Decision-Feedback Differential Detection with Optimum Detection Order Metric for Noncoherent Massive MIMO Systems

George Yammine, Robert F.H. Fischer
Institut für Nachrichtentechnik, Universität Ulm, Ulm, Germany
Email: {george.yammine, robert.fischer}@uni-ulm.de

***Abstract*—Noncoherent detection in massive MIMO systems is an attractive alternative to channel-estimation-based detection. So far, receivers employing ordered decision-feedback differential detection based on the phase quantization error of the candidate symbols of the correlation matrix have been proposed. This approach however ignores reliability information in certain cases. In this paper, we analytically derive the maximum-likelihood decision metric used to search for the optimum detection order. We compare the proposed metric to the phase-quantization-based one in terms symbol error rate performance.**

I. Introduction

Over the last years, systems where the base stations are equipped with a very large number of antennas compared to the number of users they serve, so-called *massive multiple-input/multiple-output (MIMO)*, have gained more and more attention, e.g., [8], [9]. In such massive MIMO systems, the number of channel coefficients is extremely large, presenting a challenge when acquiring accurate estimates thereof. Additionally, as the number of pilot sequences is limited, these estimates may suffer from additional interference due to pilot contamination [3].

By employing *noncoherent detection* schemes [11], the need for channel estimation is eliminated. Notably, *multiple-symbol differential detection* (MSDD), or the lower-complexity *decision-feedback differential detection* (DFDD) can be applied. Additionally, *noncoherent decision-feedback equalization* (nDFE) [4] enhances performance in the multi-user case.

To maximize the performance of noncoherent detection, the DFDD should be performed with an optimized ordering. In [11], this is achieved via a sorting based on the phase quantization error, i.e., the symbol with the smallest phase deviation from its quantized value is detected. However, since only the phase information is taken into consideration, this may lead to choosing very unreliable symbols for detection in certain cases. Hence, a different metric for optimizing the detection order is required.

In this paper, we derive the *maximum-likelihood* metric, based on the statistics of the elements of the correlation matrix, for optimized ordering in noncoherent massive MIMO detection. The performance of the phase-quantization-error-based approach [11], [4] to the newly proposed is compared.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — FI 982/12-1.

The paper is organized as follows. In Sec. II, a recapitulation of the system model and multi-user detection in noncoherent massive MIMO systems is presented. In Sec. III, we derive the maximum-likelihood metric to be used in the detection algorithm. Numerical simulation results are provided in Sec. IV. Finally, Sec. V gives some conclusions.

II. System Model and Noncoherent Detection

A *multi-user uplink scenario* where a small number of single-antenna users transmit to a central receiver employing N_{rx} antennas, cf. Fig. 1, is considered. The base station array has a much larger number of antennas than users, i.e., $N_{\mathsf{rx}} \gg N_{\mathsf{u}}$.

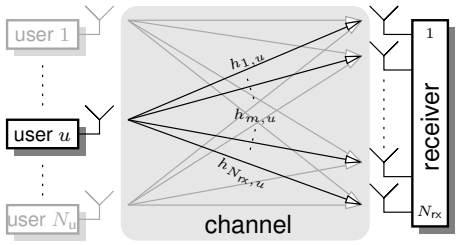

Figure 1. Multi-user massive MIMO uplink system [11]. N_{u} users transmit to a central base station with $N_{\mathsf{rx}} \gg N_{\mathsf{u}}$ antennas.

At each discrete time step k, each user u transmits an M-ary differentially-encoded phase-shift keying (DPSK) symbol $b_{k,u}$ drawn from the constellation

$$\mathcal{M} \stackrel{\text{def}}{=} \{\mathrm{e}^{\mathrm{j}2\pi\cdot i/M} \mid i = 0, 1, \ldots, M-1\} \,. \tag{1}$$

The transmit symbols are generated from the PSK data symbols $a_{k,u} \in \mathcal{M}$ as [11]

$$b_{k,u} = a_{k,u} b_{k-1,u} \,, \quad b_{0,u} = 1 \,. \tag{2}$$

The coefficients $h_{m,u}$ describing the channel between user u and antenna m are i.i.d.ly randomly drawn. Assuming a block-fading model, these coefficients remain constant over each burst of N_{bl} symbols.

