Diversity Combining via Universal Orthogonal Space-Time Transformations

Elad Domanovitz and Uri Erez
Dept. EE-Systems, Tel Aviv University, Israel

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD.**

Receiver diversity combining methods play a key role in combating the detrimental effects of fading in wireless communication and other applications. Commonly used linear diversity combining methods include maximal-ratio combining, equal-gain combining and antenna selection. A novel linear combining method is proposed where a universal, i.e., channel independent, orthogonal dimension-reducing space-time transformation is applied prior to quantization of the signals. The scheme may be considered as the counterpart of Alamouti modulation, and more generally of orthogonal space-time block codes.

I. Introduction

In wireless communication, diversity methods play a central role in combating the detrimental effects of severe channel variation (fading). Of the many techniques that have been developed over the years with this goal, an important class involves the use of multiple receive antennas. With sufficient separation between the antennas, each antenna may be viewed as a branch receiving the transmitted signal multiplied by an approximately independent fading coefficient. Diversity is achieved as the probability that the signal is severely affected by fading on all branches simultaneously is greatly reduced. The number of such (roughly) independent branches is commonly referred to as the diversity order. A classical survey of receive diversity techniques is [1]. A more recent account that also considers multiple-input multiple-output channels is [2].

We introduce a new linear diversity-combining scheme utilizing orthogonal space-time block codes. The key difference between the proposed scheme and traditional linear combining schemes is that it is *universal*. That is, the combining weights (in the proposed scheme, the space-time transformation) do not depend on the the channel realization.

To understand the potential benefits of the contribution, consider a receiver as depicted in Figure 1. A key feature of modern device architectures is the decomposition of the unit into separate functional blocks (that can be located at different physical locations, i.e., distributed processing). These blocks are connected by interfaces and a major design goal is to reduce the bandwidth between different blocks.

The proposed scheme can assist in the interface from the analog domain to the digital one, simplifying analog-to-digital conversion (ADC) and thus also reducing power consumption. This application is described in Section III. The scheme can also assist in reducing the bit rate of the digital interface between different digital blocks. For example, as described in Section IV, in a centralized (cloud) radio access network setting, this reduction will be in the fronthaul links between the relays (remote radio head units) and the central (cloud) processing unit.

II. Description of the Scheme for Two Receive Antennas

Consider a 2×1 SIMO channel, with channel coefficient h_1 and h_2, as depicted in Figure 1. The signal received at antenna

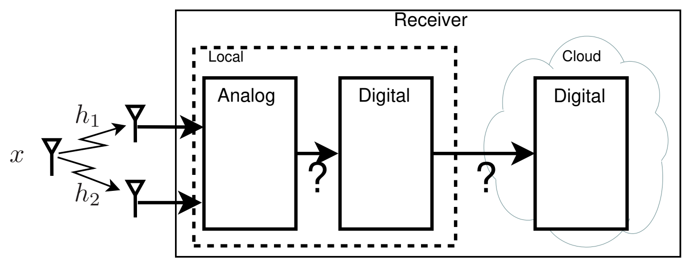

Fig. 1. Basic scenario: receiver architecture for a 2×1 SIMO channel.

$i = 1, 2$, at discrete time t, is

$$s_i(t) = h_i x(t) + n_i(t). \tag{1}$$

We assume that the noise $n_i(t)$ is i.i.d. over space and time with samples that are circularly-symmetric complex Gaussian random variables with unit variance

The scheme works on batches of two time instances and for our purposes, it will suffice to describe it for time instances $t = 1, 2$. Let us stack the four complex samples received over $T = 2$ time instances, two over each antenna, into an 8×1 real vector:

$$\mathbf{s} = [s_{1R}(1)s_{1I}(1)s_{2R}(1)s_{2I}(1)s_{1R}(2)s_{1I}(2)s_{2R}(2)s_{2I}(2)]^T, \tag{2}$$

where x_R and x_I denote the real and imaginary parts of a complex number x. We similarly define the stacked noise vector $\mathbf{n}$. Likewise, we define

$$\mathbf{x} = [x_R(1)\, x_I(1)\, x_R(2)\, x_I(2)]^T. \tag{3}$$

Next, we form a 4-dimensional real vector $\mathbf{y}$ by applying to the vector $\mathbf{s}$ the transformation $\mathbf{y} = \mathbf{P}\mathbf{s}$ where

$$\mathbf{P} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{4}$$

Note that unlike conventional linear diversity-combining schemes, here the combining matrix $\mathbf{P}$ is *universal*, i.e., it does not depend on the channel coefficients.

