SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

June 2016, Vol. 59 062302:1–062302:14
doi: 10.1007/s11432-015-5390-y

Adaptive robust Max-SLNR precoder for MU-MIMO-OFDM systems with imperfect CSI

Feng SHU[1,2,3*], Juanjuan TONG[1], Xiaohu YOU[2],
Chen GU[1] & Jiajun WU[1]

[1]*School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing* 210094*, China;*
[2]*National Mobile Communications Research Laboratory, Southeast University, Nanjing* 210096*, China;*
[3]*Ministerial Key Laboratory of JGMT, Nanjing University of Science and Technology, Nanjing* 210094*, China*

Received May 13, 2015; accepted June 15, 2015; published online September 6, 2015

Abstract The accuracy of channel state information (CSI) available at a base station (BS) has a direct impact on the performance of precoding in wideband multi-user multiple input, multiple output-orthogonal frequency division multiplexing (MIMO-OFDM) systems and depends on many factors, including: the delay between estimation and beamforming at the BS (also called the CSI delay), Doppler spread, the channel estimation method used, the average transmit power of pilot symbols, and the average number of pilot symbols that must be estimated per channel parameter. In this paper, the coefficient of CSI error needed to adapt to fading channels is modeled as a function of Doppler spread, CSI delay, and signal-to-noise ratio (SNR). In terms of the Gaussian-Markov CSI error model, an adaptive robust maximum signal-to-leakage-and-noise ratio (Max-SLNR) precoder is designed to track the statistical parameters of CSI error. The Doppler spread and SNR can be obtained through real-time estimation based on orthogonal pilot patterns. Simulation results show that, compared to non-adaptive robust and non-robust precoders of Max-SLNR, the proposed adaptive robust Max-SLNR precoder performs much better in terms of bit error rate (BER). Moreover, as either the average number of training symbols per channel parameter or the average transmit power increases, the BER performance of the proposed precoder approaches that of a precoder with ideal CSI.

Keywords multiuser MIMO, OFDM, Max-SLNR, adaptive, robust precoder, beamforming.

Citation Shu F, Tong J J, You X H, et al. Adaptive robust Max-SLNR precoder for MU-MIMO-OFDM systems with imperfect CSI. Sci China Inf Sci, 2016, 59(6): 062302, doi: 10.1007/s11432-015-5390-y

1 Introduction

Owing to its high-spectral efficiency, multiuser multiple input, multiple output-orthogonal frequency division multiplexing (MIMO-OFDM) has been adopted by fourth generation (4G) mobile communication standards such as LTE-advanced and is expected to remain a strong potential technology for use in beyond-4G (5G) mobile networks and the future-generation WLAN standards [1–7]. The key problem in MU-MIMO-OFDM systems is how to effectively control and compress interference among users in order to achieve a high sum-rate or reliable information transmission [3]. To solve this problem, several transmit

* Corresponding author (email: shufeng@njust.edu.cn)

© Science China Press and Springer-Verlag Berlin Heidelberg 2015

beamforming methods have been developed, including, among others: zero forcing (ZF); maximum signal-to-leakage-and-noise ratio (Max-SLNR) [7]; minimum mean square; error (MMSE) [8]; and singular value decomposition [9]. However, to implement these schemes, the transmitter must first have the channel state information (CSI). CSI can be obtained either through reverse channel estimation (in the time-division-duplex (TDD) model) or through feedback (in frequency-division-duplex (FDD) model). Because there is noise and fading in wireless channels, there always exists uncertainty or error in the estimated CSI, usually resulting from feedback delay, quantization error, or Gaussian estimation error. Because the performance of a system will be influenced by channel uncertainty, it is very useful to design robust precoding algorithms to avoid mismatched CSI. Several studies in the literature focus on using the zero forcing (ZF) criterion in order to develop robust beamforming algorithms that can be implemented at either the transmitter or receiver. In [10], the authors derive an expression for sum-rate based on ZF precoding assuming that there is imperfect CSI. They derive a robust factor that can convert the problem of robust multiuser MIMO scheduling into a traditional problem. By using this robust factor, global and local optimal scheduling schemes are obtained via different iterative algorithms. In [11, 12], the influences of imperfect CSI on bit error rate (BER) performance are investigated in the context of receiver ZF beamforming. A novel BER expression that takes imperfect CSI into account is derived for an M-QAM (quadrature amplitude modulation) signal in a TDD downlink MU-MIMO system using a ZF precoder, and BER performance is analyzed under different modulation schemes. Simulations have shown that, in time varying channels, CSI delay causes more performance degradation than channel estimation error at high signal to noise ratios (SNRs). Ref. [12] modeled the estimation error as independent complex Gaussian random variables in order to derive closed-form tight approximations for both the post-processing SNR distribution and BER for MIMO ZF receivers with M-QAM and M-PSK (phase shift keying) modulated signals in the presence of channel estimation error. Taking error variance from imperfect CSI into account, Refs. [8, 13] designed two robust precoders: Max-SLNR and MMSE, respectively. Using the concept of SLNR, Refs. [14,15] were able to maximize the average received signal power and ensure that leakage power was larger than a pre-specified threshold; by doing so, the non-convex optimization problem could easily be reformulated as a convex deterministic problem based on the expected values of the leakage power. In [16], the authors discuss the design of a statistically robust beamforming scheme through the maximization of the conditional expectation of the SLNR and the incorporation of a matched-filter combining vector into the beamformer design. In addition, several robust precoders based on differing criteria in order to improve the robustness of systems have been proposed, such as those in [17–21].

Because of the time-varying property of fading channels, channel estimation error has time-varying (i.e., not fixed) statistical moments in its variance or in its covariance matrix. In the studies above, the covariance matrix or the variance of the channel estimation error was usually assumed to be fixed, which is not true in actual fading channels. To track the uncertainty of estimated CSI error in a fading channel, it is very important to construct an adaptive robust precoder that can follow changes in the channel. In this study, the Max-SLNR criterion is used to design an adaptive robust precoder that can adapt to varying channel. Through a real-time estimation of Doppler spread and SNR on pilot grids, we are able to periodically update the beamforming matrices for each user at a BS.

This paper is organized as follows. In Section 2, the system model for a wideband multi-user MIMO-OFDM system is presented. In Section 3, an adaptive robust Max-SLNR precoder is proposed and designed, and Section 4 discusses a method for adaptively estimating the coefficients of a CSI error model. Numerical results and analysis are given in Section 5, and the final section concludes the paper.

Notations: throughout the paper, matrices, vectors, and scalars are denoted by bold upper case, bold lower case, and lower case lettering, respectively. The symbols $(\bullet)^*$ and $(\bullet)^{\mathrm{H}}$ denote matrix conjugate and conjugate transpose, respectively. $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix, and $\mathrm{E}(\bullet)$ denotes the expectation operator.

2 System model

The diagram block of a wide-band MU-MIMO-OFDM discrete baseband system is shown in Figure 1.

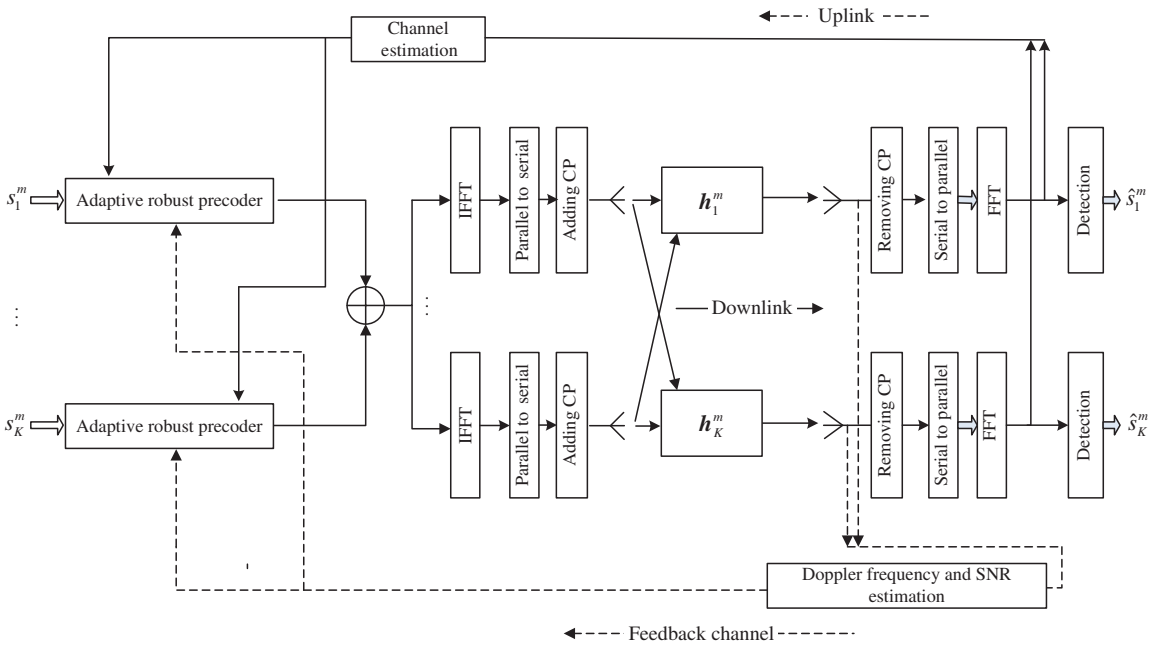

Figure 1 Diagram block of wide-band MU-MIMO-OFDM system.

