Signal Processing 232 (2025) 109932

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Analog–digital precoding based on mutual coupling considering the actual radiation performance of MIMO antenna arrays

Jianchuan Wei [a], Xiaoming Chen [a,*], Ruihai Chen [a], Chongwen Huang [b], Xiaoyu Huang [a], Wei E.I. Sha [b], Mérouane Debbah [c]

[a] School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, 710049, China
[b] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, 310027, China
[c] KU 6G Research Center, Khalifa University of Science and Technology, Abu Dhabi, 127788, United Arab Emirates

ARTICLE INFO

Keywords:
Antenna
Mutual coupling
Multiple-input multiple-output (MIMO)
Optimization
Precoding
Minimum mean square error (MMSE)
Zero-forcing (ZF)

ABSTRACT

Mutual coupling among antenna array elements can have impacts on not only the elements' mutual impedance but also the radiation characteristics, which is usually not considered appropriately in previous precoding researches. In this paper, an analog–digital precoding scheme based on mutual coupling is proposed considering the actual radiation performance of multiple-input multiple-output antenna arrays. Instead of reducing or eliminating the mutual coupling effect as in the conventional approach, this paper focuses on exploiting it to improve the system performance. Open-circuit radiation pattern is firstly introduced in the system model to further evaluate the effect of mutual coupling on the antenna radiation performance. Conventional linear precoding is applied in the digital domain. As for the analog domain, to minimize the noise amplification factor and thus maximize the receive signal-to-noise ratio of the system through exploiting the mutual coupling effect, the proper load impedances and antenna weighting coefficients are selected by convex optimization and Newton method. Simulation results show that the proposed scheme performs significantly better than the conventional one without mutual coupling in zero forcing and minimum mean square error systems. Besides, greater performance gain in MIMO systems with smaller antenna spacing, higher dimensionality and larger apparent power are observed.

1. Introduction

Multiple-input multiple-output (MIMO) technology deeply utilizes the spatial wireless resources by using spatial diversity and antenna multiplexing, greatly improving the spectrum efficiency without additional bandwidth or transmit power requirement [1,2]. MIMO technology has become an important component of modern wireless communication standards and is widely used in various wireless communication systems, including IEEE 802.11 (Wi-Fi), long-term evolution (LTE), the fifth generation (5G), and others [3].

However, in MIMO systems, mutual coupling is usually inevitable among transmitting or receiving antennas due to the limitation of physical space. In the initial low-dimensional MIMO communication systems, the mutual coupling problem between antennas in the MIMO system can be solved through reasonable spatial layout and larger antenna spacing. However, with the development of massive MIMO technology [4,5] and the proposal of holographic MIMO (HMIMO) [6–8] technology, deploying a large number of antenna elements in a compact space has become a necessary form for transmitter and receiver arrays in future communication systems. In this case, the non-negligible mutual coupling will cause significant decrease in various key performance indicators of the multi-antenna system [9–11]. Therefore, how to effectively overcome the mutual coupling effect between antennas is one of the major challenges faced in the development and application of MIMO multi-antenna technology.

Many techniques have been proposed in [12–15] to alleviate the performance loss caused by mutual coupling effects. In [12], a new type of defected ground structure was applied to microstrip array, and good decoupling performance was achieved. The authors in [13] proposed a dielectric block decoupling technique to suppress the in-band coupling in base station (BS) antenna array. An electromagnetic bandgap ground was proposed to mitigate the mutual coupling in [14]. An orthogonal directional pattern decoupling technique to reduce the coupling effect between 5G MIMO mobile phone antennas was proposed in [15]. Most of these existing methods mainly consider eliminating or minimizing the mutual coupling effect between MIMO antennas from the perspective of antenna design.

* Corresponding author.
E-mail address: xiaoming.chen@mail.xjtu.edu.cn (X. Chen).

https://doi.org/10.1016/j.sigpro.2025.109932
Received 18 November 2024; Received in revised form 15 December 2024; Accepted 31 January 2025
Available online 8 February 2025