Collecting N_{bl} time steps over the N_{rx} antennas, the receive block $\boldsymbol{R}$ is constructed and given by

$$\boldsymbol{R} = \sum\nolimits_{u=1}^{N_{\mathsf{u}}} \boldsymbol{h}_u \boldsymbol{b}_u + \boldsymbol{N} \,, \tag{3}$$

where the column vector $\boldsymbol{h}_u \stackrel{\text{def}}{=} [h_{1,u}, \ldots, h_{N_{\text{rx}},u}]^{\mathsf{T}}$ collects the channel coefficients for user u, and the row vector $\boldsymbol{b}_u \stackrel{\text{def}}{=} [b_{0,u}, \ldots, b_{N_{\text{bl}}-1,u}]$ is the transmit symbols stream of user u [11]. The zero-mean circular-symmetric Gaussian noise samples $n_{m,k}$ with variance σ_{n}^2 are collected in the matrix $\boldsymbol{N}$.

To specify the channel, the average receive power $P_{m,u} = \mathrm{E}\{|h_{m,u}|^2\}$ induced by user u on each antenna m, called the *power-space profile (PSP)* is used [11], [13].

At the base station, noncoherent detection of the symbols of user u is applied. In case of DPSK, *multiple-symbol differential detection (MSDD)* [2] is known to improve performance significantly. Alternatively, *decision-feedback differential detection (DFDD)* [1], [12] can be applied to avoid the large computational complexity of MSDD.

In [11], a correlation-based decision-feedback differential detection method was utilized. To that end, first, the $N_{\text{bl}} \times N_{\text{bl}}$ *correlation matrix* is constructed

$$\boldsymbol{Z}_u \stackrel{\text{def}}{=} \boldsymbol{R}^{\mathsf{H}} \boldsymbol{W}_u \boldsymbol{R} \,, \tag{4}$$

with the user-specific diagonal weighting matrix

$$\boldsymbol{W}_u \stackrel{\text{def}}{=} \mathbf{diag}(w_{1,u}, \ldots, w_{N_{\text{rx}},u}) \,. \tag{5}$$

Then, optimally-sorted DFDD is applied. To obtain the sorting, at each detection step ℓ, the candidate symbol

$$\bar{z}_{u,\ell,\bar{\kappa}_\ell} \stackrel{\text{def}}{=} \sum\nolimits_{n=0}^{\ell-1} \hat{b}_{\ell,u} z_{u,n,\bar{\kappa}_\ell} \,, \tag{6}$$

which is the sum of the entries of $\boldsymbol{Z}_u$ multiplied by their corresponding already-detected symbols $\hat{b}_{\ell,u}$ for each previous step n, or feedback information, with the smallest phase quantization error

$$\hat{\kappa}_\ell = \underset{\bar{\kappa} \in \{1,\ldots,N_{\text{bl}}\} \setminus \{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}{\operatorname{argmin}} \left| \Delta\mathcal{Q}_{\mathsf{PSK}}\{\bar{z}_{u,\ell,\bar{\kappa}}\} \right| , \tag{7}$$

with

$$\Delta\mathcal{Q}_{\mathsf{PSK}}\{z\} \stackrel{\text{def}}{=} \mathrm{mod}_{\text{s},2\pi}\left(\arg(z) - \frac{2\pi}{M} \left\lfloor \frac{M}{2\pi} \cdot \arg(z) \right\rceil \right) \tag{8}$$

is detected ($\mathrm{mod}_{\text{s},2\pi}(\cdot)$ is the symmetrical modulo operation, i.e., reduction into the interval $(-\pi, +\pi]$).

To reduce multi-user interference, *noncoherent decision-feedback equalization (nDFE)* [4] over the users can be employed, leading to significant performance improvements.

III. Maximum-Likelihood Decision Metric

It is well known that large performance improvements, e.g., reduced error rates, can be achieved by taking the decisions in an optimized order [11]. In the context of noncoherent massive MIMO, the sorting has been so far based on the phase quantization error (8) of the candidate symbols.