In this system, the BS employs J transmit antennas, the number of users is K (with each user terminal equipped with one receive antenna), the total number of subcarriers is N, and the length of the cyclic prefix is L. The mth received symbol for user k at time instant $t\delta T$, where δT denotes the sampling period equal to $1/B$, is given as follows:

$$y_k^m(t) = \sum_{l=0}^{L}\sum_{j=1}^{J} h_{kj}^m(l)x_{kj}^m(t-l) + \sum_{k'=1,\ k'\neq k}^{K}\sum_{l=0}^{L}\sum_{j=1}^{J} h_{kj}^m(l)x_{k'j}^m(t-l) + v_k^m(t), \tag{1}$$

where $x_{kj}^m(t)$ denotes the sample taken at time t of the transmit signal of the mth symbol transmitted from antenna j to user k, $h_{kj}^m(t)$ is the corresponding time-domain channel impulse response (CIR), and $v_k^m(t)$ is the additive white Gaussian noise with zero mean and variance σ^2. The first and second terms on the right-hand side of (1) are the desired signal and the interference sum from other users, respectively.

In slow time-variant channels, inter-channel interference (ICI) can be omitted and the mth frequency-domain receive OFDM symbol of the n subcarrier of user k can be written as follows:

$$Y_k^{mn} = \boldsymbol{G}_k^{mn}\boldsymbol{X}_k^{mn} + N_k^{mn} = \boldsymbol{G}_k^{mn}\boldsymbol{w}_k^{mn}s_k^{mn} + \boldsymbol{G}_k^{mn}\sum_{k'=1,\ k'\neq k}^{K}\boldsymbol{w}_{k'}^{mn}s_{k'}^{mn} + V_k^{mn}, \tag{2}$$

where s_k^{mn} denotes the corresponding data symbol transmitted from the BS to user k, $\boldsymbol{w}_k^{mn}$ is the linear precoder mapping subcarrier n of the mth data symbol for user k to all J antennas, V_k^{mn} is independently identical distribution (i.i.d) Gaussian noise with zero mean and variance σ^2, and $\boldsymbol{G}_k^{mn}$ is the frequency-domain channel matrix corresponding to the subcarrier n of the mth symbol for user k. Combining matrices $\boldsymbol{G}_k^{mn}$ from all subcarriers forms a large frequency-domain matrix

$$\boldsymbol{G}_k^m = \left[\left(\boldsymbol{G}_k^{m1}\right)^{\mathrm{T}}, \left(\boldsymbol{G}_k^{m2}\right)^{\mathrm{T}}, \ldots, \left(\boldsymbol{G}_k^{mN}\right)^{\mathrm{T}}\right]^{\mathrm{T}} = \boldsymbol{F}\boldsymbol{h}_k^m\boldsymbol{F}_J^{\mathrm{H}}, \tag{3}$$

where $\boldsymbol{h}_k^m$ is the equivalent time-domain channel matrix corresponding to the mth OFDM symbol for user k, $\boldsymbol{F}$ is the normalized $N\times N$ FFT matrix, and $\boldsymbol{F}_J = \boldsymbol{F}\otimes\boldsymbol{I}_J$. s_k^{mn} satisfies the transmit constraint $\mathrm{E}[s_k^{mn}s_k^{mn*}] = P_T$. Based on the Gaussian-Markov error model in [11], the availability of imperfect CSI at the BS allows the estimated frequency-domain channel matrix to be modeled as [11]

$$\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} = \beta\boldsymbol{G}_k^{mn} + \sqrt{1-(\beta)^2}\boldsymbol{G}_{e,k}^{mn} \quad (0<\beta<1), \tag{4}$$

where β is a constant. In practical systems, the time variance in a wireless channel causes the statistics parameters (such as variance and the covariance matrix) of the channel estimation error to vary; therefore, we must revise the constant error model to the following time-variant model:

$$\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} = \beta_k^m \boldsymbol{G}_k^{mn} + \sqrt{1-(\beta_k^m)^2}\boldsymbol{G}_{e,k}^{mn} \quad (0 < \beta_k^m < 1), \tag{5}$$

where Δm is the delay (number of OFDM symbols) between channel estimation and beamforming, $\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}$ is the imperfect frequency-domain channel matrix obtained through channel estimation at the $(m-\Delta m)$th OFDM symbol, $\boldsymbol{G}_k^{mn}$ denotes the ideal channel matrix at the m OFDM symbol, and $\boldsymbol{G}_{e,k}^{mn}$ represents the errors produced from channel estimation error and channel time-variance, which is assumed to have independent complex Gaussian elements with variance σ_e^2. In (5), β_k^m is the error model coefficient of the mth symbol for user k and the coefficient $\sqrt{1-(\beta_k^m)^2}$ measures the accuracy of CSI imperfection. Unlike in [11], the parameter β_k^m in (5) is time-variant and user-variant. In the case where $\beta_k^m = 1$, the transmitter can obtain a perfect CSI. In real fading wireless channels, an estimation error always exists and the corresponding error model coefficient β_k^m varies with the channel environment owing to channel-specific time and frequency selectivity as well as spatial selectivity. Thus, it is very useful to have an adaptive robust precoder that can follow channel variance. To design an adaptive beamformer, the values of β_k^m and $\boldsymbol{G}_k^{(m-\Delta m)n}$ are first computed. Then, given these two values, $\boldsymbol{G}_k^{mn}$ can be anticipated as the conditional expected value $\mathrm{E}\{\boldsymbol{G}_k^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\}$, where E is the expected value operation. Finally, using the estimated $\boldsymbol{G}_k^{mn}$, the beamforming vector $\boldsymbol{w}_k^{mn}$ is designed.

In order to simplify the estimation and derivation of β_k^m below, we assume that the pilot patterns for the various users are designed to be orthogonal, thus eliminating interference among users in the training phrase. Prior to beamforming at the BS, all users send their respective pilot symbols to the BS simultaneously. Because the JKL parameters must be estimated, in a wideband multi-pair two-way network, each user needs at least L pilot grids in order to complete one-time channel estimation. In the case where there are L transmit time-frequency pilot grids, the received pilot frequency-domain symbol over subchannel n and antenna j of the BS can be equivalently written as

$$\widetilde{Y}_{kj}^{mn} = H_{kj}^{mn}\widetilde{s}_{kj}^{mn} + \widetilde{V}_{kj}^{mn}. \tag{6}$$

We can readily obtain the value of G_{kj}^{mn} by making use of the channel reciprocity between the uplink and the downlink. If the number of transmit time-frequency pilot grids exceeds L, then an SNR gain factor α_k can be introduced to represent the increased number of pilot grids, as will be later discussed.

