ICASSP 2026 - 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 979-8-3315-6701-9/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICASSP55912.2026.11462350

FROM WMMSE TO XMMSE: AN OLD MELODY SUNG IN A FAST NEW TUNE

Zhexian Yang and Ziping Zhao

School of Information Science and Technology, ShanghaiTech University, Shanghai, China

ABSTRACT

Weighted sum-rate (WSR) maximization via linear beamforming is a central problem in wireless communications. A widely adopted approach to this problem is the WMMSE algorithm, which reformulates WSR maximization as an iteratively matrix-weighted sum-MSE minimization problem. In this paper, we revisit this classic problem and propose a novel transformation that reduces WSR maximization to a sum-regularized-MSE minimization problem. While both formulations build upon the fundamental relation between the achievable rate and the MSE matrix, WMMSE relies on an iterative approximation, whereas the proposed transformation exploits an exact rate–MSE relation. Based on this formulation, we develop an efficient algorithm, termed XMMSE, which alternately updates the transmit and receive beamformers. We further show that XMMSE admits a tighter surrogate function within the minorization–maximization framework, thereby improving convergence behavior. Numerical results further demonstrate that XMMSE achieves superior beamforming performance compared with several classical benchmark methods.

***Index Terms*—** Linear beamformer, MIMO interfering broadcast channel, weighted MMSE, weighted sum-rate maximization, minorization-maximization.

1. INTRODUCTION

Weighted sum-rate (WSR) maximization is a fundamental problem in communication systems for the design of linear transmit beamformers. This problem is generally non-convex and NP-hard [1], and iterative optimization algorithms are therefore commonly employed to obtain stationary solutions. In [2], an iterative algorithm is proposed for WSR maximization in multiple-input single-output (MISO) broadcast channels based on the uplink–downlink mean square error (MSE) duality, where a geometric program is solved at each iteration. An asynchronous distributed pricing algorithm for the WSR maximization problem in MISO interference channels was developed in [3], and it was shown that the algorithm converges monotonically to a stationary point under certain sufficient conditions. Similarly, [4] addresses WSR maximization in multiple-input multiple-output (MIMO) interference channels by optimizing transmit covariance matrices, where each receiver solves a local convex subproblem. A common limitation of these approaches is that they do not admit closed-form solutions at each iteration, which can result in relatively high computational complexity.

The weighted MMSE (WMMSE) criterion was introduced for MIMO channels in [5], where a generalized formulation was proposed to enable the design of specialized linear precoders through the adjustment of diagonal weight matrices. By exploiting the relationship between the mean square error (MSE) and the signal-to-interference-plus-noise ratio [6], the WMMSE algorithm for MIMO broadcast channels was subsequently developed in [7]. That work formulated a matrix-weighted sum-MSE minimization problem and established the connection between the optimal weights and the MSE matrix via the Karush–Kuhn–Tucker conditions. This line of research was further extended in [8] to MIMO interfering broadcast channels, where the authors demonstrated the equivalence between WSR maximization and weighted sum-MSE minimization, and interpreted the WMMSE algorithm as an alternating minimization procedure for solving the latter problem. Along a similar line, the fractional programming approach proposed in [9] addresses WSR maximization by exploiting essentially the same underlying principles as those of the WMMSE framework [8].

In this paper, we revisit the WSR maximization problem through a novel transformation that yields an equivalent formulation, which can be interpreted as a sum-regularized-MSE minimization problem. Based on this formulation, we propose the XMMSE algorithm, which alternates between updating the transmit and receive beamformers. Similar to WMMSE, XMMSE builds upon the relation between the achievable rate and the MSE matrix; however, while WMMSE relies on an iterative approximation, the proposed transformation exploits an exact rate–MSE relation. From the perspective of the minorization–maximization (MM) framework, XMMSE admits a tighter quadratic lower bound that more closely approximates the WSR objective function, thereby leading to improved convergence behavior. In numerical simulations, we compare XMMSE with existing WSR beamforming methods, including zero-forcing (ZF), regularized zero-forcing (RZF), maximum ratio transmission (MRT), and WMMSE [7, 8]. The results show that XMMSE achieves superior beamforming performance compared with ZF, RZF, and MRT. Moreover, XMMSE converges faster than WMMSE, with the performance advantage becoming more pronounced as the signal-to-noise ratio (SNR) increases.

2. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a general MIMO interfering broadcast channel with K transmitters, where the k-th transmitter communicates with a set of I_k intended receivers over a shared spectrum, and transmissions from different transmitters mutually interfere. The numbers of antennas at transmitter k and its i_k-th receiver are M_k and N_{i_k}, respectively, for $k = 1, \ldots, K$ and $i_k = 1, \ldots, I_k$. Let $\mathbf{V}_{i_k} \in \mathbb{C}^{M_k \times D_{i_k}}$ denote the linear beamforming matrix used by transmitter k for receiver i_k, and let $\mathbf{s}_{i_k} \in \mathbb{C}^{D_{i_k}}$ be the corresponding data vector with D_{i_k} spatial streams. The channel from transmitter l to receiver i_k is denoted by $\mathbf{H}_{i_k,l} \in \mathbb{C}^{N_{i_k} \times M_l}$. The received signal at receiver i_k is

$$\mathbf{y}_{i_k} = \underbrace{\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{s}_{i_k}}_{\text{desired signal}} + \underbrace{\sum_{(l,j)\neq(k,i)} \mathbf{H}_{i_k,l}\mathbf{V}_{j_l}\mathbf{s}_{j_l} + \mathbf{n}_{i_k}}_{\text{interference plus noise}}. \quad (1)$$

$$
\begin{aligned}
\mathrm{WSR}(\mathbf{V}) &= \int_0^1 \sum_{(k,i)} \omega_{i_k} \operatorname{tr}\left(\left(\mathbf{I}+\tau\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\right)^{-1}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\right)\mathrm{d}\tau \\
&\overset{(4.a)}{=} \int_0^1 \sum_{(k,i)} \omega_{i_k} \operatorname{tr}\left(\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1/2}\left(\mathbf{I}+\tau\mathbf{F}_{i_k}^{-1/2}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1/2}\right)^{-1}\mathbf{F}_{i_k}^{-1/2}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\right)\mathrm{d}\tau \\
&= \int_0^1 \sum_{(k,i)} \omega_{i_k} \operatorname{tr}\left(\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\left(\mathbf{F}_{i_k}+(1-\tau)\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\right)^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\right)\mathrm{d}\tau
\end{aligned} \tag{4}
$$

We assume that the signal covariance satisfies $\mathbb{E}\left[\mathbf{s}_{i_k}\mathbf{s}_{i_k}^{\mathsf{H}}\right] = \mathbf{I}$, and that the additive noise $\mathbf{n}_{i_k}$ at receiver i_k follows a circularly symmetric complex Gaussian distribution with mean $\mathbf{0}$ and covariance $\sigma_{i_k}^2\mathbf{I}$. Let matrix $\mathbf{\Gamma}_{i_k} = \mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}$ with $\mathbf{F}_{i_k} = \sum_{(l,j)\neq(k,i)}\mathbf{H}_{i_k,l}\mathbf{V}_{j_l}\mathbf{V}_{j_l}^{\mathsf{H}}\mathbf{H}_{i_k,l}^{\mathsf{H}} + \sigma_{i_k}^2\mathbf{I}$ for $i = 1,\ldots,I_k$ and $k = 1,\ldots,K$. The WSR maximization problem[1] is formulated as

$$
\begin{aligned}
\max_{\mathbf{V}} \quad & \sum_{(k,i)} \omega_{i_k} \log\det\left(\mathbf{I}+\mathbf{\Gamma}_{i_k}\right) =: \mathrm{WSR}(\mathbf{V}) \\
\text{s.t.} \quad & \sum_{i=1}^{I_k} \operatorname{tr}\left(\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\right) \leq P_k, k = 1,\ldots,K,
\end{aligned} \tag{2}
$$

where P_k denotes the transmit power budget of transmitter k.

3. XMMSE ALGORITHM FOR WSR MAXIMIZATION

In this section, we first reformulate the WSR maximization problem (2) as a sum-regularized-MSE minimization problem. We then propose a novel algorithm, termed XMMSE, to solve this problem.