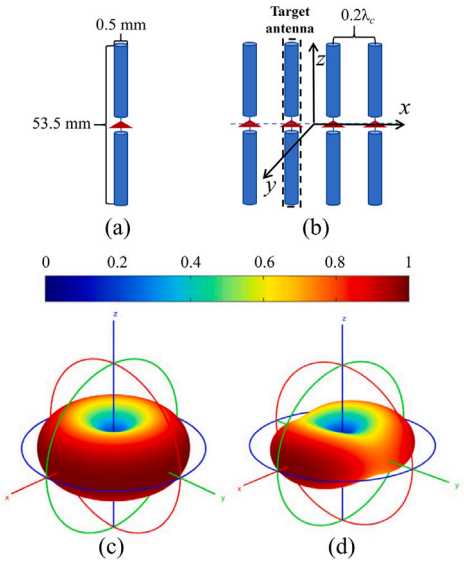

Fig. 1. (a) Design parameters of the dipole element. (b) Structure diagram of a 4-element uniform linear array (ULA) composed of half-wavelength dipoles with $0.2\lambda_c$ antenna spacing. (c) Isolated radiation pattern of the half-wavelength dipole. (d) Open-circuit radiation pattern of the target antenna (second port) in the ULA.

In [16–19], a joint analog–digital precoding method is put forward to address the antenna mutual coupling problem from the signal processing perspective. Unlike the traditional belief that mutual coupling effects will reduce the system performance, this approach considers utilizing the mutual coupling effect between the antenna elements to further improve the system performance. In the scheme proposed in [16–19], each antenna is assumed to be equipped with an adjustable load impedance (e.g., a variable capacitance diode), which can be used to control the mutual coupling effect by tuning the value of each load impedance so as to improve the system performance.

However, in the mutual coupling precoding scheme proposed in [16–19], the communication system model with antenna coupling is mainly established under the assumption of a minimum scattering antenna [20]. This hypothesis assumes that an open-circuited antenna has no effect on other antennas in the array. The hypothesis works well for antennas with small electrical size ($<0.1\lambda_c$, where λ_c is the carrier wavelength). However, the minimum scattering antenna hypothesis is not valid for most practical antennas in base stations or terminal devices (i.e., the open-circuit pattern[1] of the antenna can still be distorted by mutual coupling effects). Fig. 1(c) and (d) show the significant difference between the isolated pattern and open-circuit pattern of the second antenna element in a 4-element linear array composed of half-wavelength dipoles (carrier frequency is assumed as 2.6 GHz) with $0.2\lambda_c$ inter-element spacing, which demonstrate that the half-wavelength dipole does not fulfill the minimum scattering antenna hypothesis. (The design parameters of each dipole element are shown in Fig. 1(a), and the schematic diagram of the 4-element dipole array is presented in Fig. 1(b). The realized gain of the 4-element dipole array is 4.23 dBi at 2.6 GHz.) Hence, the distortions of the antenna patterns caused by mutual coupling effects have been overlooked in the previous works on mutual coupling precoding.

To address the aforementioned issue, we further consider the mutual coupling impact on the actual radiation performance of the antenna array with mutual coupling precoding in this paper. The actual radiation pattern of the antenna array is decomposed into the product of the open-circuit pattern and the current coefficient of each element and a more comprehensive channel model including the actual radiation performance of the antenna is established using the channel transfer function [21]. At this point, the influence of antenna mutual coupling and load impedance variation on the antenna pattern can be accurately characterized by the modified system model. Therefore, the mutual coupling precoding proposed in this work is applicable to arbitrary antennas.

Additionally, the antenna weighting coefficients frequently used in phased array systems for beamforming [22–24] have also been incorporated into the modified system model. Our mathematical analysis reveals that for a given set of optimized current coefficients of the mutual coupling precoding system, the load impedances can be linearly expressed in terms of the antenna weighting coefficients. This allows the optimization of load impedances in the system model to be converted into the optimization of antenna weighting coefficients. Moreover, an analysis of the feasible domain of current coefficients shows that varying the antenna weighting coefficients enables a larger feasible current coefficient domain compared to adjusting the load impedances.

Based on these findings, this work proposes a joint analog–digital mutual coupling precoding scheme that considers the actual radiation performance of the antenna. In the digital domain, the traditional linear precoding is used. Convex optimization and interior point methods are used to determine the optimal value of each load impedance or antenna weighting coefficient in the analog domain, minimizing the noise amplification factor of the digital precoder and maximizing the received signal-to-interference-plus-noise ratio (SINR) to improve the system performance.

For clarity, the main contributions of this paper are summarized as follows:

(1) A precoding scheme based on mutual coupling effects is proposed with consideration of the actual radiation characteristics of the antennas and realistic constraints. Unlike the traditional thought of reducing or eliminating the mutual coupling effects, this scheme focuses on exploiting the mutual coupling to improve the system performance, providing a promising solution to the high mutual coupling challenges may faced by ultra-massive MIMO and holographic MIMO systems.