3 Proposed Max-SLNR-based adaptive robust beamformer

Given the estimated channel matrix $\boldsymbol{G}_k^{(m-\Delta m)n}$ and β_k^m, the concept of leakage in [7,22] and the revised Gaussian-Markov error model in (5) can be used to define the SLNR for subcarrier n of the mth symbol for user k with an imperfect channel matrix as

$$\mathrm{SLNR}_k^{mn}(\boldsymbol{w}_k^{mn}) = \frac{\mathrm{E}\left[(s_k^{mn})^{\mathrm{H}}(\boldsymbol{w}_k^{mn})^{\mathrm{H}}(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn}\boldsymbol{w}_k^{mn}s_k^{mn} \mid (\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}{\sigma^2 + \mathrm{E}\left[(s_k^{mn})^{\mathrm{H}}(\boldsymbol{w}_k^{mn})^{\mathrm{H}}\left(\boldsymbol{G}_{(-k)}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{(-k)}^{mn}\boldsymbol{w}_k^{mn}s_k^{mn} \mid (\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}, \tag{7}$$

where $\boldsymbol{G}_{(-k)}^{mn}$ is the extended channel matrix

$$\boldsymbol{G}_{(-k)}^{mn} = \left[(\boldsymbol{G}_1^{mn})^{\mathrm{T}}\cdots\left(\boldsymbol{G}_{k-1}^{mn}\right)^{\mathrm{T}}\left(\boldsymbol{G}_{k+1}^{mn}\right)^{\mathrm{T}}\cdots(\boldsymbol{G}_K^{mn})^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{8}$$

excluding the matrix vector $\boldsymbol{G}_k^{mn}$.

Using the generalized Rayleigh-Ritz theorem in [7], the precoder $\boldsymbol{w}_k^{mn,o}$ for maximizing the SLNR can be obtained from the eigenvector corresponding to the largest eigen-value of the matrix $\boldsymbol{A}_k^{mn}$

$$\boldsymbol{A}_k^{mn} = \frac{\mathrm{E}\left[(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn} \mid (\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}{\frac{\sigma^2}{P_T}\boldsymbol{I}_J + \mathrm{E}\left[\left(\boldsymbol{G}_{(-k)}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{(-k)}^{mn} \mid (\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}. \tag{9}$$

In accordance with the derivation in Appendix A, $\boldsymbol{A}_k^{mn}$ above can be further simplified as

$$\boldsymbol{A}_k^{mn} = \frac{\left(\frac{1}{\beta_k^m} - \frac{\sigma_e^2\sqrt{1-(\beta_k^m)^2}}{\left(1-(\beta_k^m)^2\right)\beta_k^m\sigma_e^2+(\beta_k^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} + \frac{J\sigma_e^2\left(1-(\beta_k^m)^2\right)}{\left(1-(\beta_k^m)^2\right)\sigma_e^2+(\beta_k^m)^2}\boldsymbol{I}_J}{\frac{\sigma^2}{P_T}\boldsymbol{I}_J + \sum\limits_{k'=1,k'\neq k}^{K}\left(\left(\frac{1}{\beta_{k'}^m} - \frac{\sigma_e^2\sqrt{1-(\beta_{k'}^m)^2}}{\left(1-(\beta_{k'}^m)^2\right)\beta_{k'}^m\sigma_e^2+(\beta_{k'}^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n} + \frac{J\sigma_e^2\left(1-(\beta_{k'}^m)^2\right)}{\left(1-(\beta_{k'}^m)^2\right)\sigma_e^2+\beta_{k'}^2}\boldsymbol{I}_J\right)}. \tag{10}$$

If we ignore the error in channel estimation, i.e., if we assume that $\beta = 1$ and $\sigma_e^2 = 0$, the solution (7) reduces to the solution with ideal CSI given in [7].

4 Real-time estimation of Markov-Gaussian error model coefficients

For convenience in further derivation, in this section the error model coefficient β_k^m is decomposed into product of two factors:

$$\beta_k^m = \beta_{k,d}^m\beta_{k,e}^m, \tag{11}$$

These factors are related to time variance in the channel and to the channel estimation error, respectively; the parameter $\beta_{k,d}^m$ measures the effect of channel mismatching caused by Doppler spread, whereas $\beta_{k,e}^m$ measures the effect of channel mismatching owing to estimation error.

In order to successfully estimate β_k^m in real-time, we assume that the BS broadcasts L_p continuous pilot OFDM symbols, each having the same length and transmit power as the data OFDM symbols. After the pilot OFDM symbol passes through a channel, the sampling point t of the pilot time-domain OFDM symbol of user k received from the jth antenna of the BS is expressed as $\overline{y}_{kj}^m(t)$. Each user estimates its own SNR and Doppler spread by using the corresponding received OFDM symbol and feeds back these estimated values to the BS. Based on these values, the BS computes β_k^m and updates the beamforming matrix at BS.

4.1 Error parameter associated with Doppler spread

In the following subsection, we will show how to compute the parameter $\beta_{k,d}^m$. To simplify this derivation, we assume that $\beta_{k,e}^m = 1$. $\beta_{k,d}^m$ can be obtained using the MMSE criterion. The complex channel gain from subcarrier n of user k to the jth antenna of the BS at the $(m-\Delta m)$th OFDM symbol is modeled as

$$\widehat{G}_{kj}^{(m-\Delta m)n} = \beta_{k,d}^m G_{kj}^{mn} + \sqrt{1-(\beta_{k,d}^m)^2}G_{e,kj}^{mn}, \tag{12}$$

where Δm is the corresponding delayed number of OFDM symbols, $\Delta m T_s$ is the time delay between channel estimation and beamforming, and T_s is the total length of the OFDM symbol. The first and second terms on the right-hand side of the above equation denote the ideal channel gain at the mth OFDM symbol and the error caused by the Δm-symbol delay, respectively.

The MSE between the received and estimated transmitted signals can be given as follows

$$\mathrm{E}\left(\|\,\widetilde{Y}_{kj}^{mn} - \widehat{G}_{kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right) = \mathrm{E}\left(\|\,G_{kj}^{mn}\widetilde{s}_{kj}^{mn} + \widetilde{V}_{kj}^{mn} - \beta_{k,d}^m G_{kj}^{(m+\Delta m)n}\widetilde{s}_{kj}^{mn} - \sqrt{1-(\beta_{k,d}^m)^2}G_{e,kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right). \tag{13}$$

As proven in Appendix B, minimizing the MSE by differentiating the above expression with respect to $\beta_{k,d}^m$ yields

$$\beta_{k,d}^m = \frac{\Re_{GG,k}\left(\Delta m, 0\right)}{\Re_{GG,k}\left(0,0\right)}, \tag{14}$$

where $\Re_{GG,k}(\Delta m, \Delta n) = \mathrm{E}(G_k^{mn}(G_k^{(m+\Delta m)(n+\Delta n)})^*)$ denotes the time-frequency channel correlation function. Note that, in this derivation, $\Delta n = 0$. For a time-varying Rayleigh fading channel, $\Re_{GG,k}(\Delta m, 0)$ is given as

$$\Re_{GG,k}(\Delta m, 0) = J_0\left(2\pi f_{d,k}^m \Delta m T_s\right), \tag{15}$$

where function $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind:

$$J_0(x) = \frac{1}{x}\int_0^{\pi} \mathrm{e}^{\mathrm{i}x\cos\theta}\mathrm{d}\theta, \tag{16}$$

and $f_{d,k}^m$ is the Doppler spread for user k. In terms of (14), we may compute the $\beta_{k,d}^m$ given $f_{d,k}^m$. By exploiting the cyclic property of the CP of the OFDM symbol, the Doppler spread $f_{d,k}^m$ can be estimated as in [23–27]:

$$\widehat{f}_{d,k}^m = \frac{1}{2\pi T_u}\sqrt{\frac{4\sum_{j=1}^{J}\sum_{t=0}^{L-1}\left(|\overline{y}_{kj}^m(t)|^2 - |\overline{y}_{kj}^m(t+N)|^2\right)}{\sum_{j=1}^{J}\sum_{t=0}^{L-1}\left(|\overline{y}_{kj}^m(t)|^2 + |\overline{y}_{kj}^m(t+N)|^2\right)}}, \tag{17}$$

where T_u is the useful length of OFDM symbol.