3.1. Sum-regularized-MSE Minimization

Let $\mathbf{U}_{i_k} \in \mathbb{C}^{N_{i_k}\times D_{i_k}}$ denote the linear receive beamformer for receiver i_k. The estimated signal is $\widehat{\mathbf{s}}_{i_k} = \mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{y}_{i_k}$. The MSE matrix $\mathbf{E}_{i_k}$ is

$$
\begin{aligned}
\mathbf{E}_{i_k} &= \mathbb{E}_{\left\{\mathbf{s}_{i_k},\mathbf{n}_{i_k}\right\}}\left[\left(\widehat{\mathbf{s}}_{i_k}-\mathbf{s}_{i_k}\right)\left(\widehat{\mathbf{s}}_{i_k}-\mathbf{s}_{i_k}\right)^{\mathsf{H}}\right] \\
&= \mathbf{I} - \mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k} - \mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k} \\
&\quad + \mathbf{U}_{i_k}^{\mathsf{H}}\left(\sum_{(l,j)}\mathbf{H}_{i_k,j}\mathbf{V}_{j_l}\mathbf{V}_{j_l}^{\mathsf{H}}\mathbf{H}_{i_k,j}^{\mathsf{H}} + \sigma_{i_k}^2\mathbf{I}\right)\mathbf{U}_{i_k}.
\end{aligned}
$$

Then, we have the following result.

Theorem 1. *The WSR maximization problem in* (2) *is equivalent to the following sum-regularized-MSE minimization problem*

$$
\begin{aligned}
\min_{\mathbf{V}} \quad & \sum_{(k,i)} \omega_{i_k} \int_0^1 \min_{\mathbf{U}} \operatorname{tr}\left(\mathbf{E}_{i_k}-\tau\mathbf{R}_{i_k}\right)\mathrm{d}\tau \\
\text{s.t.} \quad & \sum_{i=1}^{I_k} \operatorname{tr}\left(\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\right) \leq P_k, k = 1,\ldots,K,
\end{aligned} \tag{3}
$$

where $\mathbf{R}_{i_k} = \mathbb{E}_{\mathbf{s}_{i_k}}\left[\widehat{\mathbf{r}}_{i_k}\widehat{\mathbf{r}}_{i_k}^{\mathsf{H}}\right]$, *and* $\widehat{\mathbf{r}}_{i_k} = \mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{s}_{i_k}$ *represents the estimated desired signal.*

[1] Throughout this paper, we use $\mathbf{V}$ and $\mathbf{U}$ to collectively denote the sets of beamformers $\{\mathbf{V}_{i_k}\}$ and $\{\mathbf{U}_{i_k}\}$, respectively.

Proof. We first introduce the following lemma.

Lemma 2. *For any matrix* $\mathbf{X} \succeq \mathbf{0}$*, we have*

$$
\log\det\left(\mathbf{I}+\mathbf{X}\right) = \int_0^1 \operatorname{tr}\left(\left(\mathbf{I}+\tau\mathbf{X}\right)^{-1}\mathbf{X}\right)\mathrm{d}\tau.
$$

Proof. According to the Newton-Leibniz formula, for any differentiable function $f(\mathbf{X})$, we have

$$
f(\mathbf{X}) - f(\mathbf{X}_0) = \int_0^1 \left\langle \nabla_{\mathbf{X}} f\left(\mathbf{X}_0+\tau\left(\mathbf{X}-\mathbf{X}_0\right)\right), \mathbf{X}-\mathbf{X}_0\right\rangle \mathrm{d}\tau.
$$

Setting $f(\mathbf{X}) = \log\det\left(\mathbf{I}+\mathbf{X}\right)$ and $\mathbf{X}_0 = \mathbf{0}$ leads to the result. □

According to Lemma 2, by letting $\mathbf{X} = \mathbf{\Gamma}_{i_k}$, it follows that

$$
\mathrm{WSR}(\mathbf{V}) = \sum_{(k,i)} \omega_{i_k} \int_0^1 \operatorname{tr}\left(\left(\mathbf{I}+\tau\mathbf{\Gamma}_{i_k}\right)^{-1}\mathbf{\Gamma}_{i_k}\right)\mathrm{d}\tau.
$$

Since $\mathbf{\Gamma}_{i_k} = \mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}$, we have (4), where $(4.a)$ uses the property $\left(\mathbf{I}+\mathbf{AB}\right)^{-1}\mathbf{A} = \mathbf{A}\left(\mathbf{I}+\mathbf{BA}\right)^{-1}$.