(2) A more comprehensive system model is developed that incorporates the actual radiation performance of antennas via the channel transfer function. This model is broadly applicable to arbitrary antennas, rather than being limited to the restrictive minimum scattering antenna assumption used in previous works. Using this model, the effects of antenna mutual coupling, load impedances, and antenna weighting coefficients on the antenna pattern can be accurately evaluated.

(3) The optimization of adjustable load impedances and antenna weighting coefficients are discussed separately and a closed-form solution for the optimal current coefficient[2] that achieves the best SINR performance with fixed antenna weighting coefficients is deduced. Our mathematical analysis shows that varying antenna weighting coefficients results in a larger feasible current coefficient domain compared to adjusting load impedances, leading to better performance. Additionally, unlike adjustable impedances, varying antenna weighting coefficients is not affected by insertion loss, simplifying hardware implementation.

The rest of the paper is organized as follows: Section 2 introduces a point-to-point MIMO precoding system model that considers the mutual coupling and actual radiation performance of antennas. Section 3 provides the specific details on the proposed precoding methods that exploit the mutual coupling effect, and discusses the precoding methods based on antenna weighting coefficients and load impedances.

[1] The open-circuit radiation pattern of an antenna element in the antenna array is defined as the radiation pattern when it is excited while all the other antenna elements (ports) are open-circuited.

[2] The current coefficient represents the excitation weight of the open-circuit radiation pattern, which is determined by the antenna weighting coefficients and impedance [25–27].

Section 4 presents the numerical results to validate the superiority of the proposed method. Conclusions are given in Section 5.

Notation: Letters a, **a**, and **A** represent scalars, vectors and matrices, respectively. $|\mathbf{a}^T|$ denotes the operation of transposing **a** and then taking the modulus value for each element. $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$, $\mathbf{A}^\dagger$, $\|\mathbf{A}\|$, $\|\mathbf{A}\|_F$, $\|\mathbf{A}\|_\infty$ and $\mathrm{E}\{\mathbf{A}\}$ denote transpose, Hermitian (conjugate transpose), inverse, pseudo-inverse, 2 norm, Frobenius norm, infinite norm and expectation of **A**, respectively. $\mathbf{A}^{-H}$ denotes the Hermitian (conjugate transpose) of **A** and then inverse it. $\mathbf{A}_{i,j}$ represents the (i, j)th element of **A**. $|\cdot|$ and $(\cdot)^*$ denote the modulus and conjugate, respectively. $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary part of a complex variable. $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ give the trace and determinant of a matrix, respectively. **I** is the identity matrix. $\mathrm{diag}(\mathbf{a})$ represents a diagonal matrix with the entries of **a** on its main diagonal.

2. System model

2.1. Evaluation of the mutual coupling's impacts on the actual radiation performance of antenna arrays

Two critical metrics are introduced to evaluate the mutual coupling effects accurately in the system model, the coupling matrix [20] for the effect of mutual coupling on the circuit characteristics, and the open-circuit radiation pattern for that on the antenna radiation performance [25–27].

For the infinite uniform linear array (ULA) case, the open-circuit radiation pattern of the mth antenna (referenced to the phase center of the array) can be derived as [28]

$$F_m^{\mathrm{oc}}(\Omega) = \frac{1}{2}\left\{ Z_{\mathrm{TL},m} + \sum_{n=-\infty}^{+\infty} Z_{nm} \exp\left(\varphi_m(\Omega)\right) \right\} F^a(\Omega) \tag{1}$$

where $\Omega \triangleq (\theta, \phi)$ denotes the solid angle; $\varphi_m(\Omega) = jk_0 d(n-m)\sin\theta\cos\phi$ represents the phase difference with respect to the mth antenna; θ and ϕ represent the elevation angle and azimuth angle in the spherical coordinates; $Z_{\mathrm{TL},m}$ represents the characteristic impedance of the transmission line connected to the mth antenna port; Z_{nm} represents the mutual impedance between the mth and nth antennas in the array; k_0 represents the wavenumber in free space; d represents the spacing between antennas in the ULA, and $F^a(\Omega)$ represents the embedded radiation pattern of the antenna element in an infinite array (with one element excited and the rest is terminated with matching loads). Since (1) is for the infinite case, significant errors can be observed for the application in the finite antenna arrays, especially for the edge elements. Besides, the mutual coupling related terms in (1) also demonstrates that the open-circuit radiation pattern of the antenna is different from the isolated radiation pattern.