4.2 Error parameter associated with channel estimation error

We now proceed to estimate the parameter $\beta_{k,e}^m$ relative to channel estimation. The accuracy of channel estimation relies directly on the average SNR $\gamma_{p,k}^m$ over the pilot grids in a channel given an estimator such as MMSE, which is approximately proportional to the average SNR $\gamma_{d,k}^m$ over the data grids. As shown in Appendix C, the parameter $\beta_{k,e}^m$ obtained by adopting the LMMSE channel estimator is

$$\beta_{k,e}^m = \frac{\gamma_{p,k}^m}{1+\gamma_{p,k}^m}. \tag{18}$$

Based on the transmission power of the pilot symbols and the number of pilot symbols per channel parameter, we can approximate $\gamma_{p,k}^m$ as a linear function of $\gamma_{d,k}^m$ as follows

$$\gamma_{p,k}^m \approx \alpha_k^m \gamma_{d,k}^m, \tag{19}$$

where α is defined as

$$\alpha_k^m = \frac{P_{p,k,m}}{P_{d,k,m}}\frac{N_{p,k,m}}{N_{c,k,m}}, \tag{20}$$

where $P_{p,k,m}$ and $N_{p,k,m}$ are the power and length, respectively, of the practice training symbols for user k, $P_{d,k,m}$ is the average power of the data symbols for user k, and $N_{c,k}$ is the number of channel parameters to be estimated for user k. The cyclic property of cyclic prefix (CP) in the OFDM system means that $\gamma_{d,k}^m$ can be estimated by the following expression:

$$\hat{\gamma}_{d,k}^m = \frac{\frac{1}{L}\sum_{j=1}^{J}\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\left(\overline{y}_{kj}^m(t+N)\right)^*\right|}{\frac{1}{2L}\sum_{j=1}^{J}\left(\sum_{t=0}^{L-1}\left(\left|\overline{y}_{kj}^m(t)\right| - \left|\overline{y}_{kj}^m(t+N)\right|\right)^2\right)}, \tag{21}$$

which is given in Appendix D. Based on this, we can complete the calculation of $\beta_{k,e}^m$.

The error model coefficient β_k^m can be obtained by substituting $\beta_{k,e}^m$ and $\beta_{k,d}^m$ in (11). Substituting the error model coefficient into the adaptive robust precoder in Section 3, the precoder can be adjusted adaptively through real-time estimation of the Doppler spread and SNR at the data grid, which can improve the robustness of the system. The results will be further verified by simulation in the next section.

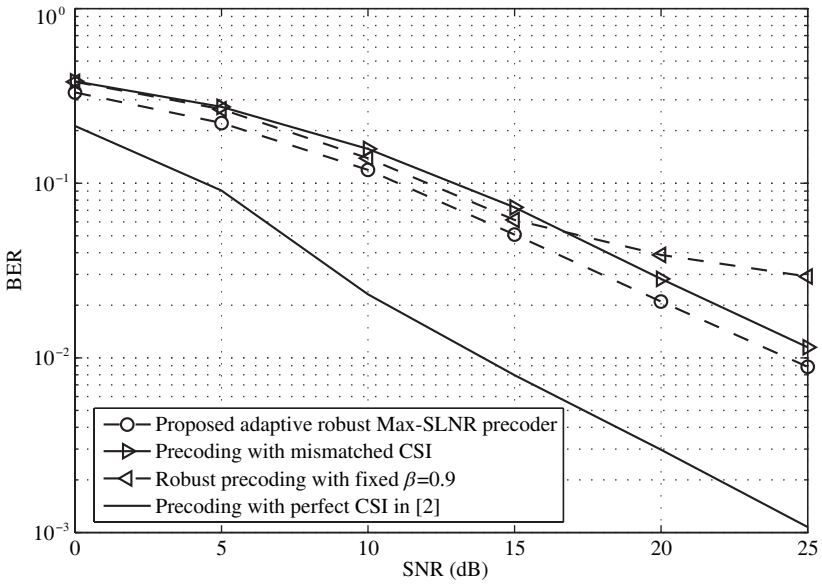

Figure 2 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 0.5$, and $f_d \Delta m T_s = 0.01$).

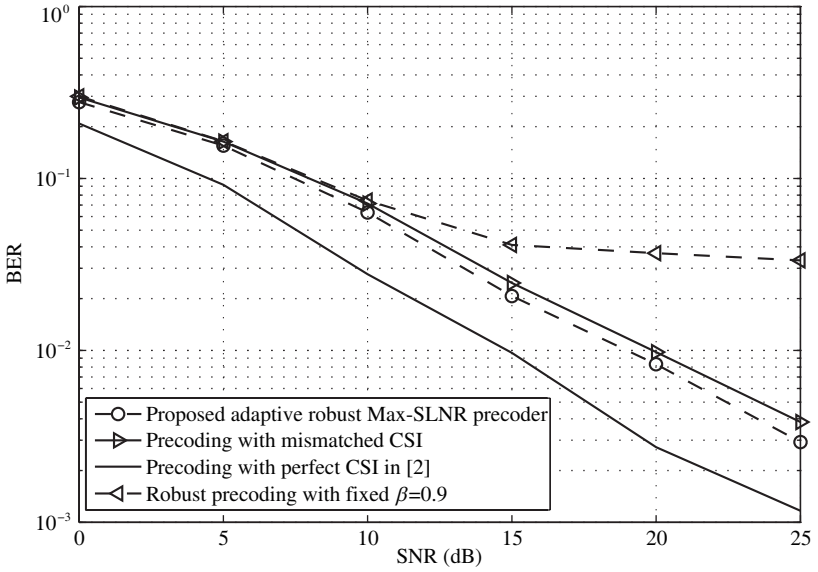

Figure 3 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 1$, and $f_d \Delta m T_s = 0.01$).

5 Simulations and discussion

The following baseband system parameters and specifications are used in the following simulation: digital modulation QPSK; channel bandwidth $BW = 1$ MHz, $N = 64$, CP length = 8, four single-antenna users ($K = 4$); one BS with four antennas ($J = 4$); and carrier frequency $f_c = 900$ MHz. The normalized Doppler spread is defined as $f_d T_u$, where f_d and T_u denote the Doppler spread and the length of useful OFDM symbol, respectively. $f_d T_u$ equals 0.01, indicating that our simulation system is a slow time-variant fading MIMO-OFDM system. At the same time, we define $\Delta m f_d T_s$, where T_s is the total length of the OFDM symbol including CP, as the delayed time difference between channel estimation and beamforming.

Figures 2, 3, and 4 show the curves of BER versus SNR of the proposed adaptive robust precoder at $\alpha = 0.5$, 1, and 4, respectively. For comparison, the BER curves of a robust precoder with fixed β, a precoder obtained by mismatched channel matrix, and a precoder with perfect CSI are presented in the three figures. Under all SNRs, the proposed algorithm performs better than the algorithm with fixed β.

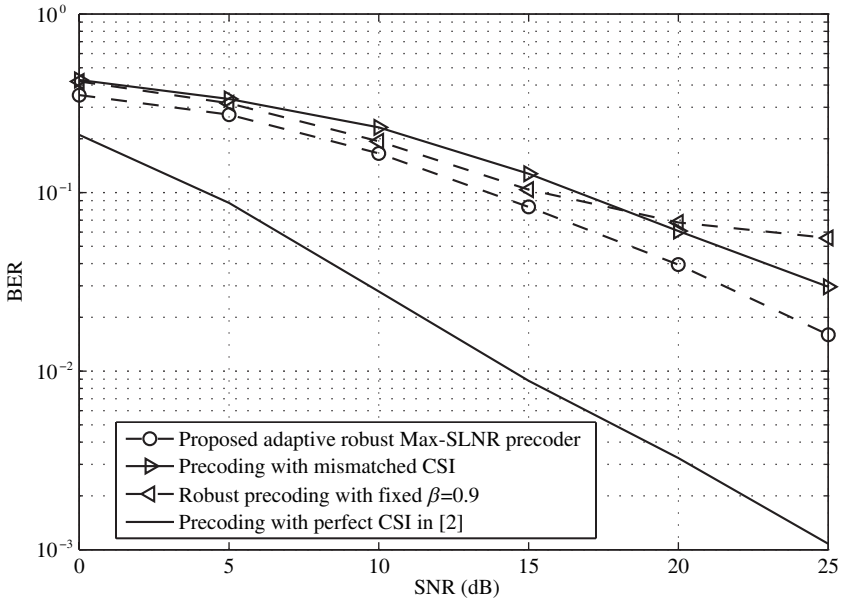

Figure 4 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 4$, and $f_d \Delta m T_s = 0.01$).