Let $\mathbf{R}_{i_k} = \mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}$. Note that $\mathbf{R}_{i_k} = \mathbb{E}_{\mathbf{s}_{i_k}}\left[\widehat{\mathbf{r}}_{i_k}\widehat{\mathbf{r}}_{i_k}^{\mathsf{H}}\right]$ is the covariance matrix of the estimated desired signal $\widehat{\mathbf{r}}_{i_k} = \mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{s}_{i_k}$, which is $\widehat{\mathbf{s}}_{i_k}$ with the interference and noise components excluded. For receiver i_k, minimizing the regularized MSE $\operatorname{tr}\left(\mathbf{E}_{i_k}-\tau\mathbf{R}_{i_k}\right)$ yields the closed-form solution

$$
\mathbf{U}_{i_k}^{\star}(\tau) = \left(\mathbf{F}_{i_k}+(1-\tau)\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\right)^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k},
$$

where $0 \leq \tau \leq 1$ is the regularization coefficient. The optimal regularized MSE is $\operatorname{tr}\left(\mathbf{I}-\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}^{\star}(\tau)\right)$. Substitute $\mathbf{U}_{i_k}^{\star}$ into the right-hand side of (4), the WSR can be reformulated as

$$
\mathrm{WSR}(\mathbf{V}) = \int_0^1 \sum_{(k,i)} \omega_{i_k} \operatorname{tr}\left(\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}^{\star}(\tau)\right)\mathrm{d}\tau.
$$

Hence, the problem (2) can be written as

$$
\begin{aligned}
\min_{\mathbf{V}} \quad & \sum_{(k,i)} \omega_{i_k} \int_0^1 \operatorname{tr}\left(\mathbf{I}-\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}^{\star}(\tau)\right)\mathrm{d}\tau \\
\text{s.t.} \quad & \sum_{i=1}^{I_k} \operatorname{tr}\left(\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\right) \leq P_k, k = 1,\ldots,K.
\end{aligned}
$$

This minimization problem is equivalent to problem (3). □

Unlike the WMMSE algorithm [7,8], which transforms the original WSR maximization problem (2) into a matrix-weighted sum-MSE minimization problem, Theorem 1 establishes an exact relation between the achievable rate and the MSE. In the following section, we propose an algorithm, named XMMSE, that alternates between updating $\mathbf{U}$ and $\mathbf{V}$ to solve the sum-regularized-MSE minimization problem in (3).

$$\mathbf{U}_{i_k}^{\mathrm{opt}} = \left((1-\tau)\,\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}} + \sum_{(l,j)\neq(k,i)} \mathbf{H}_{i_k,l}\mathbf{V}_{j_l}\mathbf{V}_{j_l}^{\mathsf{H}}\mathbf{H}_{i_k,l}^{\mathsf{H}} + \sigma_{i_k}^2\mathbf{I} \right)^{-1} \mathbf{H}_{i_k,k}\mathbf{V}_{i_k} \tag{5}$$

$$\begin{aligned}
\widetilde{\mathbf{A}}_{i_k} &= \mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{P}_{i_k}\int_0^1 (1-\tau)\left(\mathbf{I} - (1-\tau)\,\mathbf{\Lambda}_{i_k}\left(\mathbf{I}+(1-\tau)\,\mathbf{\Lambda}_{i_k}\right)^{-1}\right)^2 \mathrm{d}\tau\mathbf{P}_{i_k}^{\mathsf{H}}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k} \\
\mathbf{A}_{i_k,l} &= \mathbf{H}_{i_k,l}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{P}_{i_k}\int_0^1 \left(\mathbf{I} - (1-\tau)\,\mathbf{\Lambda}_{i_k}\left(\mathbf{I}+(1-\tau)\,\mathbf{\Lambda}_{i_k}\right)^{-1}\right)^2 \mathrm{d}\tau\mathbf{P}_{i_k}^{\mathsf{H}}\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,l} \\
\mathbf{B}_{i_k} &= \mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{P}_{i_k}\left(\mathbf{I} - \int_0^1 (1-\tau)\,\mathbf{\Lambda}_{i_k}\left(\mathbf{I}+(1-\tau)\,\mathbf{\Lambda}_{i_k}\right)^{-1}\mathrm{d}\tau\right)\mathbf{P}_{i_k}^{\mathsf{H}}
\end{aligned} \tag{6}$$