However, looking closer at this metric reveals an inherent problem. To illustrate this, the following scenario is considered. For readability, the subscript indices are dropped, but the scenario of the previous section remains, and it should be understood that the following describes the situation for $\bar{z}_{u,\ell,\bar{\kappa}_\ell}$ at Step ℓ. A candidate symbol $\bar{z}$ is given. This symbol comprise a scaled version of the useful information $\bar{b} \stackrel{\text{def}}{=} \eta b$, $b \in \mathcal{M}$, and some noise ψ, i.e.,

$$\bar{z} \stackrel{\text{def}}{=} \bar{b} + \psi \,. \tag{9}$$

The scaling η is real-valued, i.e., $\bar{b}$ shifts on its axis of direction only. Without added noise, no matter where the symbols lies, it will be quantized to the closest PSK symbol based on the deviation in phase (e.g., $\bar{b} = 0.1$ or $\bar{b} = 1.1$ will be both quantized to the same symbol $\hat{b} = 1$). However, when the complex-valued noise is superimposed, the impact upon differentially-encoded data symbols that are close to the origin of the complex-plane will be much larger in terms of the effect on the phase of these symbols than those farther from the origin.

If the distance to the origin of candidate symbols $\bar{z}$ (amplitude of $\bar{z}$) is ignored, there is a chance that the noise ψ has shifted the symbol $\bar{b}$ enough to the next neighboring symbol from the transmit alphabet. Based on phase information alone, this symbol would appear highly reliable to detect with minimal phase quantization error, compared to one farther away from the origin of the complex-plane affected by slightly more noise.

The phase-only metric, i.e., the phase quantization error, is not optimum. Hence, the optimum metric is still to be derived. Given equally-likely transmitted symbols, the optimum decision strategy would be the maximum-likelihood strategy, i.e.,

$$\hat{b} \stackrel{\text{def}}{=} \underset{b \in \mathcal{M}}{\operatorname{argmax}}\, \mathsf{f}_{\bar{\mathsf{z}}}(\bar{z} \mid b) \,. \tag{10}$$

To that end, the conditional probability density function (pdf) and statistics of the data symbols of user u, the interference caused by users $\nu \neq u$ on u, and the noise sources present in the the candidate symbols $\bar{z}$ at each detection step ℓ are required.

Returning to the specific massive MIMO model, we have the following. The data symbols b are taken from a PSK alphabet $\mathcal{M}$. The scaling factor

$$\eta \stackrel{\text{def}}{=} \ell \sum\nolimits_{m=1}^{N_{\text{rx}}} h_{m,u}^* w_{m,u} h_{m,u} \,, \tag{11}$$

is approximately real-valued Gaussian distributed [4] with mean and variance

$$\mu_\eta \stackrel{\text{def}}{=} \mathrm{E}\{\eta\} = \ell \sum\nolimits_{m=1}^{N_{\text{rx}}} w_{m,u} P_{m,u} \tag{12}$$

$$\sigma_\eta^2 \stackrel{\text{def}}{=} \mathrm{E}\{|\eta - \mu_\eta|^2\} = \ell^2 \sum\nolimits_{m=1}^{N_{\text{rx}}} w_{m,u}^2 P_{m,u}^2 \,. \tag{13}$$

This means that $\bar{b}$ is now a real-valued Gaussian bell, centered at $\mu_\eta b$ in the complex-plane on a line that intersects the origin and b. As an example, a histogram plot, measured at $E_{\text{s}}/N_0 \mathrel{\widehat{=}} 12$ dB showing the union of the different distributions of $\bar{b}$ in the 4PSK case is given in Fig. 2.

The term ψ collects the portion of the interference (that is not reduced via nDFE) and cross-interference caused by other users $\nu \neq u$, and the different noise sources. In this context, the distribution and statistics the sum of the individual terms is of interest.

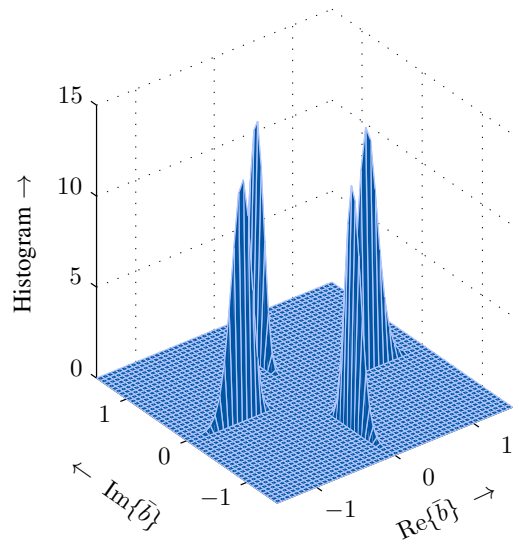

Figure 2. Histogram plot of the union of the different distributions of $\bar{b}$ in the 4PSK case, measured at $E_{\mathsf{s}}/N_0 \mathrel{\widehat{=}} 12$ dB.