Remark 1: We note that the transpose of $\mathbf{P}$ is precisely the description of the linear operation performed by Alamouti modulation [3] when expressed over the reals.

It is readily shown that the following holds

$$\begin{aligned}\mathbf{y} &= \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{U}(h_1,h_2)\mathbf{x} + \mathbf{P}\mathbf{n} \\ &= \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{U}(h_1,h_2)\mathbf{x} + \mathbf{n}', \qquad (5)\end{aligned}$$

where

$$\mathbf{U}(h_1,h_2) = \frac{1}{\|\mathbf{h}\|}\begin{bmatrix} h_{1R} & -h_{1I} & h_{2R} & -h_{2I} \\ h_{1I} & h_{1R} & -h_{2I} & -h_{2R} \\ h_{2R} & -h_{2I} & -h_{1R} & h_{1I} \\ h_{2I} & h_{2R} & h_{1I} & h_{R1} \end{bmatrix}. \qquad (6)$$

A key observation is that $\mathbf{U}(h_1,h_2)$ is an orthonormal matrix for any h_1, h_2:

$$\mathbf{U}^H(h_1,h_2)\mathbf{U}(h_1,h_2) = I, \qquad (7)$$

where $\mathbf{I}$ is the identity matrix. Further, since the rows of $\mathbf{P}$ are orthonormal, it follows that $\mathbf{n}'$ is i.i.d. and Gaussian with unit variance.

We may reconstruct (up to additive noise) the original samples by applying

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{U}^H(h_1,h_2)\cdot\mathbf{y} \\ &= \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{x} + \mathbf{n}'' \qquad (8)\end{aligned}$$

where $\mathbf{n}''$ is also i.i.d. Gaussian with unit variance.

Since the dimension (over the reals) of $\mathbf{y}$ is four rather than eight, as is the dimension of the received signal $\mathbf{s}$, we obtained a universal dimension-reducing linear-combining scheme.

III. Application to analog-to-digital conversion

In this section we demonstrate the applicability of the scheme to analog-to-digital conversion for power-limited receivers of narrowband signals.

The proposed method may be used to achieve maximal diversity order with a single radio-frequency (RF) chain and ADC, and without requiring selection and switching mechanisms that come at substantial analog hardware complexities; see discussion of hardware aspects in [2].

Consider again the scenario of a 2×1 SIMO system as depicted in Figure 1 and described in the previous section. We note that as the fading coefficients are constants (rather than impulse responses), the model assumed is that of frequency-flat fading.

The best performance may be attained by quantizing (at sufficient resolution) the output of each antenna and then using maximum-ratio combining (MRC). Applying MRC amounts to forming

$$\begin{aligned}y_{\mathrm{MRC}} &= \frac{1}{\|\mathbf{h}\|}\begin{bmatrix} h_1^* & h_2^* \end{bmatrix}\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \\ &= \frac{\|h_1\|^2 + \|h_2\|^2}{\|\mathbf{h}\|}x + \frac{h_1^* n_1 + h_2^* n_2}{\|\mathbf{h}\|} \\ &= \|\mathbf{h}\|x + n, \qquad (9)\end{aligned}$$

where n is white and Gaussian with unit variance. As the variation of $\|\mathbf{h}\|$ is much smaller than that of either h_1 or h_2, diversity is attained. When h_1 and h_2 are independent, we obtain a diversity order of 2. The precise performance of MRC under independent Rayleigh fading is well-known and may be found, e.g., in [1]. The major downside of such a system is that two RF chains and ADCs are needed.

A classic alternative to MRC that requires only one RF chain is the method of antenna selection or "selection combining". Here, rather than choosing the antenna arbitrarily, we choose the one with the higher SNR. Thus the effective channel becomes

$$y_{\mathrm{SC}} = \max(|h_1|, |h_2|)x + n, \qquad (10)$$

where again n is Gaussian noise of unit variance. While the performance does not reach that of MRC, it does attain a diversity order of 2. The precise performance under independent Rayleigh fading of selection combining is well-known and may be found, e.g., in [1]. A downside of the selection combining method is that it requires analog power measurement and switching mechanisms.