Assume a MIMO system with N_{T} transmit antennas and N_{R} receive antennas, the radiation pattern of the ULA can be expressed in terms of the open-circuit antenna patterns as

$$F^{array}(\Omega) = \sum_{m=1}^{N_{\mathrm{T}}} F_m^{\mathrm{oc}}(\Omega)\, c_m \exp\left(\varphi_m(\Omega)\right) \tag{2}$$

where c_m denotes the current coefficient of the mth antenna port. The mutual coupling effect on the circuit characteristics of the transmitter can be represented by a $N_{\mathrm{T}} \times N_{\mathrm{T}}$ coupling impedance matrix $\mathbf{\Gamma}$. For Massive MIMO or HMIMO systems with spatial constraints, the mutual coupling effect between antenna elements cannot be ignored. If the antenna array is a ULA with spacing d, the coupling matrix $\mathbf{\Gamma}$ can be expressed as

$$\mathbf{\Gamma} = \begin{bmatrix} z_{A1} & z_{12} & \cdots & z_{1N_{\mathrm{t}}} \\ z_{21} & z_{A2} & \cdots & z_{2N_{\mathrm{t}}} \\ \vdots & \vdots & \ddots & \vdots \\ z_{N_{\mathrm{t}}1} & z_{N_{\mathrm{t}}2} & \cdots & z_{AN_{\mathrm{t}}} \end{bmatrix} \tag{3}$$

where z_{Aj} denotes the self-impedance of the jth antenna, z_{mn} represents the mutual impedance between the mth and nth antennas, $m, n \in \{1, 2, \ldots, N_{\mathrm{t}}\}$. The value of z_{mn} can be obtained by the electromagnetic-field (EMF) approach [22] or the inverse Fourier transform of the Floquet impedance [28]. For simplicity, we denote the antenna weighting coefficient vector as $\boldsymbol{\omega} = \left[\omega_1, \omega_2, \ldots \omega_{N_{\mathrm{T}}}\right]^T$, the current coefficient vector as $\mathbf{c} = \left[c_1, c_2, \ldots c_{N_{\mathrm{T}}}\right]^T$, and the diagonal matrix consisting of load impedances as $\mathbf{Z}_{\mathrm{L}} = \mathrm{diag}\left(z_{\mathrm{L},1}, z_{\mathrm{L},2}, \ldots z_{\mathrm{L},N_{\mathrm{T}}}\right)$.

To calculate the system's current coefficient conveniently, we assume that in the transmitter or receiver antenna system, the connections between the source, load, and antenna are made through ideal conductors, without considering the impact of transmission lines on the system, same as the previous works [16–19]. Note that, since the transmission lines' effects can be easily included in the system model, this is not a strong assumption. The system's current coefficient can then be calculated as (see Appendix)

$$\mathbf{c} = (\mathbf{Z}_{\mathrm{L}} + \mathbf{\Gamma})^{-1}\boldsymbol{\omega}. \tag{4}$$

The radiation pattern of the array can be expressed as

$$F^{\mathrm{array}}(\Omega) = \sum_{m=1}^{N_{\mathrm{T}}} c_m F_m^{\mathrm{oc}}(\Omega) = \mathbf{g}_{\mathrm{oc}}^T(\Omega)\left(\mathbf{Z}_{\mathrm{L}} + \mathbf{\Gamma}\right)^{-1}\boldsymbol{\omega} \tag{5}$$

where $\mathbf{g}_{\mathrm{oc}}(\Omega) = \left[F_1^{\mathrm{oc}}(\Omega), F_2^{\mathrm{oc}}(\Omega), \ldots, F_{N_{\mathrm{T}}}^{\mathrm{oc}}(\Omega)\right]^T$. Eq. (5) can accurately evaluate the actual radiation performance of the MIMO antenna array with consideration of the mutual coupling effect on both circuit characteristics and radiation performance. Moreover, the relations between the load impedance $\mathbf{Z}_{\mathrm{L}}$, the antenna weighting coefficient $\boldsymbol{\omega}$ and the radiation pattern of the entire antenna array $F^{\mathrm{array}}(\Omega)$ are shown in (5). It can be seen that the change in load impedances or antenna weighting coefficients will cause significant deformation of the overall radiation pattern of the antenna array. Therefore, in a communication system, in a communication system that leverages mutual coupling by tuning load impedances or antenna weighting coefficients, the impact on the actual radiation pattern performance of the antenna array should also be considered.