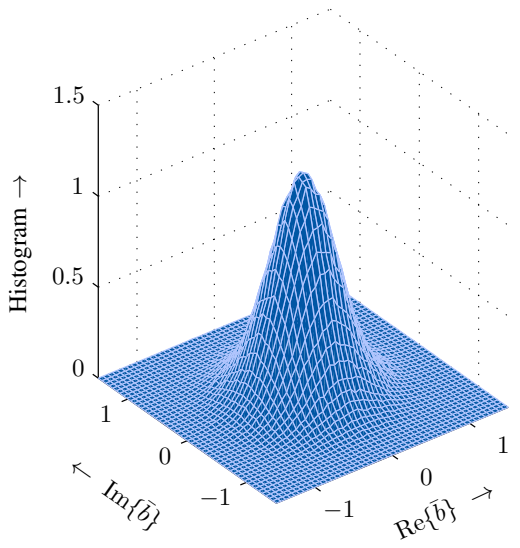

Figure 3. Histogram plot of ψ measured at $E_{\mathsf{s}}/N_0 \mathrel{\widehat{=}} 12$ dB. Since Lindeberg's condition is satisfied, the sum if the independently but not identically distributed random variables converges to a zero-mean complex-valued Gaussian distribution.

The sum is formed of independent *but not* identically distributed random variables. These different disturbances contribute a small portion to ψ, i.e., the individual variances are arbitrarily small compared to their sum (an in depth derivation of these variances are found in [14, Appendix A]). Hence, Lindeberg's condition [6] is satisfied, and the interference plus noise terms converge to a zero-mean complex-valued Gaussian distributed random variable ψ with variance ($\mathcal{D}$: set of previously-detected users)

$$\sigma_\psi^2 \stackrel{\text{def}}{=} \sum\nolimits_{\substack{\nu\notin\mathcal{D}\\ \nu\neq u}}^{N_{\mathsf{u}}} \mu_{u,\nu} + \sum\nolimits_{\substack{\nu=1\\ \nu\neq u}}^{N_{\mathsf{u}}} \sigma_{u,\nu}^2 + \sum\nolimits_{\substack{\nu,\mu=1\\ \nu<\mu}}^{N_{\mathsf{u}}} \sigma_{u,\nu,\mu}^2 + \sigma_{\mathsf{n},u}^2 + \sum\nolimits_{\substack{\nu=1\\ \nu\neq u}}^{N_{\mathsf{u}}} \sigma_{\mathsf{n},u,\nu}^2 + \sigma_{\mathsf{n},\mathsf{n}}^2 \,, \tag{14}$$

where

$$\mu_{u,\nu} = \ell\Big(\sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u} P_{m,\nu}\Big)^2 \,, \tag{15}$$

$$\sigma_{u,\nu}^2 = \ell \sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u}^2 P_{m,\nu}^2 \,, \tag{16}$$

$$\sigma_{u,\nu,\mu}^2 = 2\ell \sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u}^2 P_{m,\nu} P_{m,\mu} \,, \tag{17}$$

$$\sigma_{\mathsf{n},u}^2 = \sigma_{\mathsf{n}}^2(\ell^2+\ell) \sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u}^2 P_{m,u} \,, \tag{18}$$

$$\sigma_{\mathsf{n},u,\nu}^2 = 2\ell\sigma_{\mathsf{n}}^2 \sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u}^2 P_{m,\nu} \,, \tag{19}$$

and finally,

$$\sigma_{\mathsf{n},\mathsf{n}}^2 = \ell \sum\nolimits_{m=1}^{N_{\mathsf{rx}}} w_{m,u}^2 \,. \tag{20}$$

As example, a histogram plot, measured at $E_{\mathsf{s}}/N_0 \mathrel{\widehat{=}} 12$ dB, of the sum-noise-plus-interference term ψ is given in Fig. 3. Here, it can be clearly seen that, indeed, the sum converges to a zero-mean complex-valued Gaussian distribution.