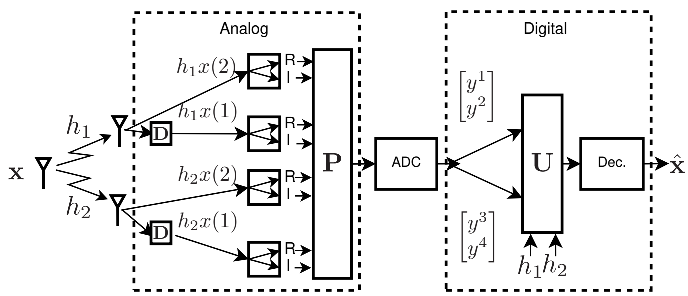

Fig. 2. Proposed receiver front end employing a universal orthogonal space-time diversity transformation.

The space-time diversity combining method described in the previous section may be applied to the problem of ADC as follows. Since the processing matrix $\mathbf{P}$ is fixed for all channels (i.e., is universal), it is relatively easy to apply in the analog domain (i.e., prior to quantization), requiring only delay, summation and negation elements.

As depicted in Figure 2, the received signals are first passed through the dimension-reducing transformation $\mathbf{P}$ to obtain the vector $\mathbf{y} = [y^1, y^2, y^3, y^4]^T$ as defined in (5) and (6). Then, a (component-wise) scalar uniform quantizer $Q(\cdot)$ is applied to $\mathbf{y}$ to obtain $\mathbf{y}_q = Q(\mathbf{y})$. We denote the quantization error vector by

$$\begin{aligned}\mathbf{e} &= \mathbf{y} - \mathbf{y}_q \\ &= \mathbf{y} - Q(\mathbf{y}). \qquad (11)\end{aligned}$$

The sequence of quantized samples is used to reconstruct an estimation of the source vector

$\hat{\mathbf{x}} = [\hat{x}_R(1), \hat{x}_I(1), \hat{x}_R(2), \hat{x}_I(2)]^T$ by applying the transformation:

$$\hat{\mathbf{x}} = \mathbf{U}(h_1, h_2)^H \mathbf{y}_q. \tag{12}$$

Using (5) and (11), we have

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{U}(h_1, h_2)^H (\mathbf{y} - \mathbf{e}) & (13) \\ &= \mathbf{U}(h_1, h_2)^H \left(\frac{\|\mathbf{h}\|}{\sqrt{2}} \mathbf{U}(h_1, h_2)\mathbf{x} + \mathbf{n}' - \mathbf{e} \right) & (14) \\ &= \frac{\|\mathbf{h}\|}{\sqrt{2}} \mathbf{x} + \mathbf{n}'' - \mathbf{e}', & (15) \end{aligned}$$

where $\mathbf{n}''$ has the same distribution as $\mathbf{n}$.

As for the quantization error $\mathbf{e}$ and its transformed variant $\mathbf{e}'$, we may invoke the standard assumption, that may be justified using subtractive dithered quantization, that it is independent of the signal (and hence of $\mathbf{x}$) and is white (i.e., its covariance matrix is the scaled identity).

We conclude that the input/output relationship of the proposed diversity combiner is identical to that of MRC, except for a power loss of a factor of two. In other words, we attain full diversity but no array gain, precisely as in the case of Alamouti space-time diversity transmission. In comparison with selection combining (without taking into account implementation losses), there a loss in the achieved SNR whereas an advantage is that no estimation of channel quality in the analog front end nor switching is required.

A comparison of the performance of the proposed method is shown in Figure 3 that plots the bit error rate of all three methods for uncoded QPSK transmission over a 2×1 Rayleigh fading channel.

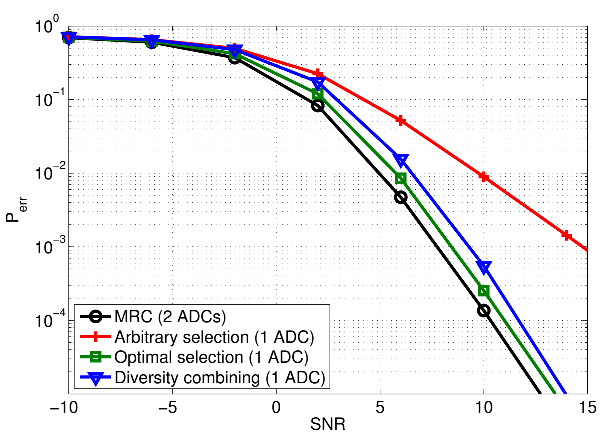

Fig. 3. Bit error rate of the proposed space-time diversity scheme and comparison with alternatives for uncoded QPSK over a 2×1 Rayleigh fading channel.