Fig. 2(b) and (d) show the comparison between the array pattern (blue dots) synthesized by the open-circuit pattern through (5) and the array pattern (solid line) simulated by the full-wave simulation of a 4-element ULA composed of half-wavelength dipoles with $0.2\lambda_c$ inter-element separation. Two different cases of load impedances at each port of the array are shown in Fig. 2(a) and (c), respectively. Good conformity between the simulation and calculation results (c.f., Fig. 2(b) and (d)) demonstrates the correctness of (5) in characterizing the mutual coupling effect and the actual array radiation pattern changes caused by the variations in load impedances or antenna weighting coefficients.

2.2. Channel model considering the actual radiation performance of the antenna arrays

To accurately assess the impact of mutual coupling effects on the actual radiation performance of antenna arrays in communication model, channel transfer function (CTF) [21] is employed. CTF describes the input–output relationship among the transmitting (Tx) and receiving (Rx) antenna ports in a given propagation channel. Typically, the propagation channel between the Tx and Rx antennas is modeled as a superposition of many scattering paths. For an $N_{\mathrm{T}} \times N_{\mathrm{R}}$ MIMO system, the channel transfer function between the sth Tx antenna and the uth Rx antenna at time t and center frequency f_c can be expressed as [29]

$$\begin{aligned} h_{u,s}(t, f_c) = \int\int\int\int & \begin{bmatrix} F_s^{\mathrm{V}}\left(\Omega^{\mathrm{Tx}}\right) \\ F_s^{\mathrm{H}}\left(\Omega^{\mathrm{Tx}}\right) \end{bmatrix}^T \mathbf{S}\left(\tau, \upsilon, \Omega^{\mathrm{Tx}}, \Omega^{\mathrm{Rx}}\right) \begin{bmatrix} F_u^{\mathrm{V}}\left(\Omega^{\mathrm{Rx}}\right) \\ F_u^{\mathrm{H}}\left(\Omega^{\mathrm{Rx}}\right) \end{bmatrix} \\ & \cdot \exp(j2\pi\upsilon t)\exp(-j2\pi f_c\tau) d\tau d\upsilon\, d\Omega^{\mathrm{Tx}} d\Omega^{\mathrm{Rx}} \end{aligned} \tag{6}$$

where τ represents the time delay; υ denotes the Doppler shift; Ω^{Tx} and Ω^{Rx} represent direction of departure (DoD) and direction of arrival (DoA), respectively; $F_u^{\mathrm{V}}(\Omega)$ and $F_u^{\mathrm{H}}(\Omega)$ are the vertical polarization (V-pol) and horizontal polarization (H-pol) antenna radiation patterns of

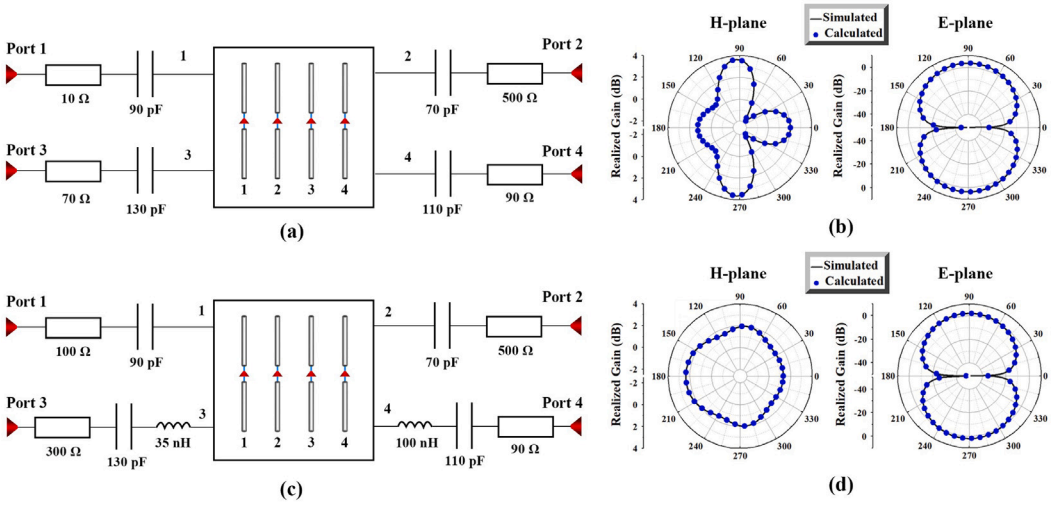

Fig. 2. Comparisons of simulated and calculated radiation pattern results of 4-port half wavelength dipole uniform linear array with $0.2\lambda_c$ inter-element spacing at 2.6 GHz. (a) Load impedance Case 1. (b) Simulated and calculated radiation patterns corresponding to Case 1. (c) Load impedance Case 2. (d) Simulated and calculated radiation patterns corresponding to Case 2.