For a given b, the candidate symbol $\bar{z}$ is a superposition of a real-valued Gaussian bell and a zero-mean complex-valued one, i.e., $\bar{z}$ conditioned on b is described by an *ellipsoidal* complex-valued Gaussian bell

$$\mathrm{f}_{\bar{\mathsf{z}}}(\bar{z} \mid b) = \frac{1}{\pi\sqrt{\det(\boldsymbol{\Sigma})}} \mathrm{e}^{-(\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}})^{\mathsf{T}}\Sigma^{-1}(\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}})} \,, \tag{21}$$

where $\det(\cdot)$ is the determinant of a matrix, and

$$\bar{\boldsymbol{z}} \stackrel{\text{def}}{=} \begin{bmatrix}\mathrm{Re}\{\bar{z}\}\\ \mathrm{Im}\{\bar{z}\}\end{bmatrix} \,, \quad \text{and} \quad \bar{\boldsymbol{b}} \stackrel{\text{def}}{=} \mu_\eta \begin{bmatrix}\mathrm{Re}\{b\}\\ \mathrm{Im}\{b\}\end{bmatrix} \,, \tag{22}$$

are the real-valued vector representation of the complex-valued quantities $\bar{z}$ and $\bar{b}$, and

$$\Sigma \stackrel{\text{def}}{=} \begin{bmatrix}\alpha^2(2\sigma_\eta^2+\sigma_\psi^2)+\beta^2\sigma_\psi^2 & 2\alpha\beta\sigma_\eta^2\\ 2\alpha\beta\sigma_\eta^2 & \beta^2(2\sigma_\eta^2+\sigma_\psi^2)+\alpha^2\sigma_\psi^2\end{bmatrix} \tag{23}$$

is the covariance matrix of the ellipsoidal distribution, with

$$\alpha \stackrel{\text{def}}{=} \mathrm{Re}\{b\} \,, \quad \text{and} \quad \beta \stackrel{\text{def}}{=} \mathrm{Im}\{b\} \,, \tag{24}$$

defining the rotation (direction of the major axis) thereof.

The optimum decision at Step ℓ is then obtained by finding the index $\hat{\kappa}_\ell$ and estimate $\hat{b}_{\hat{\kappa}_\ell}$ from the remaining candidate symbols $\bar{z}_{u,\ell,\bar{\kappa}_\ell}$ that maximizes (21),

$$\hat{b}_{\hat{\kappa}_\ell} = \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmax}} \mathrm{f}_{\bar{\mathsf{z}}}(\bar{z}_{u,\ell,\bar{\kappa}_\ell} \mid b) \,,$$

which, by ignoring constants that do not affect the outcome of the maximization, and applying $\log(\cdot)$ to the conditional pdf, is simplified to

$$\begin{aligned}\hat{b}_{\bar{\kappa}_\ell} &= \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmax}} \Big(-(\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}})^{\mathsf{T}}\Sigma^{-1}(\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}})\Big) \\ &= \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmin}} (\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}})^{\mathsf{T}}\Sigma^{-1}(\bar{\boldsymbol{z}}-\bar{\boldsymbol{b}}) \,, \end{aligned} \quad (25)$$

i.e., the optimum decision is obtained by searching for $\bar{z}_{u,\ell,\bar{\kappa}_\ell}$ smallest squared Mahalanobis distance [7] to $b \in \mathcal{M}$. Here, unlike the phase-based approach, cf. Section II, the sorting and quantization to PSK symbols are not done separately. Compared to the pseudo-code representation in [11, Fig. 3], only Lines 8 and 9 are replaced by (25). In fact, the inverse of the covariance matrix Σ can be pre-computed and stored for employed PSK alphabet at the different detection steps ℓ. This means that, at the end, no increase in complexity is incurred.

Another approach would be to implement a suboptimum metric, based on an approximation of the conditional pdf given by (21). This is obtained by ignoring part of the statistical information, in this case the variance σ_η^2.