IV. Application to "Dumb" Relaying for Multi-User Detection at a Remote Destination

Another potential application of the proposed scheme is to be employed as part of a "dumb" relay. By a "dumb" relay we mean a relay (equipped with multiple antennas) that can only apply channel-independent linear processing to the antennas outputs followed by scalar quantization, the output of which is fed into a rate-constrained bit pipe.[1]

Unlike in the previous section, the scheme we present now operates purely in the digital domain. A further difference is that we no longer assume frequency-flat fading. Rather, we will assume that after analog-to-digital conversion, a DFT operation is applied, so that we are working in the frequency domain. In other words, the static channel we will consider is to be understood to apply to a single tone. The "time" index t will correspondingly refer to subsequent uses of the same tone, or in a practical setting could apply to adjacent tones as these typically have very similar channel coefficients.

We demonstrate the application to "dumb" relaying in the context of the system described in Figure 4. Here, two single-antenna users communicate with a central receiver via two relays, each equipped with two antennas, where the medium between the users and relays is a Rayleigh fading wireless channel, whereas the relays are connected to the central receiver via bit pipes. We further assume that the operation performed at the central decoder is also "dumb" in the sense that only linear processing is performed prior to applying channel decoding of the users' codes.

The signal received at relay $i = 1, 2$ and antenna $j = 1, 2$ is given by

$$s_j^i(t) = h_{j1}^i \cdot x_1(t) + h_{j2}^i \cdot x_2(t) + n_j^i(t), \tag{16}$$

and the corresponding channel matrix of relay i is

$$\mathbf{H}^i = \begin{bmatrix} h_{11}^i & h_{12}^i \\ h_{21}^i & h_{22}^i \end{bmatrix}. \tag{17}$$

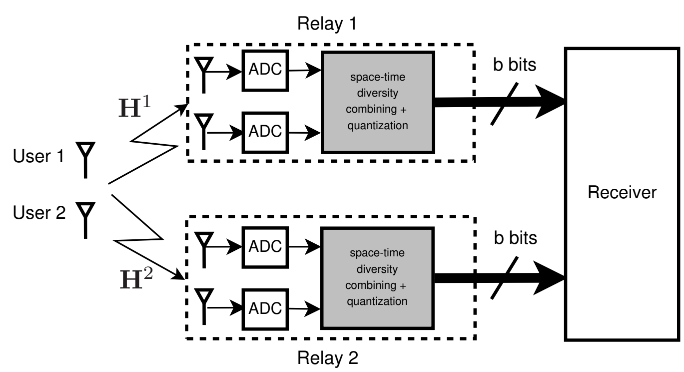

Fig. 4. Two-user virtual MIMO system formed by two two-antenna relays connected to a receiver via rate-constrained fronthaul links.

The question now arises as to how best to utilize the finite number of bits available per sample in quantizing the output of the two antennas. Note that not only do MRC and selection diversity depend on the use of channel state information (which is precluded by the definition of a "dumb" relay), due to the distributed nature of the problem, both MRC and

[1] This definition is similar to the definition of an instantaneous relay (see, e.g., [4] and [5]), with the additional requirement of linearity while allowing a small delay at the relay.

selection combining are also ineffective as the base station is interested in recovering both signals.

We now demonstrate that while keeping the bit rate fixed, each relay can nonetheless provide diversity gains to *both users* using the proposed diversity combining method, precisely since it makes no use of channel state information at the linear combining stage, rather only in the reconstruction stage.