3.2. The XMMSE Algorithm

First, with the transmit beamformers $\mathbf{V}$ fixed, we solve the regularized-MSE minimization problem to obtain the update for receive beamformers $\mathbf{U}$ in (5). Then, given $\mathbf{U}$, the objective function of problem (3) can be expressed as

$$\sum_{(k,i)} \mathrm{tr}\left(\mathbf{V}_{i_k}^{\mathsf{H}}\widetilde{\mathbf{A}}_{i_k}\mathbf{V}_{i_k} + \sum_{(l,j)\neq(k,i)} \mathbf{V}_{j_l}^{\mathsf{H}}\mathbf{A}_{i_k,l}\mathbf{V}_{j_l} - 2\Re\left(\mathbf{B}_{i_k}^{\mathsf{H}}\mathbf{V}_{i_k}\right)\right),$$

where matrices

$$\begin{aligned}
\widetilde{\mathbf{A}}_{i_k} &= \int_0^1 (1-\tau)\,\mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}\mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}\,\mathrm{d}\tau, \\
\mathbf{A}_{i_k,l} &= \int_0^1 \mathbf{H}_{i_k,l}^{\mathsf{H}}\mathbf{U}_{i_k}\mathbf{U}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,l}\,\mathrm{d}\tau, \\
\mathbf{B}_{i_k} &= \int_0^1 \mathbf{H}_{i_k,k}^{\mathsf{H}}\mathbf{U}_{i_k}\,\mathrm{d}\tau, i=1,\dots,I_k; k,l=1,\dots,K.
\end{aligned}$$

Let $\mathbf{P}_{i_k}\mathbf{\Lambda}_{i_k}\mathbf{P}_{i_k}^{\mathsf{H}}$ denote the eigenvalue decomposition of $\mathbf{\Gamma}_{i_k}$. Applying the matrix inversion lemma, we obtain (6). For $i=1,\dots,I_k$ and $j,k=1,\dots,K$, denote $\widetilde{\mathbf{P}}_{i_k} = \mathbf{F}_{i_k}^{-1}\mathbf{H}_{i_k,k}\mathbf{V}_{i_k}\mathbf{P}_{i_k}$. After evaluating the integrals in (6), we obtain

$$\begin{aligned}
\widetilde{\mathbf{A}}_{i_k} &= \omega_{i_k}\mathbf{H}_{i_k,k}^{\mathsf{H}}\widetilde{\mathbf{P}}_{i_k}\mathrm{diag}\left\{ \frac{1}{\lambda_{i_k,n_{i_k}}}\left(\frac{\log(1+\lambda_{i_k,n_{i_k}})}{\lambda_{i_k,n_{i_k}}} - \frac{1}{1+\lambda_{i_k,n_{i_k}}} \right)\right\}\widetilde{\mathbf{P}}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,k}, \\
\mathbf{A}_{i_k,l} &= \omega_{i_k}\mathbf{H}_{i_k,l}^{\mathsf{H}}\widetilde{\mathbf{P}}_{i_k}\mathrm{diag}\left\{ \frac{1}{1+\lambda_{i_k,n_{i_k}}} \right\}\widetilde{\mathbf{P}}_{i_k}^{\mathsf{H}}\mathbf{H}_{i_k,l}, \\
\mathbf{B}_{i_k} &= \omega_{i_k}\mathbf{H}_{i_k,k}^{\mathsf{H}}\widetilde{\mathbf{P}}_{i_k}\mathrm{diag}\left\{ \frac{\log(1+\lambda_{i_k,n_{i_k}})}{\lambda_{i_k,n_{i_k}}} \right\}\mathbf{P}_{i_k}^{\mathsf{H}}.
\end{aligned}$$

For $i=1,\dots,I_k$ and $k=1,\dots,K$, we further denote matrix

$$\mathbf{A}_{i_k} = \widetilde{\mathbf{A}}_{i_k} + \sum_{(l,j)\neq(k,i)} \mathbf{A}_{j_l,k}. \tag{7}$$

After interchanging the order of summation in the quadratic term, the update of $\mathbf{V}_{i_k}$ can be decoupled across all transmitters, leading to the following optimization problem

$$\begin{aligned}
\min_{\{\mathbf{V}_{i_k}\}_{i=1}^{I_k}} \quad & \sum_{i=1}^{I_k} \mathrm{tr}\left(\mathbf{V}_{i_k}^{\mathsf{H}}\mathbf{A}_{i_k}\mathbf{V}_{i_k} - 2\Re\left(\mathbf{B}_{i_k}^{\mathsf{H}}\mathbf{V}_{i_k}\right)\right) \\
\text{s.t.} \quad & \sum_{i=1}^{I_k} \mathrm{tr}\left(\mathbf{V}_{i_k}\mathbf{V}_{i_k}^{\mathsf{H}}\right) \leq P_k,
\end{aligned} \tag{8}$$