In the low-to-moderate SNR regime, the variance of the differentially-encoded data σ_η^2 is much smaller than that of interference-plus-noise term σ_ψ^2. This means that instead of the ellipsoidal complex Gaussian bells, the conditional pdf of $\bar{z}_{u,\ell,\bar{\kappa}_\ell}$ can be approximated by a spherical complex-valued Gaussian distribution, centered at $\mu_\eta b$

$$\mathsf{f}_{\bar{\mathsf{z}},\text{approx.}}(\bar{z}_{u,\ell,\bar{\kappa}_\ell} \mid b) = \frac{1}{\pi\sigma_\psi^2}\mathrm{e}^{-\frac{|\bar{z}_{u,\ell,\bar{\kappa}_\ell}-\mu_\eta b|^2}{\sigma_\psi^2}} \,. \quad (26)$$

Searching for $\hat{\kappa}_\ell$ that maximizes the approximate pdf gives

$$\hat{b}_{\bar{\kappa}_\ell} = \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmax}} \mathsf{f}_{\bar{\mathsf{z}},\text{approx.}}(\bar{z}_{u,\ell,\bar{\kappa}_\ell} \mid b) \,,$$

which is simplified by applying $\log(\cdot)$ to $\mathsf{f}_{\bar{\mathsf{z}},\text{approx.}}(\bar{z}_{u,\ell,\bar{\kappa}_\ell} \mid b)$ and ignoring constants that are irrelevant for the outcome of the maximization,

$$\begin{aligned}\hat{b}_{\bar{\kappa}_\ell} &= \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmax}} \Big(-|\bar{z}_{u,\ell,\bar{\kappa}_\ell}-\mu_\eta b|^2\Big) \\ &= \underset{\substack{b\in\mathcal{M}\\ \bar{\kappa}\in\{1,\ldots,N_{\mathsf{bl}}\}\setminus\{\hat{\kappa}_1,\ldots,\hat{\kappa}_{\ell-1}\}}}{\operatorname{argmin}} |\bar{z}_{u,\ell,\bar{\kappa}_\ell}-\mu_\eta b|^2 \,, \end{aligned} \quad (27)$$

i.e., the problem reduces to a minimum squared Euclidean distance search.

As before, Lines 8 and 9 of the pseudo-code found in [11, Fig. 3] are replaced. With this approximation, the only required statistic is μ_η.

IV. Numerical Simulations

Numerical simulations were conducted for an $N_{\mathsf{u}} = 3$ multi-user scenario. The users are positioned at $m_1 = 20$ (■), $m_2 = 50$ (■), and $m_3 = 85$ (■), and transmit a burst of $N_{\mathsf{bl}} = 201$ (200 data and 1 reference) symbols, taken from a quaternary ($M = 4$) DPSK alphabet. The central base station is equipped with an $N_{\mathsf{rx}} = 100$ uniform linear array. The antennas at the receiver side are defined by the number of equidistant zeros in their horizontal radiation pattern N_{notch} (all elements in the array utilize the same pattern).

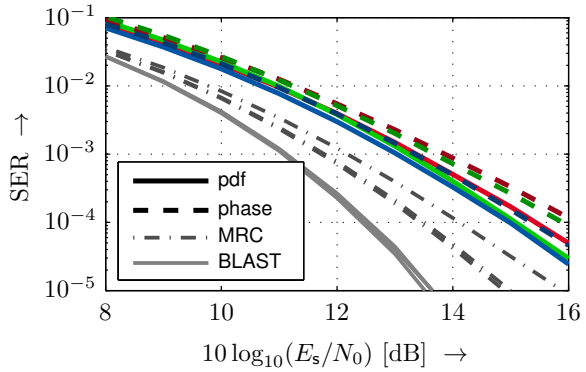

Figure 4. Symbol error rate vs. E_{s}/N_0 (in dB) for a high multi-user interference scenario. $N_{\mathsf{u}} = 3$; $m_1 = 20$ (■), $m_2 = 50$ (■), and $m_3 = 85$ (■). Uniform linear array with $N_{\mathsf{rx}} = 100$ antennas, with $N_{\mathsf{notch}} = 0$. Burst length $N_{\mathsf{bl}} = 201$. nDFE user detection order: (■)→(■)→(■).

The geometrical channel model of [13] is presumed, which is described via the power-space profiles of all users. The PSPs are known at the receiver, but the actual channel realizations are not. Power control at the transmitters is assumed, such that the sum receive power $E_{\mathsf{h},u} = \sum_m P_{m,u}$ of each user u is normalized to one, i.e., $E_{\mathsf{h},u} = E_{\mathsf{h}} = 1$. The average received energy per symbol is $E_{\mathsf{s}} = E_{\mathsf{h}}$. The signal-to-noise ratio is defined as $E_{\mathsf{s}}/N_0 = 1/\sigma_{\mathsf{n}}^2$ (N_0 is the (two-sided in ECB domain) noise power spectral density) due to PSK signaling. Each channel coefficient $h_{m,u}$ is drawn i.i.d. from zero-mean circular-symmetric complex Gaussian distributions with variance $P_{m,u}$.