Assuming both relays use the proposed space-time diversity combining scheme, the signal passed to the cloud from relay i is given by

$$\mathbf{y}^i = \mathbf{U}(h_{11}^i, h_{21}^i)\mathbf{x}_1 + \mathbf{U}(h_{12}^i, h_{22}^i)\mathbf{x}_2 + \mathbf{n}'^i,$$

where $\mathbf{x}_j$ represents the real representation of the signal transmitted by user j over the two time instances according to the notation in (3). Thus, at the cloud we obtain the effective channel

$$\begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \end{bmatrix} = \underbrace{\left[\begin{array}{c|c} \mathbf{U}(h_{11}^1, h_{12}^1) & \mathbf{U}(h_{21}^1, h_{22}^1) \\ \hline \mathbf{U}(h_{11}^2, h_{12}^2) & \mathbf{U}(h_{21}^2, h_{22}^2) \end{array}\right]}_{\mathcal{G}} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{n}'^1 \\ \mathbf{n}'^2 \end{bmatrix}.$$

Note that the effective matrix $\mathcal{G}$ has the desirable property that each of the four submatrices is orthogonal. Thus, it is expected that applying linear equalization to the effective channel followed by a slicer (or in general, a decoder) will exhibit some diversity gain.

The performance of the proposed scheme is demonstrated in Figure 5 for a simple scenario where the users transmit uncoded 16-QAM symbols. At the central receiver linear MMSE equalization is applied, followed by a slicer. As a baseline for comparison, we consider a relay that quantizes and forwards the output of an arbitrary relay; or alternatively, a relay that quantizes and forwards the output of both relays but with half the number of bits allocated to each quantizer. [2] The latter is referred to as "no combining" in Figure 5.

Substantial improvement may be seen with respect to the baseline schemes when a low bit error probability is desired, where we have considered quantization rates of 4, 6 and 8 bits per sample for each relay.[3]

V. Extensions to More than Two Antennas

As in the case of space-time modulation for channel coding, extension of the scheme to more receive antennas is possible, albeit with some loss.

A natural approach is to try utilizing the theory of orthogonal designs. It should be noted however that it is well known that the decoding delay (number of time instances stacked together) roughly grows exponentially with the number of antennas. Another possible avenue is to try to follow the approach of quasi-orthogonal space-time codes as developed in [6]–[8]. We demonstrate both approaches.

[2]Since we assume "dumb" relays, the quantization of the inputs to the receiver was performed using a fixed (SNR independent) loading factor, taken as three times the standard deviation of the noise-free input to the quantizer.

[3]As the gain are more pronounced at high SNR, we chose to demonstrate the performance of 16-QAM rather than QPSK transmission.

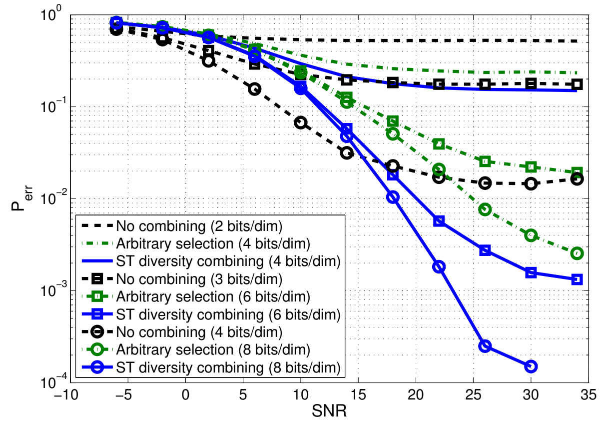

Fig. 5. Bit error rate achieved using "dumb" relaying using the proposed diversity-combing scheme and comparison to baseline relaying schemes for uncoded 16-QAM transmission over a Rayleigh fading environment, where the receiver employs MMSE equalization.

Attempting to apply orthogonal designs, one immediately confronts a basic obstacle due to the fact that rate-1 complex orthogonal designs do not exist beyond the case of two antennas [9]. We next demonstrate the problem that arises and also show how it may be resolved by judiciously combining balanced rate-$1/2$ orthogonal designs [10] (which include the four basic OSTBCs described in [9] for 2-8 antennas) with repeated quantization used in conjunction with multiplicative dithering. For the sake of concreteness and ease of exposition, we demonstrate the method for the case of a SIMO system with $M = 4$ receive antennas.

The received signals are given by (1) where now $i = 1, \ldots, M$ (with $M = 4$). We proceed by stacking $T = 8$ time instances of the received signal from the 4 antennas and build an effective real-valued vector by decomposing each entry into its real and imaginary components, just as is done in (2). This yields for $M = 4$, a vector $\mathbf{s}$ of dimension $2 \times 4 \times 8 = 64$.

