "Capacity and coding for quantized MIMO systems," in *Proc. 2006 IEEE International Wireless Commun. Mobile Comput. Conf.*, pp. 1387–1392.

[5] J. Singh, O. Dabeer, and U. Madhow, "Capacity of the discrete-time AWGN channel under output quantization," in *Proc. 2008 IEEE International Symp. Inf. Theory*, pp. 1218–1222.

[6] J. Boutros and E. Viterbo, "Signal space diversity: a power and bandwidth efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1453-1467, July 1998.

[7] J. Boutros, E. Viterbo, C. Rastello, and J. C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian Channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 502–518, Mar. 1996.

[8] Q. Gu, *RF System Design of Transceivers for Wireless Communications*, 1st edition. Springer, 2005.

[9] S. I. Liu and C. C. Chang, "CMOS analog divider and four-quadrant multiplier using pool circuits," *IEEE J. Solid-State Circuits*, vol. 30, no. 9, pp. 1025-1029, Sep. 1995.

[10] E. Bayer-Fluckiger, F. Oggier, and E. Viterbo, "New algebraic constructions of rotated $\mathbb{Z}^n$-lattice constellations for the Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 702–714, Apr. 2004.

[11] S. K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "Modulation diversity in fading channels with quantized receiver." Available: www.arxiv.org,arXiv:1104.1155v1[cs.IT], Apr. 2011.

[12] Z. Wu, T. Fu, X. Wang, L. Zhang, Y. Gao, C. Ma, and W. Wang, "A novel coding-rotated-modulation OFDM scheme," in *Proc. 2009 IEEE International Conf. Commun. Technol. Appl.*, pp. 517–520.

Saif Khan Mohammed (S'08-M'11) received the B.Tech. degree in computer science and engineering from the Indian Institute of Technology (I.I.T.), New Delhi, India, in 1998. Subsequently, he received his Ph.D. degree in Electrical and Communication Engineering from the Indian Institute of Science (I.I.Sc.), Bangalore, India, December 2010. Since October 2010, he is a post doctoral researcher at the Communication Systems Division of the Electrical Engineering Department (ISY) at the University of Linköping, Sweden.

From 1998 to 2000, he was employed with Philips, Inc., Bangalore, as an ASIC Design Engineer. From 2000 to 2003, he worked with Ishoni Networks, Inc., Santa Clara, CA, as a Senior Chip Architecture Engineer. From 2003 to 2007, he was employed with Texas Instruments, Bangalore, as a Systems and Algorithms Designer in the wireless systems group.

Dr. Saif K. Mohammed was awarded the Young Indian Researcher Fellowship for the year 2008-09 by the Italian Ministry of University and Research (MIUR). His main research interests include wireless communication using large antenna arrays, coding and signal processing for wireless communication systems, and statistical signal processing.

Emanuele Viterbo (M'95-SM'04-F'11) was born in Torino, Italy, in 1966. He received his degree (Laurea) in Electrical Engineering in 1989 and his Ph.D. in 1995 in Electrical Engineering, both from the Politecnico di Torino, Torino, Italy. From 1990 to 1992 he was with the European Patent Office, The Hague, The Netherlands, as a patent examiner in the field of dynamic recording and error-control coding. Between 1995 and 1997 he held a post-doctoral position in the Dipartimento di Elettronica of the Politecnico di Torino. In 1997-98 he was a post-doctoral research fellow in the Information Sciences Research Center of AT&T Research, Florham Park, NJ, USA. He became first Assistant Professor (1998) then Associate Professor (2005) in Dipartimento di Elettronica at Politecnico di Torino. In 2006 he became Full Professor in DEIS at University of Calabria, Italy. Since September 2010 he is Full Professor in the ECSE Department at Monash University, Melbourne, Australia.