The required windowing coefficients $w_{m,u}$ of each user u are obtained through offline numerical optimization using MATLAB's sequential quadratic programming solver. Additionally, the user detection order for nDFE is obtained during this process.

A. High Multi-User Interference

To begin with, $N_{\mathsf{notch}} = 0$ is chosen, leading to a channel model with a high overlap in the PSPs (high multi-user interference) of the three users (details of the channel model can be found in [13]). The symbol error rate (SER) vs. E_{s}/N_0 (in dB) is plotted in Fig. 4. The user detection order in the nDFE scheme is given as user 3 (■) followed by user 1 (■) and then user 2 (■). Here, a gain of about 1 dB for user 1 (■) and user 2 (■) is achieved when applying detection with pdf-based (maximum-likelihood metric) detection (solid curves), compared to the straightforward phase-quantization-based approach (dashed curves). In case of user 3 (■), we see an increase in performance of approximately 0.5 dB.

SER curves for coherent detection using maximum-ratio combining (MRC), and BLAST[1] [5] with perfect channel

[1]Bell Laboratories Layered Space-Time: successive interference cancellation over the users.

Table I
SYMBOL ERROR RATE OF NONCOHERENT DETECTION USING THE OPTIMAL AND SUBOPTIMAL METRIC AT HIGH SNR. FOR COMPARISON, THE SERS OF THE CONVENTIONAL PHASE-BASED DFDD IS GIVEN. ALL SER VALUES IN THE TABLE ARE $\times 10^{-5}$.

		E_s/N_0 (in dB)			
		17	17.5	18	18.5
SER (Optimal)	user 1 (■)	1.471	0.820	0.419	0.241
	user 2 (■)	0.787	0.408	0.209	0.120
	user 3 (■)	0.640	0.308	0.126	0.060
SER (Suboptimal)	user 1 (■)	1.485	0.842	0.424	0.245
	user 2 (■)	0.788	0.407	0.209	0.121
	user 3 (■)	0.644	0.311	0.127	0.061
SER (Phase-based)	user 1 (■)	3.602	2.0	1.226	0.746
	user 2 (■)	3.07	1.832	1.118	0.638
	user 3 (■)	1.31	0.651	0.294	0.165

knowledge, which is superior to matched filtering only (MRC), are given for reference in dash-dot and solid gray, respectively.

Noteworthy, the curves for detection based on the suboptimum metric (not shown) are almost indistinguishable from that for the optimal one. This means that the approximation holds well in this SNR range. In fact, only when going to higher SNRs, a measurable difference in the SERs can be seen. To highlight this, the error rates measured at $E_s/N_0 \mathrel{\widehat{=}} 17$ dB to $E_s/N_0 \mathrel{\widehat{=}} 18.5$ dB in 0.5 dB increments are collected in Table I. Although the difference is small, indeed the suboptimal metric using the squared Euclidean distance performs slightly worse, i.e., information which given by the Mahalanobis distance is lost. For comparison, the SERs using phase-based DFDD are given.

B. Low Multi-User Interference

Now, $N_{\text{notch}} = 4$ is set. This leads to a channel model with a much smaller overlap in the power-space profiles of the three users, when compared to the previous case. nDFE use the same user detection order ((■)→(■)→(■)). The symbol error rate vs. E_s/N_0 (in dB) is plotted in Fig. 5. Here, all users benefit from a small increase in SER performance, minimizing the gap between DFDD/nDFE and coherent detection with perfect channel knowledge. In fact, noncoherent detection using the ML metric performs almost as good as MRC.

It is worth noting that the power efficiency loss incurred due to noisy channel estimates is not considered here. This means that DFDD/nDFE with pdf-based sorting would outperform coherent detection by a large margin.

V. CONCLUSIONS

In this paper, the maximum-likelihood metric for optimum sorting for correlation-based detection in the noncoherent massive MIMO setting was derived. Employing this new metric incurs no additional complexity in the detection algorithms. As in the phase-quantization-error-based approach, only statistical knowledge at the receiver is required.