By reinterpreting the rate-$1/2$ orthogonal design of 4 transmit antennas (see [9]), we arrive at a 8×64 transformation matrix $\mathbf{P}$.[4] Next, we form a 8×1 real vector $\mathbf{y}$ by applying to the effective received vector $\mathbf{s}$, formed in the manner described in (2), the transformation $\mathbf{y} = \mathbf{P}\mathbf{s}$.

It can be shown that the following holds

$$\begin{aligned} \mathbf{y} &= \frac{\sqrt{2}\|\mathbf{h}\|}{\sqrt{8}}\mathbf{U}(h_1, h_2, h_3, h_4)\mathbf{x} + \mathbf{U}\mathbf{n} \\ &= \frac{\|\mathbf{h}\|}{2}\mathbf{U}(h_1, h_2, h_3, h_4)\mathbf{x} + \mathbf{n}', \end{aligned} \tag{18}$$

where $\mathbf{U}(h_1, h_2, h_3, h_4)$ is a unitary matrix.[5] Here, the vector $\mathbf{x}$ is the 16-dimensional real representation of the transmitted signal over $T = 8$ time instances, formed analogously to (3).

[4]The specific form of $\mathbf{P}$ can be found in Equation (36) in [11].

[5]The specific form of $\mathbf{U}(h_1, h_2, h_3, h_4)$ is given in Equation (29) in [11].

We note that rows of $\mathbf{U}(h_1, h_2, h_3, h_4)$ are orthogonal for any values of $h_1, \ldots, h_4$. It follows that $\mathbf{n}'$ is white (and Gaussian with unit variance).

The problem with using a non-rate 1 orthogonal design now becomes clear. Unlike $\mathbf{U}(h_1, h_2)$ (see (6)) which is square, $\mathbf{U}(h_1, h_2, h_3, h_4)$ is non-square and hence is non-invertible. We overcome this obstacle by passing the same observation vector $\mathbf{s}$ via a "dithered" version of $\mathbf{P}$, such that another set of 8 mutually orthogonal measurement rows is attained. Specifically, let us define a 4 dimensional vector $d = (d_1, d_2, d_3, d_4)$ where d_i are complex numbers of unit magnitude (pure phases). We form a dithered version of the antenna outputs as

$$\tilde{s}_i(t) = d_i \cdot s_i(t), \tag{19}$$

where d_i does not depend on t. We assume that the d_i are drawn at random as i.i.d. uniform phases.

We may associate with $\tilde{s}_i(t)$, $t = 1, \ldots, T = 8$, the effective 64-dimensional real vector $\tilde{\mathbf{s}}$. Next, we obtain another 8-dimensional real vector $\tilde{\mathbf{u}}$ by applying to the vector $\tilde{\mathbf{s}}$ the transformation $\tilde{\mathbf{y}} = \mathbf{P}\tilde{\mathbf{s}}$. We therefore obtain

$$\tilde{\mathbf{y}} = \frac{\|\mathbf{h}\|}{2}\mathbf{U}(d_1h_1, d_2h_2, d_3h_3, d_4h_4)\mathbf{x} + \mathbf{n}'', \tag{20}$$

where $\mathbf{n}''$ is distributed as $\mathbf{n}'$.

Note that the dithers drawn implicitly (via (19)) may be absorbed in $\mathbf{P}$, thus defining a "dithered" combining matrix $\mathbf{P}_{\text{dith}}$. Combining (18) and (20), we have

$$\underbrace{\begin{bmatrix} \mathbf{y} \\ \hline \tilde{\mathbf{y}} \end{bmatrix}}_{\mathbf{y}_{\text{eff,dith}}} = \underbrace{\frac{\|\mathbf{h}\|}{2}\begin{bmatrix} \mathbf{U}(h_1, h_2, h_3, h_4) \\ \hline \mathbf{U}(d_1h_1, d_2h_2, d_3h_3, d_4h_4) \end{bmatrix}}_{\mathcal{F}_{\text{dith}}}\mathbf{x} + \begin{bmatrix} \mathbf{n}' \\ \mathbf{n}'' \end{bmatrix}. \tag{21}$$

Finally, we apply component-wise quantization to obtain

$$\mathbf{y}_q = Q(\mathbf{y}_{\text{eff,dith}}). \tag{22}$$

We may then recover an estimate of $\mathbf{x}$ by applying the inverse of $\mathcal{F}$ to $\mathbf{y}$ or a linear MMSE estimator.