Algorithm 1 The XMMSE Algorithm for WSR Beamforming

Input: The channel $\{\mathbf{H}_{i_k,j}\}$ and noise power $\sigma_{i_k}^2$.
1: Initialize $t \leftarrow 0$ and $\mathbf{V}^{(0)}$.
2: **while** not convergence **do**
3: Update beamformers $\mathbf{U}^{(t+1)}$ according to (5)
4: Update beamformers $\mathbf{V}^{(t+1)}$ according to (9).
5: $t \leftarrow t+1$.
6: **end while**
Output: Beamformers $\mathbf{V}^{(t)}$.

which obtains a solution as [8]

$$\mathbf{V}_{i_k}^{\mathrm{opt}} = \left(\mathbf{A}_{i_k} + \mu_k^{\star}\mathbf{I}\right)^{\dagger}\mathbf{B}_{i_k}, \tag{9}$$

where $\mu_k^{\star} \geq 0$ is the optimal Lagrange multiplier. Let $\mathbf{Q}_{i_k}\mathbf{\Sigma}_{i_k}\mathbf{Q}_{i_k}^{\mathsf{H}}$ be the eigenvalue decomposition of $\mathbf{A}_{i_k}$. Define $\mathbf{G}_{i_k} = \mathbf{B}_{i_k}^{\mathsf{H}}\mathbf{Q}_{i_k}$, and express it as $\mathbf{G}_{i_k} = [\mathbf{g}_{i_k,1},\dots,\mathbf{g}_{i_k,M_k}]$. For $k=1,\dots,K$, the optimal Lagrange multiplier $\mu_k^{\star}$ must satisfy

$$\mu_k^{\star} = \min\left\{ \mu_k \geq 0 : \sum_{i=1}^{I_k}\sum_{n=1}^{M_k} \frac{\|\mathbf{g}_{i_k,n}\|_2^2}{\left(\mathbf{\Sigma}_{i_k}(n,n)+\mu_k\right)^2} \leq P_k \right\}.$$

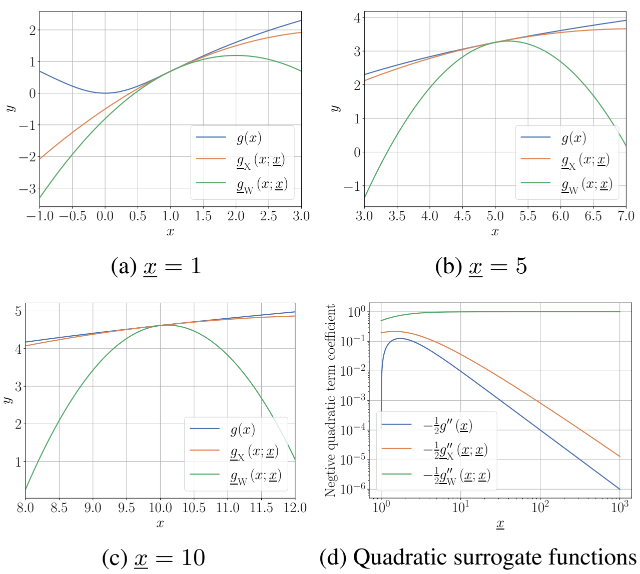
We can apply a line search method to determine the optimal value $\mu_k^{\star}$ within the closed interval $[0, b_k]$, where $b_k = \sqrt{\sum_{i=1}^{I_k}\|\mathbf{B}_{i_k}\|_F^2/P_k}$. The overall XMMSE algorithm is summarized in Algorithm 1. Furthermore, the following result demonstrates the theoretical convergence properties of the XMMSE algorithm.

Theorem 3. *Any limit point $\mathbf{V}^{\star}$ of the generated sequence by XMMSE is a stationary point of the WSR maximization problem* (2).