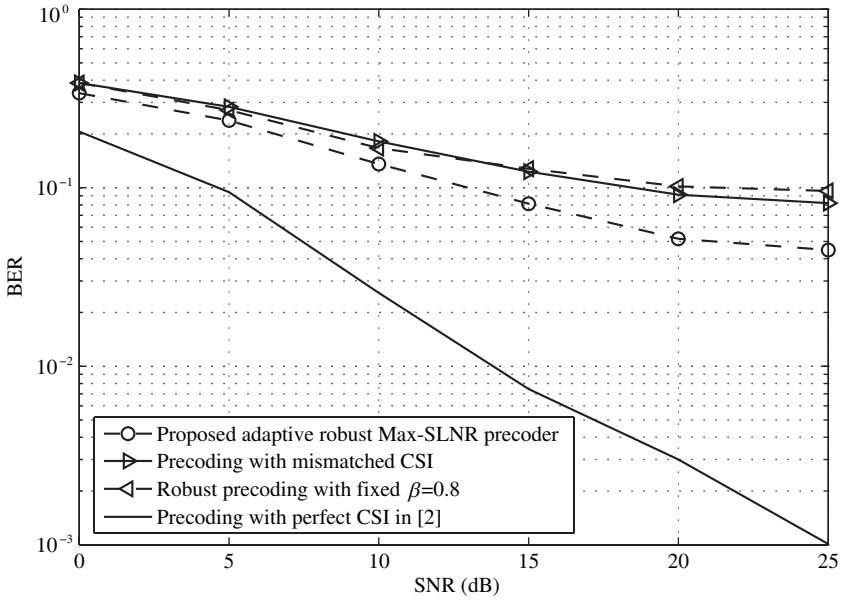

Figure 5 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 0.5$, and $f_d \Delta m T_s = 0.05$).

Given a fixed $\Delta m f_d T_s = 0.01$, the BER performance of the proposed beamformer tends toward the ideal BER performance as α increases from 0.5 to 4.

Figures 5, 6, and 7 show the curves of BER versus SNR of the proposed adaptive robust beamformer for three different values of α ($\alpha = 0.5$, 1, and 4). From these figures, it is seen that, given a fixed α, the SNR gains of the proposed method over other algorithms increase as $\Delta m f_d T_s$ increases from 0.01 to 0.05.

Figure 8 plots the curves of BER versus α of the proposed adaptive robust precoder based on Max-SLNR at three fixed SNRs (5, 10, and 15 dB). From Figure 8, as α increases, the BER performance converges to the performance of the Max-SLNR with ideal CSI as a lower bound. This effect is primarily the result of the fact that a large α corresponds to more training symbols per channel parameter/higher transmit pilot power; in other words, a large α provides higher channel estimation accuracy. Additionally, the convergence rate of the BER performance at SNR = 15 dB is faster than those at SNR = 10 and

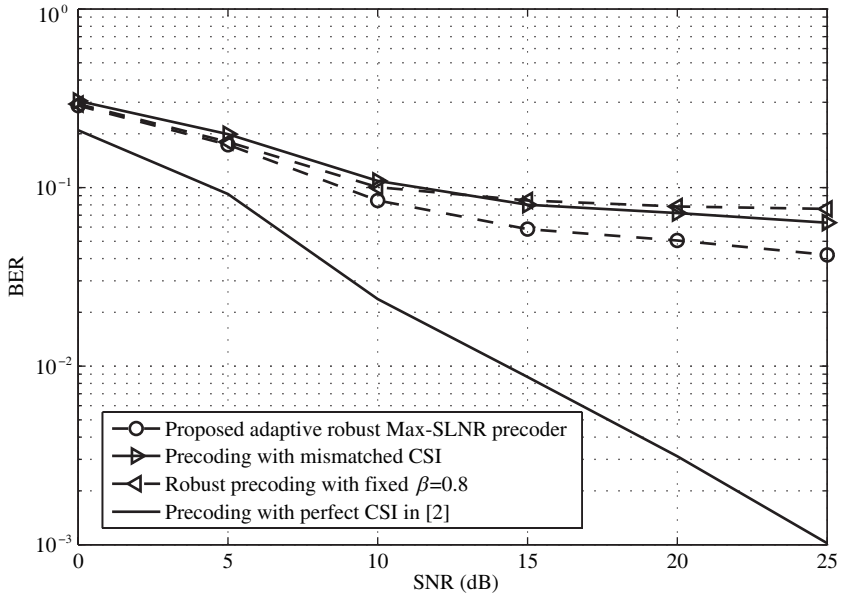

Figure 6 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 1$, and $f_d \Delta m T_s = 0.05$).

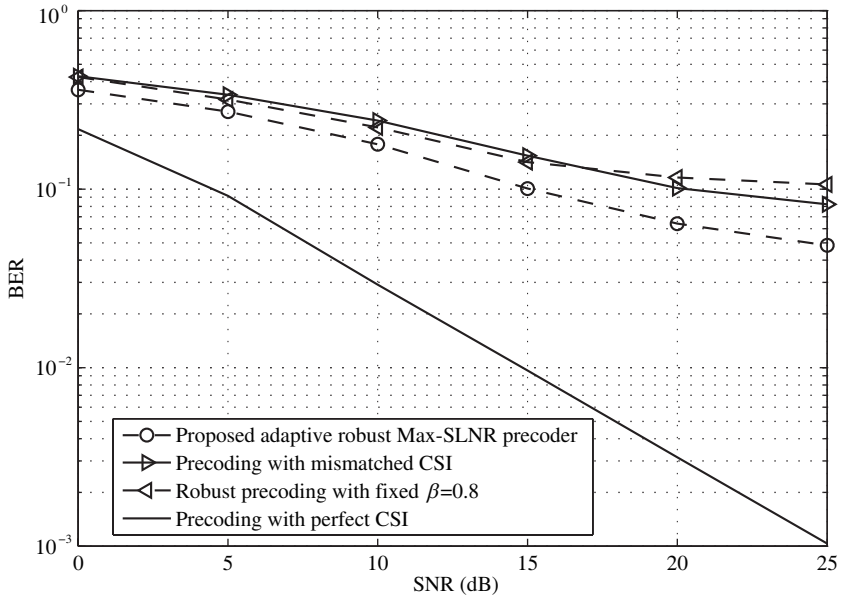

Figure 7 Curves of BER versus SNR of the proposed adaptive robust beamformer ($\alpha = 4$, and $f_d \Delta m T_s = 0.05$).

5 dB.

Figure 9 plots the curves of sum-rate versus α of the proposed adaptive robust precoder based on Max-SLNR at three fixed SNRs (5, 10, and 15 dB). From Figure 9, as α increases, the sum-rate performance approaches the performance of the Max-SLNR with ideal CSI as an upper bound, for the same reason as discussed in conjunction with Figure 8.

6 Conclusion

In this paper, an adaptive robust Max-SLNR precoder that can take into account some of the Doppler spread and channel estimation-induced uncertainty or variation in the channel matrix was proposed. Simulation results show that the proposed adaptive robust precoder performs better than existing robust

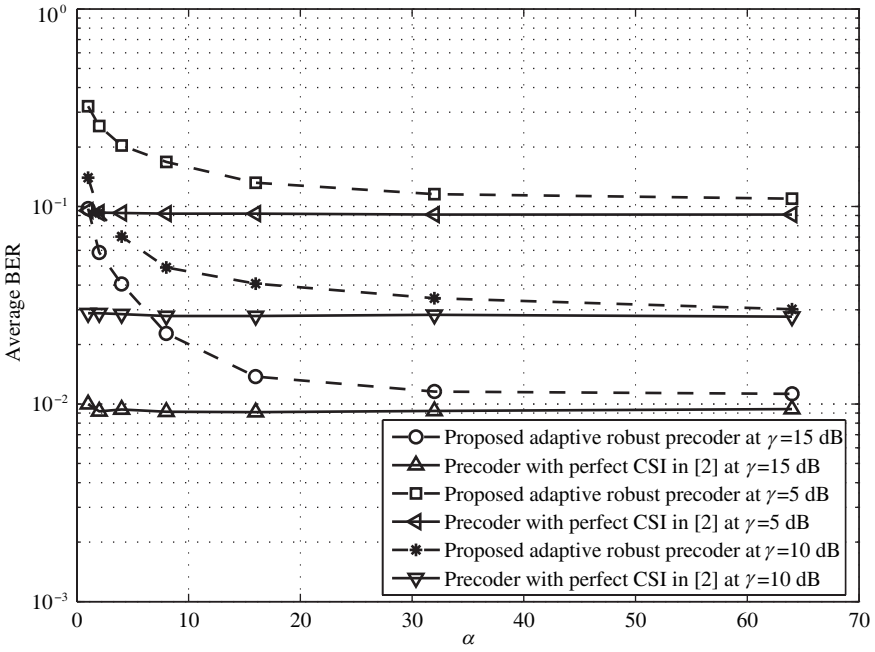

Figure 8 Curves of BER versus α of the proposed adaptive robust beamformer at γ = 5 dB, 15 dB, and 25 dB.