This improved receiver presents a low-complexity high-performing alternative to coherent detection with channel estimation. In certain scenarios, noncoherent detection with pdf-based sorting performs similarly to coherent detection.

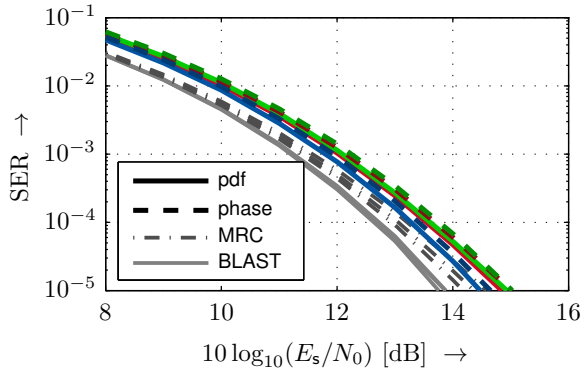

Figure 5. Symbol error rate vs. E_s/N_0 (in dB) for a low multi-user interference scenario. $N_u = 3$; $m_1 = 20$ (■), $m_2 = 50$ (■), and $m_3 = 85$ (■). Uniform linear array with $N_{rx} = 100$ antennas, with $N_{\text{notch}} = 4$. Burst length $N_{bl} = 201$. nDFE user detection order: (■)→(■)→(■).

When considering the loss due to noisy channel estimation in the coherent case, pdf-based DFDD/nDFE would provide a much better error-rate performance.

REFERENCES

[1] F. Adachi, M. Sawahashi. Decision-Feedback Multiple-Symbol Differential Detection for Mary DPSK. *Electronics Letters*, vol. 29, no. 15, pp. 1385–1387, July 1993.

[2] D. Divsalar, M.K. Simon. Multiple-Symbol Differential Detection of MPSK. *IEEE Trans. Communications*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[3] O. Elijah, C.Y. Leow, T.A. Rahman, S. Nunoo, S.Z. Iliya. A Comprehensive Survey of Pilot Contamination in Massive MIMO–5G System. *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 905–923, 2016.

[4] R.F.H. Fischer, M. Bense. Noncoherent Decision-Feedback Equalization in Massive MIMO Systems. In *Proceedings of International Zurich Seminar (IZS)*, pp. 112–115, Zurich, Switzerland, Feb. 2014.

[5] G.J. Foschini, D. Chizhik, M.J. Gans, C. Papadias, R.A. Valenzuela. Analysis and Performance of Some Basic Space-Time Architectures. *IEEE J. Sel. Areas in Communications*, vol. 21, no. 3, pp. 303–320, Apr. 2003.

[6] J.W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, vol. 15, no. 1, pp. 211–225, 1922.

[7] P.C. Mahalanobis. On the Generalized Distance in Statistics. *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

[8] T.L. Marzetta. Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas. *IEEE Trans. Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[9] F. Rusek, D. Persson, B.K. Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, F. Tufvesson. Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays. *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] A. Schenk. *Coding, Modulation, and Detection in Impulse-Radio Ultra-Wideband Communications*, PhD Thesis. University Erlangen-Nürnberg, 2012.

[11] A. Schenk, R.F.H. Fischer. Noncoherent Detection in Massive MIMO Systems. In *Proceedings of International ITG/IEEE Workshop on Smart Antennas*, Stuttgart, Germany, Mar. 2013.

[12] R. Schober, W.H. Gerstacker, J.B. Huber. Decision-Feedback Differential Detection of MDPSK for Flat Rayleigh Fading Channels. *IEEE Trans. Communications*, vol. 47, no. 7, pp. 1025–1035, July 1999.

[13] G. Yammine, R.F.H. Fischer, C. Waldschmidt. On the Influence of the Antenna Pattern in Noncoherent Massive MIMO Systems. In *Proceedings of The Twelfth International Symposium on Wireless Communication Systems*, Brussels, Belgium, Aug. 2015.

[14] G. Yammine, R.F.H. Fischer. Feedback-Aware Noncoherent Receivers for Massive MIMO Systems. In *Proceedings of International ITG/IEEE Workshop on Smart Antennas*, Hamburg, Germany, Feb. 2020.