Remark. It is worth noting that [10] recently proposed the so-called reduced WMMSE algorithm for the WSR maximization problem in MIMO broadcast channels, achieving a computational complexity that scales linearly with the number of transmit antennas. This setting corresponds to a special case of the MIMO interfering broadcast channel with a single transmitter. Consequently, similar ideas in [10] can be incorporated into XMMSE to address this scenario.

4. COMPARISON BETWEEN XMMSE AND WMMSE

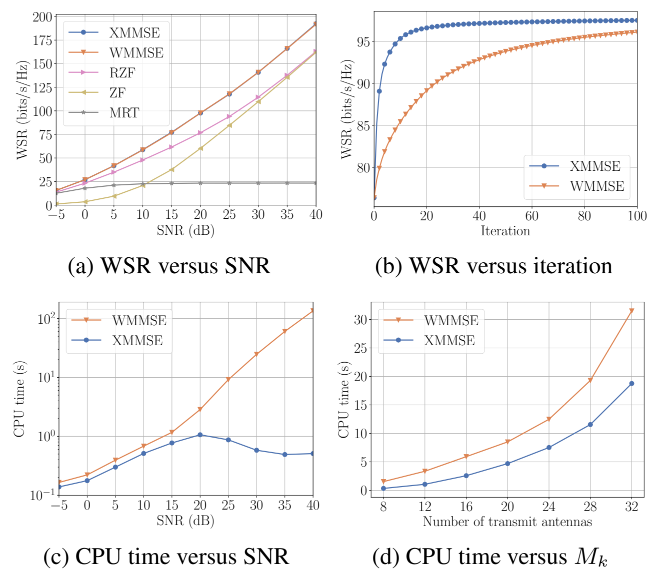
The per-iteration computational complexity of XMMSE is of the same order as that of WMMSE in terms of the numbers of base stations, data streams, transmit antennas, and receive antennas. In the WMMSE algorithm [7, 8], after updating the receive beamformers and the matrix weights, the resulting objective function becomes



(a) $\underline{x} = 1$ (b) $\underline{x} = 5$

(c) $\underline{x} = 10$ (d) Quadratic surrogate functions

Fig. 1. Comparison between XMMSE and WMMSE.



(a) WSR versus SNR (b) WSR versus iteration

(c) CPU time versus SNR (d) CPU time versus M_k

Fig. 2. The performance of the WSR maximization algorithms.

quadratic with respect to the transmit beamformers. It was shown in [11] that this quadratic function serves as a lower bound of the original WSR objective at given values of transmit beamformers within the MM framework. Similarly, in the proposed XMMSE algorithm, updating the receive beamformers also leads to a quadratic surrogate function. More importantly, the resulting quadratic surrogate in XMMSE provides a tighter lower bound than that induced by WMMSE. To illustrate this point in an abstract and intuitive way, we consider a simple scalar example that captures the essential structure of the WSR objective function. Let $x \in \mathbb{R}$ and define $g(x) = \log\left(1 + x^2\right)$. Given a reference point $\underline{x} \in \mathbb{R}$, the quadratic surrogate functions corresponding to XMMSE and WMMSE are given, respectively, by

$$\underline{g}_{\mathrm{X}}(x;\underline{x}) = \left(\frac{1}{1+\underline{x}^2} - \frac{\log\left(1+\underline{x}^2\right)}{\underline{x}^2}\right)(x-\underline{x})^2 + \underline{l}(x;\underline{x}),$$
$$\underline{g}_{\mathrm{W}}(x;\underline{x}) = \left(\frac{1}{1+\underline{x}^2} - 1\right)(x-\underline{x})^2 + \underline{l}(x;\underline{x}),$$

where $\underline{l}(x) = \frac{2\underline{x}}{1+\underline{x}^2}(x-\underline{x}) + g(\underline{x})$ denotes the first-order Taylor approximation of $g(x)$ at $\underline{x}$. Since the logarithm function is concave, we have $\log\left(1+\underline{x}^2\right) \leq \underline{x}^2$. Consequently, the difference between the two quadratic surrogate functions can be expressed as

$$\underline{g}_{\mathrm{X}}(x;\underline{x}) - \underline{g}_{\mathrm{W}}(x;\underline{x}) = \left(1 - \frac{\log\left(1+\underline{x}^2\right)}{\underline{x}^2}\right)(x-\underline{x})^2 \geq 0.$$

Fig. 1 illustrates the difference between the two quadratic lower bounds. As $\underline{x}$ increases, the gap between the surrogates induced by XMMSE and WMMSE becomes larger. In particular, as shown in Fig. 1(d), the quadratic surrogate associated with XMMSE closely matches the second-order Taylor expansion of $g(x)$ at $\underline{x}$. This observation is further corroborated by the numerical results presented in Section 5.