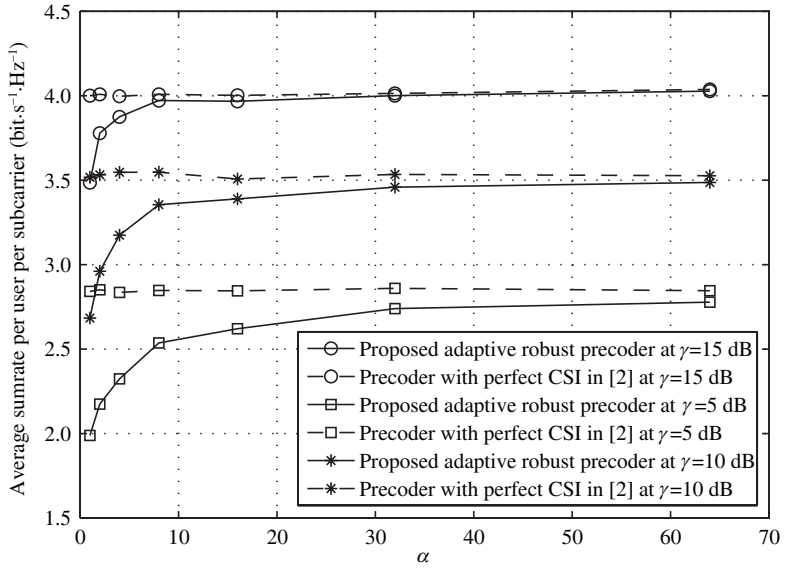

Figure 9 Curves of sum-rate versus α of the proposed adaptive robust beamformer at γ = 5 dB, 15 dB, and 25 dB.

and non-robust algorithms. As either the number of training symbols or the transmit power becomes large, the BER and sum-rate performance of the proposed adaptive method approach those of a precoder with perfect CSI.

Acknowledgements This work was supported in part by the open research fund of National Mobile Communications Research Laboratory, Southeast University (Grant No. 2013D02), the Fundamental Research Funds for the Central Universities (Grant No. 30920130122004), and the National Natural Science Foundation of China (Grant Nos. 61271230, 61472190).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Gesbert D, Hanly S, Huang H, et al. Multi-cell MIMO cooperative networks: a new look at interference. IEEE Sel Area Commun, 2010, 28: 1380–1408
2 Chen R H, Geirhofer S, Sayana K, et al. Downlink MIMO in LTE-advanced: SU-MIMO vs MU-MIMO. IEEE Commun Mag, 2012, 50: 140–147
3 Karaa H, Adve R S, Tenenbaum A J. Linear precoding for multiuser MIMO-OFDM systems. In: IEEE International Conference on Communications (ICC), Glasgow, 2007. 2797–2802
4 Wang J, Xie X, Zhang Q. A way to reduce ICI for multi-user MIMO-OFDM system with precoding. In: IEEE International Conference on Advanced Computer Control, Singapore, 2009. 134–137
5 Shin Y, Kang T, Kim H. An efficient resource allocation for multiuser MIMO-OFDM systems with zero-forcing beamformer. In: IEEE International Symposium on Personal, Indoor and Mobile Radio Commun, Athens, 2007. 1–5
6 Castanheira D, Silva A, Gameiro A. Linear and nonlinear precoding schemes for centralized multicell MIMO-OFDM systems. IEEE Wirel Pers Commun, 2013, 72: 759–777
7 Sadek M, Aissa S. Leakage based precoding for multi-user MIMO-OFDM systems. IEEE Trans Wirel Commun, 2011, 10: 2428–2433
8 Sung H, Lee S R, Lee I. Generalized channel inversion methods for multiuser MIMO systems. IEEE Trans Commun, 2009, 57: 3489–3499
9 Liu W, Yang L L, Hanzo L. SVD-assisted multiuser transmitter and multiuser detector design for MIMO systems. IEEE Trans Veh Technol, 2009, 58: 1016–1021
10 Mao J L, Gao J C, Liu Y A, et al. Robust multiuser MIMO scheduling algorithms with imperfect CSI. Sci China Inf Sci, 2012, 55: 815–826
11 Zhou B, Jiang L, Zhang L, et al. Impact of imperfect channel state information on TDD downlink multiuser MIMO system. In: Proceedings of IEEE Wirel Commun and Networking Conference (WCNC), Cancun, 2011. 1823–1828
12 Wang C, Au E K S, Murch R D, et al. On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error. IEEE Trans Wirel Commun, 2007, 6: 805–810
13 Sadek M, Tarighat A, Sayed A H. A leakage-based precoding scheme for downlink multi-user MIMO channels. IEEE Trans Wirel Commun, 2007, 6: 1711–1721
14 Chung P J, Du H. Robust SLNR downlink beamforming based on Markov's inequality. In: IEEE International Conference on Communications (ICC), Ottawa, 2012. 3627–3631
15 Chung P J, Du H, Chen M. Robust transmit beamforming for multi-user MIMO systems using a probabilistic constraint approach. In: Signal Processing of 11th International Conference. Beijing: IEEE, 2012, 2: 1482–1485
16 Shen H, Xu W, Jin S, et al. Joint transmit and receive beamforming for multiuser MIMO downlinks with channel uncertainty. IEEE Trans Veh Technol, 2014, 63: 2319–2335
17 Merline A, Thiruvengadam S J. Design of optimal linear precoder and decoder for MIMO channels with per antenna power constraint and imperfect CSI. IEEE Wirel Pers Commun, 2014, 75: 1251–1263
18 Khalid F, Speidel J. Robust hybrid precoding for multiuser MIMO wireless communication systems. IEEE Trans Wirel Commun, 2014, 13: 1–11
19 Tseng F, Wang Y, Hsu C, et al. Robust tomlinson-harashima precoder design with random vector quantization in MIMO systems. IEEE Commun Lett, 2014, 18: 265–268
20 Sun Y, Wu M, Guo Q, et al. Robust non-linear precoder for multiuser MISO systems based on delay and channel quantization. Wirel Pers Commun, 2013, 72: 1993–2014
21 Pan C, Su C, Chen M. Robust precoding for multicell MIMO downlink systems. In: IEEE International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, 2013. 1–6
22 Shu F, Lu Y Z, Chen Y. High-sum-rate beamformers for multi-pair two-way relay networks with amplify-and-forward relaying strategy. Sci China Inf Sci, 2014, 57: 022312
23 Wang J, Wen O Y, Li S. Soft-output MMSE MIMO detector under imperfect channel estimation. In: IEEE Wirel Commun and Networking Conference, Las Vegas, 2008. 1334–1338
24 Cai J P, Song W T, Li Z. Doppler spread estimation for mobile OFDM systems in rayleigh fading channels. IEEE Trans Consum Electron, 2003, 49: 973–977
25 Tepedelenlioglu C, Giannakis G. On velocity estimation and correlation properties of narrow-band mobile communication channels. IEEE Trans Veh Technol, 2001, 50: 1039–1052
26 Ijaz A, Awoseyila A B, Evans B G. Low-complexity time-domain SNR estimation for OFDM systems. IEEE Electron Lett, 2011, 47: 1154–1156
27 Li Y, Xu X, Zhang D D, et al. Optimal pilots design for frequency offsets and channel estimation in OFDM modulated single frequency networks. Sci China Inf Sci, 2014, 57: 042301

Appendix A Derivation of the proposed robust Max-SLNR precoder

Below, we will show how to compute $\mathrm{E}\left[(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]$, which can be used to simplify (9).