As mentioned above, another approach to extend the basic scheme to more antennas is to borrow ideas from quasi-orthogonal space-time codes. As an example for a quasi-orthogonal space-time linear combining matrix, we construct a matrix $\mathbf{P}_{\text{quasi}}$ by taking half of the columns of $\mathbf{P}$, specifically columns $1 - 16$ and $49 - 64$, scaling by $\sqrt{2}$ to maintain orthonormality. This results in

$$\mathbf{y}_{\text{eff,quasi}} = \underbrace{\frac{\|\mathbf{h}\|}{2}\mathbf{U}_{\text{quasi}}(h_1, h_2, h_3, h_4)}_{\mathcal{F}_{\text{quasi}}}\mathbf{x} + \mathbf{n}' \tag{23}$$

where $\mathbf{U}_{\text{quasi}}(h_1, h_2, h_3, h_4)$ is given by Equation (35) in [11].

We tested the performance attained with both combining matrices in the scenario considered in Section III. Specifically, Figure 6 depicts the performance achieved for uncoded QPSK transmission when using different linear-combining schemes, for the case Rayleigh fading 4×1 SIMO channel.

First, we observe that both $\mathbf{P}_{\text{dith}}$ and $\mathbf{P}_{\text{quasi}}$ achieve similar performance. Note, however, that $\mathbf{P}_{\text{quasi}}$ utilizes four consecutive symbols rather than eight and hence induces less latency. On the other hand, the construction of $\mathbf{P}_{\text{dith}}$ can be readily extended to more antennas.

We further observe that as both variants of space-time diversity combining do not achieve orthogonality, the gap from optimal combining (MRC) is larger. Optimal antenna selection has a gap of 3.5 dB from MRC. Yet as the number of antennas increases, the complexity of performing optimal selection increases as well.

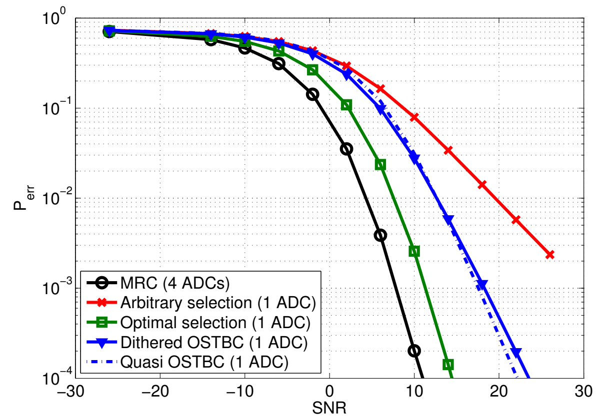

Fig. 6. Bit error rate for uncoded QPSK transmission over a 4×1 Rayleigh fading channel with reception employing the proposed space-time diversity combining scheme, using $\mathbf{P}_{\text{dith}}$ and $\mathbf{P}_{\text{quasi}}$, and comparison with alternatives.

References

[1] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, no. 6, pp. 1075–1102, 1959.

[2] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, 2004.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[4] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3413–3431, 2007.

[5] M. N. Khormuji and M. Skoglund, "On instantaneous relaying," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3378–3394, 2010.

[6] O. Tirkkonen, A. Boariu, and A. Hottinen, "Minimal non-orthogonality rate 1 space-time block code for 3+ tx antennas," in *Spread Spectrum Techniques and Applications, 2000 IEEE Sixth International Symposium on*, vol. 2. IEEE, 2000, pp. 429–432.

[7] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 1–4, 2001.

[8] N. Sharma and C. B. Papadias, "Improved quasi-orthogonal codes through constellation rotation," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 332–335, 2003.

[9] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[10] S. S. Adams, J. Davis, N. Karst, M. K. Murugan, B. Lee, M. Crawford, and C. Greeley, "Novel classes of minimal delay and low PAPR rate 1/2 complex orthogonal designs," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2254–2262, 2011.

[11] E. Domanovitz and U. Erez, "Diversity combining via universal orthogonal space-time transformations," *arXiv preprint arXiv:1712.04038*, 2017.