5. NUMERICAL RESULTS

We conduct numerical simulations for a MIMO interference channel with $I_k = 1$ for $k = 1, \ldots, K$, where $K = 4$, $M_k = 16$, $N_{i_k} = D_{i_k} = 4$, and $\omega_{i_k} = 1$ for all k. The noise powers are identical across receivers, i.e., $\sigma_{i_k}^2 = \sigma^2$, and under a total transmit power P, the SNR is defined as $\mathrm{SNR} = 10\log_{10}(P/\sigma^2)$. The simulations are conducted on a desktop computer equipped with an Intel Xeon W CPU. All results are averaged over 1000 independent channel realizations drawn from $\mathcal{CN}(0,1)$. We compare XMMSE with benchmark schemes including ZF, RZF, MRT, and WMMSE [7, 8], where all iterative algorithms are initialized using the RZF beamformer. As shown in Fig. 2(a), both XMMSE and WMMSE achieve higher WSR than the other benchmarks, and under the same convergence criterion they attain comparable WSR performance. Fig. 2(b) shows the convergence behavior at $\mathrm{SNR} = 20\,\mathrm{dB}$, where XMMSE converges faster than WMMSE. The fast convergence behavior of XMMSE becomes more evident at higher SNRs, as shown in Fig. 2(c). Finally, Fig. 2(d) reports the average CPU time with $N_{i_k} = D_{i_k} = 4$ and M_k varying from 8 to 32 in a fully loaded setting ($N_k = M_k/N_{i_k}$) at $\mathrm{SNR} = 20\,\mathrm{dB}$, where XMMSE consistently requires less computation time than WMMSE across all antenna configurations.

6. CONCLUSION

In this paper, we have studied the WSR maximization problem in MIMO interference broadcast channels. We have proposed a novel equivalent transformation of the WSR objective, leading to a sum-regularized-MSE minimization formulation, and have developed the corresponding XMMSE algorithm. The advantage of XMMSE over WMMSE has been theoretically established within the MM framework. Numerical results have further demonstrated that XMMSE achieves superior beamforming performance compared with several classical benchmark methods.

7. REFERENCES

[1] Zhi-Quan Luo and Shuzhong Zhang, "Dynamic Spectrum Management: Complexity and Duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[2] Shuying Shi, Martin Schubert, and Holger Boche, "Rate Optimization for Multiuser MIMO Systems With Linear Processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.

[3] Changxin Shi, Randall A. Berry, and Michael L. Honig, "Monotonic Convergence of Distributed Interference Pricing in Wireless Networks," in *2009 IEEE International Symposium on Information Theory*, June 2009, pp. 1619–1623.

[4] Seung-Jun Kim and Georgios B. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.

[5] H. Sampath, P. Stoica, and A. Paulraj, "Generalized Linear Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[6] D.P. Palomar, J.M. Cioffi, and M.A. Lagunas, "Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[7] Søren Skovgaard Christensen, Rajiv Agarwal, Elisabeth De Carvalho, and John M. Cioffi, "Weighted Sum-Rate Maximization using Weighted MMSE for MIMO-BC Beamforming Design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[8] Qingjiang Shi, Meisam Razaviyayn, Zhi-Quan Luo, and Chen He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sept. 2011.

[9] Kaiming Shen and Wei Yu, "Fractional Programming for Communication Systems—Part I: Power Control and Beamforming," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, May 2018.

[10] Xiaotong Zhao, Siyuan Lu, Qingjiang Shi, and Zhi-Quan Luo, "Rethinking WMMSE: Can Its Complexity Scale Linearly With the Number of BS Antennas?," *IEEE Transactions on Signal Processing*, vol. 71, pp. 433–446, 2023.

[11] Zepeng Zhang, Ziping Zhao, Kaiming Shen, Daniel P. Palomar, and Wei Yu, "Discerning and Enhancing the Weighted Sum-Rate Maximization Algorithms in Communications," *arXiv:2311.04546*, Nov. 2023.

[12] Ying Sun, Prabhu Babu, and Daniel P. Palomar, "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, Feb. 2017.