Conditioned on β_k^m and $\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}$, we have

$$\mathrm{E}\left[(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left[\left(\frac{1}{\beta_k^m}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} - \frac{\sqrt{1-(\beta_k^m)^2}}{\beta_k^m}\boldsymbol{G}_{e,k}^{mn}\right)^{\mathrm{H}}\left(\frac{1}{\beta_k^m}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} - \frac{\sqrt{1-(\beta_k^m)^2}}{\beta_k^m}\boldsymbol{G}_{e,k}^{mn}\right)\Big|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]\\
&= \frac{1}{(\beta_k^m)^2}\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} + \left(\frac{1-(\beta_k^m)^2}{(\beta_k^m)^2}\right)\mathrm{E}\left(\left(\boldsymbol{G}_{e,k}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{e,k}^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right)\\
&\quad - \frac{\sqrt{1-(\beta_k^m)^2}}{(\beta_k^m)^2}\mathrm{E}\left(\left(\boldsymbol{G}_{e,k}^{mn}\right)^{\mathrm{H}}\Big|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right)\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\\
&\quad - \frac{\sqrt{1-(\beta_k^m)^2}}{(\beta_k^m)^2}\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\mathrm{E}\left(\boldsymbol{G}_{e,k}^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right). \qquad \text{(A1)}
\end{aligned}
$$

Furthermore, from [23] we have

$$
\begin{aligned}
\mathrm{E}\left(\left(\boldsymbol{G}_{e,k}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{e,k}^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right) &= \mathrm{E}\left(\sum_{m=1}^{J}\left(G_{e,kj}^{mn}\right)^{\mathrm{H}}G_{e,kj}^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right)\\
&= \sum_{m=1}^{J}\mathrm{E}\left(\left(G_{e,kj}^{mn}\right)^{\mathrm{H}}G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right), \qquad \text{(A2)}
\end{aligned}
$$

where $G_{e,kj}^{mn}$ and $\widehat{G}_{kj}^{(m-\Delta m)n}$ denote the jth column of matrix $\boldsymbol{G}_{e,k}^{mn}$ and $\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}$, respectively.

$$
\begin{aligned}
\mathrm{E}\left(\left(G_{e,kj}^{mn}\right)^{\mathrm{H}}G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right) &= \mathrm{cov}\left[\left(G_{e,kj}^{mn}\right)^{\mathrm{H}}G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right]\\
&\quad + \mathrm{E}\left[\left(G_{e,kj}^{mn}\right)^{\mathrm{H}}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right]\mathrm{E}\left[G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right], \qquad \text{(A3)}
\end{aligned}
$$

where the notation $\mathrm{cov}(\cdot)$ denotes the covariance operator. The element of $\boldsymbol{G}_k^{mn}$ has independent real and imaginary parts with zero mean and unit variance; then we have

$$
\begin{aligned}
\mathrm{E}\left[G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right] &= \mathrm{E}\left[G_{e,kj}^{mn}\left(G_{e,kj}^{mn}\right)^*\right]\mathrm{E}\left[\widehat{G}_{kj}^{(m-\Delta m)n}\left(\widehat{G}_{kj}^{(m-\Delta m)n}\right)^*\right]^{-1}\widehat{G}_{kj}^{(m-\Delta m)n}\\
&= \sigma_e^2\frac{1}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\widehat{G}_{kj}^{(m-\Delta m)n}\\
&= \frac{\sigma_e^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\widehat{G}_{kj}^{(m-\Delta m)n}. \qquad \text{(A4)}
\end{aligned}
$$

The covariance matrix $\mathrm{cov}\left[\left(G_{e,kj}^{mn}\right)^* G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right]$ can be given as [23]

$$
\begin{aligned}
&\mathrm{cov}\left[\left(G_{e,kj}^{mn}\right)^* G_{e,kj}^{mn}|(\beta_k^m, \widehat{G}_{kj}^{(m-\Delta m)n})\right]\\
&= \mathrm{E}\left[\left(G_{e,kj}^{mn}\right)^* G_{e,kj}^{mn}\right] - \mathrm{E}\left[\left(\widehat{G}_{kj}^{(m-\Delta m)n}\right)^* G_{e,kj}^{mn}\right]\mathrm{E}\left[\left(\widehat{G}_{kj}^{(m-\Delta m)n}\right)^*\widehat{G}_{kj}^{(m-\Delta m)n}\right]^{-1}\mathrm{E}\left[\left(G_{e,kj}^{mn}\right)^*\widehat{G}_{kj}^{(m-\Delta m)n}\right]\\
&= \sigma_e^2 - \sqrt{1-(\beta_k^m)^2}\sigma_e^2\frac{1}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\sqrt{1-(\beta_k^m)^2}\sigma_e^2\\
&= \frac{\sigma_e^2(\beta_k^m)^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}. \qquad \text{(A5)}
\end{aligned}
$$

By substituting (A3), (A4), and (A5) into (A2), it yields

$$
\begin{aligned}
&\mathrm{E}\left(\left(\boldsymbol{G}_{e,k}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{e,k}^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right)\\
&= \frac{J\sigma_e^2(\beta_k^m)^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\boldsymbol{I}_J + \left(\frac{\sigma_e^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\right)^2\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}. \qquad \text{(A6)}
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathrm{E}\left[(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn}|(\beta_k^m, \widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]\\
&= \frac{1}{(\beta_k^m)^2}\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} - 2\frac{\sqrt{1-(\beta_k^m)^2}}{(\beta_k^m)^2}\frac{\sigma_e^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\\
&\quad + \frac{1-(\beta_k^m)^2}{\beta_k^2}\left(\frac{J(\beta_k^m)^2\sigma_e^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\boldsymbol{I}_J + \left(\frac{\sigma_e^2}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\right)^2\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)\\
&= \left(\frac{1}{\beta_k^m} - \frac{\sigma_e^2\sqrt{1-(\beta_k^m)^2}}{\left(1-(\beta_k^m)^2\right)\beta_k^m\sigma_e^2 + (\beta_k^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n} + \frac{J\sigma_e^2\left(1-(\beta_k^m)^2\right)}{\left(1-(\beta_k^m)^2\right)\sigma_e^2 + (\beta_k^m)^2}\boldsymbol{I}_J. \qquad \text{(A7)}
\end{aligned}
$$

Similarly, we can obtain

$$\begin{aligned}
&\mathrm{E}\left[\left(\boldsymbol{G}_{(-k)}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{(-k)}^{mn}|(\beta_k^m,\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]\\
&=\sum_{k'=1,k'\neq k}^{K}\mathrm{E}\left[(\boldsymbol{G}_{k'}^{mn})^{\mathrm{H}}(\boldsymbol{G}_{k'}^{mn})|(\beta_{k'}^m,\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n})\right]\\
&=\sum_{k'=1,k'\neq k}^{K}\left[\left(\frac{1}{\beta_{k'}^m}-\frac{\sigma_e^2\sqrt{1-(\beta_{k'}^m)^2}}{\left(1-(\beta_{k'}^m)^2\right)\beta_{k'}^m\sigma_e^2+(\beta_{k'}^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n}+\frac{J\sigma_e^2\left(1-(\beta_{k'}^m)^2\right)}{\left(1-(\beta_{k'}^m)^2\right)\sigma_e^2+(\beta_{k'}^m)^2}\boldsymbol{I}_J\right].
\end{aligned}\tag{A8}$$

Therefore the matrix $\boldsymbol{A}_k^{mn}$ can be given as follows

$$\begin{aligned}
\boldsymbol{A}_k^{mn}&=\frac{\mathrm{E}\left[(\boldsymbol{G}_k^{mn})^{\mathrm{H}}\boldsymbol{G}_k^{mn}\mid(\beta_k^m,\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}{\frac{\sigma^2}{P_T}\boldsymbol{I}_J+\mathrm{E}\left[\left(\boldsymbol{G}_{(-k)}^{mn}\right)^{\mathrm{H}}\boldsymbol{G}_{(-k)}^{mn}\mid(\beta_k^m,\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n})\right]}\\
&=\frac{\left(\frac{1}{\beta_k^m}-\frac{\sigma_e^2\sqrt{1-(\beta_k^m)^2}}{(1-(\beta_k^m)^2)\beta_k^m\sigma_e^2+(\beta_k^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_k^{(m-\Delta m)n}+\frac{J\sigma_e^2(1-(\beta_k^m)^2)}{(1-(\beta_k^m)^2)\sigma_e^2+(\beta_k^m)^2}\boldsymbol{I}_J}{\frac{\sigma^2}{P_T}\boldsymbol{I}_J+\sum\limits_{k'=1,k'\neq k}^{K}\left(\left(\frac{1}{\beta_{k'}^m}-\frac{\sigma_e^2\sqrt{1-(\beta_{k'}^m)^2}}{\left(1-(\beta_{k'}^m)^2\right)\beta_{k'}^m\sigma_e^2+(\beta_{k'}^m)^3}\right)^2\left(\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n}\right)^{\mathrm{H}}\widehat{\boldsymbol{G}}_{k'}^{(m-\Delta m)n}+\frac{J\sigma_e^2\left(1-(\beta_{k'}^m)^2\right)}{\left(1-(\beta_{k'}^m)^2\right)\sigma_e^2+\beta_{k'}^2}\boldsymbol{I}_J\right)}.
\end{aligned}\tag{A9}$$

This completes the simplification of matrix $\boldsymbol{A}_k^{mn}$.