Prof. Emanuele Viterbo is an ISI Highly Cited Researcher since 2009. He is Associate Editor of *IEEE Transactions on Information Theory, European Transactions on Telecommunications and Journal of Communications and Networks,* and Guest Editor for *IEEE Journal of Selected Topics in Signal Processing: Special Issue Managing Complexity in Multiuser MIMO Systems.*

In 1993 he was visiting researcher in the Communications Department of DLR, Oberpfaffenhofen, Germany. In 1994 and 1995 he was visiting the Ecole Nationale Supérieure des Télécommunications (E.N.S.T.), Paris. In 2003 he was visiting researcher at the Maths Department of EPFL, Lausanne, Switzerland. In 2004 he was visiting researcher at the Telecommunications Department of UNICAMP, Campinas, Brazil. In 2005, 2006 and 2009 he was visiting researcher at the ITR of UniSA, Adelaide, Australia. In 2007 he was visiting fellow at the Nokia Research Center, Helsinki, Finland.

Dr Emanuele Viterbo was awarded a NATO Advanced Fellowship in 1997 from the Italian National Research Council. His main research interests are in lattice codes for the Gaussian and fading channels, algebraic coding theory, algebraic space-time coding, digital terrestrial television broadcasting, digital magnetic recording, and irregular sampling.

Yi Hong (S'00-M'05-SM'10) is currently a lecturer at the Department of Electrical and Computer Systems Eng., at Monash University, Melbourne, Australia. She received her Ph.D. degree in Electrical Engineering and Telecommunications from the University of New South Wales (UNSW), Sydney, Australia, in Oct. 2004. She then worked at the Institute of Telecom. Research, University of South Australia, Australia; at the Institute of Advanced Telecom., Swansea University, UK; and at University of Calabria, Italy.

During her PhD, she received an International Postgraduate Research Scholarship (IPRS) from the Commonwealth of Australia; a supplementary Engineering Award from the School of Electrical Engineering and Telecommunications, UNSW; and a Wireless Data Communication System Scholarship from UNSW. She received the NICTA-ACoRN Earlier Career Researcher award for a paper presented at the Australian Communication Theory Workshop (AUSCTW), Adelaide, Australia, 2007.

Dr. Hong is an Associate Editor for *European Transactions on Telecommunications* and an IEEE Senior member. She was Technical Program Committee Chair of AUSCTW'11, Melbourne, Australia. She was the Publicity Chair at the IEEE Information Theory Workshop 2009, Sicily, Italy. She is a Technical Program Committee member for various IEEE conferences such as IEEE ICC 2011, VTC 2011, PIMRC and WCNC 2008. Her research interest include information and communication theory with applications to telecommunication engineering.

A. Chockalingam (S'92-M'93-SM'98) was born in Rajapalayam, Tamil Nadu, India. He received the B.E. (Honors) degree in Electronics and Communication Engineering from the P. S. G. College of Technology, Coimbatore, India, in 1984, the M.Tech. degree (with specialization in satellite communications) from the Indian Institute of Technology, Kharagpur, India, in 1985, and the Ph.D. degree in Electrical Communication Engineering (ECE) from the Indian Institute of Science (IISc), Bangalore, India, in 1993.

During 1986 to 1993, he worked with the Transmission R&D division of the Indian Telephone Industries Limited, Bangalore. From December 1993 to May 1996, he was a Postdoctoral Fellow and an Assistant Project Scientist at the Department of Electrical and Computer Engineering, University of California, San Diego. From May 1996 to December 1998, he served Qualcomm, Inc., San Diego, CA, as a Staff Engineer/Manager in the systems engineering group. In December 1998, he joined the faculty of the Department of ECE, IISc, Bangalore, India, where he is a Professor, working in the area of wireless communications and networking.

Dr. Chockalingam is a recipient of the Swarnajayanti Fellowship from the Department of Science and Technology, Government of India. He served as an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from May 2003 to April 2007. He currently serves as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He served as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on Multiuser Detection for Advanced Communication Systems and Networks). Currently, he serves as a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (). He is a Fellow of the Institution of Electronics and Telecommunication Engineers and the Indian National Academy of Engineering.