Appendix B Derivation of parameter $\boldsymbol{\beta_{k,d}^m}$

In terms of the MSE principle, substituting (12) in (13) yields

$$\begin{aligned}
\mathrm{E}\left(\|\,\widetilde{Y}_{kj}^{mn}-\widehat{G}_{kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right)&=\mathrm{E}\left(\|\,G_{kj}^{mn}\widetilde{s}_{kj}^{mn}+\widetilde{V}_{kj}^{mn}-\beta_{k,d}^m G_{kj}^{(m+\Delta m)n}\widetilde{s}_{kj}^{mn}-\sqrt{1-(\beta_{k,d}^m)^2}G_{e,kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right)\\
&=\mathrm{E}\left(\left[G_{kj}^{mn}\widetilde{s}_{kj}^{mn}+\widetilde{V}_{kj}^{mn}-\beta_{k,d}^m G_{kj}^{(m+\Delta m)n}\widetilde{s}_{kj}^{mn}-\sqrt{1-(\beta_{k,d}^m)^2}G_{e,kj}^{mn}\widetilde{s}_{kj}^{mn}\right]\right.\\
&\quad\left.\cdot\left[G_{kj}^{mn}\widetilde{s}_{kj}^{mn}+\widetilde{V}_{kj}^{mn}-\beta_{k,d}^m G_{kj}^{(m+\Delta m)n}\widetilde{s}_{kj}^{mn}-\sqrt{1-(\beta_{k,d}^m)^2}G_{e,kj}^{mn}\widetilde{s}_{kj}^{mn}\right]^{\mathrm{H}}\right)\\
&=\Re_{GG,k}(0,0)-2\beta_{k,d}^m\Re_{GG,k}(\Delta m,0)+(\beta_{k,d}^m)^2\Re_{GG,k}(0,0)+\sigma^2,
\end{aligned}\tag{B1}$$

where $\Re_{GG,k}(\Delta m,\Delta n)=\mathrm{E}\left(G_k^{mn}\left(G_k^{(m+\Delta m)(n+\Delta n)}\right)^*\right)$ denotes the time-frequency channel correlation function, and note that $\Delta n=0$ in this paper. We take the derivative of $\mathrm{E}\left(\|\,\widetilde{Y}_{kj}^{mn}-\widehat{G}_{kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right)$ with respect to $\beta_{k,d}^m$ and set it equal to zero; then we have

$$\frac{\partial\mathrm{E}\left(\|\,\widetilde{Y}_{kj}^{mn}-\widehat{G}_{kj}^{mn}\widetilde{s}_{kj}^{mn}\|^2\right)}{\partial\beta_{k,d}^m}=-2\Re_{GG,k}(\Delta m,0)+2\Re_{GG,k}(0,0)\beta_{k,d}^m=0,\tag{B2}$$

which yields

$$\beta_{k,d}^m=\frac{\Re_{GG,k}(\Delta m,0)}{\Re_{GG,k}(0,0)}.\tag{B3}$$

This completes the simplification of matrix $\boldsymbol{B}_k^{mn}$.

Appendix C Derivation of parameter $\boldsymbol{\beta_{k,e}^m}$

In what follows, when the MMSE-based channel estimator is used, the estimated channel complex gain is given by the form

$$\begin{aligned}
\widehat{G}_{kj}^{(m-\Delta m)n}&=\Re_{G\widetilde{Y}}\Re_{\widetilde{Y}\widetilde{Y}}^{-1}\widetilde{Y}_{kj}^{mn}=\frac{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\left(\widetilde{s}_{kj}^{mn}\right)^*\right)}{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\widetilde{s}_{kj}^{mn}\left(\widetilde{s}_{kj}^{mn}\right)^*\right)+\sigma^2}\left(G_{kj}^{mn}\widetilde{s}_{kj}^{mn}+\widetilde{V}_{kj}^{mn}\right)\\
&=\frac{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\widetilde{s}_{kj}^{mn}\left(\widetilde{s}_{kj}^{mn}\right)^*\right)}{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\widetilde{s}_{kj}^{mn}\left(\widetilde{s}_{kj}^{mn}\right)^*\right)+\sigma^2}G_{kj}^{mn}+\frac{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\left(\widetilde{s}_{kj}^{mn}\right)^*\right)\widetilde{V}_{kj}^{mn}}{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\widetilde{s}_{kj}^{mn}\left(\widetilde{s}_{kj}^{mn}\right)^*\right)+\sigma^2}\\
&\approx\frac{\gamma_{p,k}}{1+\gamma_{p,k}}G_{kj}^{mn}+\frac{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\left(\widetilde{s}_{kj}^{mn}\right)^*\right)\widetilde{V}_{kj}^{mn}}{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\widetilde{s}_{kj}^{mn}\left(\widetilde{s}_{kj}^{mn}\right)^*\right)+\sigma^2},
\end{aligned}\tag{C1}$$

where $\gamma_{p,k}^m = \frac{\mathrm{E}\left(G_{kj}^{mn}\left(G_{kj}^{mn}\right)^*\right)\mathrm{E}\left(\tilde{s}_{kj}^{mn}\left(\tilde{s}_{kj}^{mn}\right)^*\right)}{\sigma^2}$ is the equivalent SNR of pilot grids, which can be estimated by the SNR of data grids. From the first term on the right-hand side of (C1), we have

$$\hat{\beta}_{k,e}^m = \frac{\gamma_{p,k}^m}{1+\gamma_{p,k}^m}. \tag{C2}$$

Appendix D CP-based SNR estimation

In accordance with [26], the self-correlated function of the OFDM symbol is equal to the sum of signal and noise powers:

$$\frac{1}{2L}\left(\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\right|^2+\sum_{t=N}^{N+L-1}\left|\overline{y}_{kj}^m(t)\right|^2\right)\approx \widehat{P}_{R,kj}^m+\widehat{P}_{I,kj}^m+\widehat{P}_{N,kj}^m, \tag{D1}$$

where $\widehat{P}_{R,kj}^m$ denotes the average power of useful data signal at the receive signal, $\widehat{P}_{I,kj}^m$ is the residual interference from other users, $\widehat{P}_{N,kj}^m$ is the average power of AWGN noise in the channel.

Making use of the cyclic property of CP of the OFDM symbol, the correlation function between CP and its corresponding part in the OFDM symbol is approximated to be the average useful symbol power as

$$\hat{P}_{R,kj}^m = \frac{1}{L}\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\left(\overline{y}_{kj}^m(t+N)\right)^*\right|, \tag{D2}$$

due to the independent property of noise.

Substituting (D2) in (D1) yields

$$\begin{aligned}\hat{P}_{N,kj}^m &= \frac{1}{2L}\left(\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\right|^2+\sum_{t=N}^{N+L-1}\left|\overline{y}_{kj}^m(t)\right|^2\right)-\hat{P}_{R,kj}^m\\ &= \frac{1}{2L}\left(\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\right|^2+\sum_{t=N}^{N+L-1}\left|\overline{y}_{kj}^m(t)\right|^2\right)-\frac{1}{L}\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\left(\overline{y}_{kj}^m(t+N)\right)^*\right|\\ &= \frac{1}{2L}\left(\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\right|^2+\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t+N)\right|^2\right)-\frac{1}{2L}\sum_{t=0}^{L-1}2\left|\overline{y}_{kj}^m(t)\left(\overline{y}_{kj}^m(t+N)\right)^*\right|\\ &= \frac{1}{2L}\left(\sum_{t=0}^{L-1}\left(\left|\overline{y}_{kj}^m(t)\right|-\left|\overline{y}_{kj}^m(t+N)\right|\right)^2\right).\end{aligned} \tag{D3}$$

Combining (D2) and (D3) forms the estimated average SNR at data grids as follows:

$$\hat{\gamma}_{d,k}^m = \frac{\sum_{j=1}^J \hat{P}_{R,kj}^m}{\sum_{j=1}^J \hat{P}_{N,kj}^m} = \frac{\frac{1}{L}\sum_{j=1}^J\sum_{t=0}^{L-1}\left|\overline{y}_{kj}^m(t)\left(\overline{y}_{kj}^m(t+N)\right)^*\right|}{\frac{1}{2L}\sum_{j=1}^J\left(\sum_{t=0}^{L-1}\left(\left|\overline{y}_{kj}^m(t)\right|-\left|\overline{y}_{kj}^m(t+N)\right|\right)^2\right)}. \tag{D4}$$