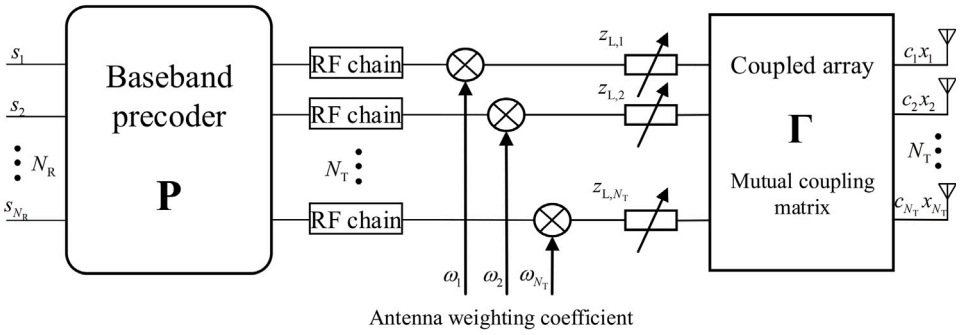

Fig. 3. Diagram of a N-port antenna transmitter system loaded with adjustable impedance and antenna weighting coefficients. The multi-antenna system can consist of any type of antennas and array configuration.

the uth Rx antenna at solid angle Ω, respectively; $F_s^{\mathrm{V}}(\Omega)$ and $F_s^{\mathrm{H}}(\Omega)$ are defined analogously for the sth Tx antenna. Note that the phase reference point of each port's radiation pattern in (6) is at the center of the array. Additionally, the matrix $\mathbf{S}$ in (6) is the so-called scattering function [29], which can be written in the Geometrically-Based Stochastic Channel Model (GSCM) as

$$\begin{aligned}\mathbf{S}(\tau,\upsilon,\Omega^{\mathrm{Tx}},\Omega^{\mathrm{Rx}}) = \sum_{m=1}^{M}\sqrt{P_m}\mathbf{A}\delta(\tau-\tau_m)\,\delta(\upsilon-\upsilon_m)\\ \cdot\delta\left(\Omega^{\mathrm{Tx}}-\Omega_m^{\mathrm{Tx}}\right)\delta\left(\Omega^{\mathrm{Rx}}-\Omega_m^{\mathrm{Rx}}\right)\end{aligned} \tag{7}$$

where τ_m, υ_m, Ω_m^{Tx}, Ω_m^{Rx} and P_m represent the time delay, Doppler shift, DoD, DoA and power of the mth path, respectively; M represents the total number of paths, $\delta(\cdot)$ is the Dirac delta function, and $\mathbf{A}$ is the polarization matrix

$$\mathbf{A}=\begin{bmatrix}\exp\left(j\boldsymbol{\Phi}_m^{\mathrm{VV}}\right) & \sqrt{\kappa_{1,m}^{-1}}\exp\left(j\boldsymbol{\Phi}_m^{\mathrm{VH}}\right)\\ \sqrt{\kappa_{2,m}^{-1}\chi_m^{-1}}\exp\left(j\boldsymbol{\Phi}_m^{\mathrm{HV}}\right) & \sqrt{\chi_m^{-1}}\exp\left(j\boldsymbol{\Phi}_m^{\mathrm{HH}}\right)\end{bmatrix} \tag{8}$$

with $\boldsymbol{\Phi}_m^{\mathrm{VV}}$, $\boldsymbol{\Phi}_m^{\mathrm{HV}}$, $\boldsymbol{\Phi}_m^{\mathrm{VH}}$ and $\boldsymbol{\Phi}_m^{\mathrm{HH}}$ representing the initial phase of vertical-to-vertical, horizontal-to-vertical, vertical-to-horizontal, and horizontal-to-horizontal polarization in the mth path, respectively; $\kappa_{1,m}$ and $\kappa_{2,m}$ being the cross-polarization ratios (XPRs) of the mth path, and χ_m denoting the co-polarization ratio (CPR) of the mth path.

By substituting (7) and (8) into (6), we can obtain

$$\begin{aligned}h_{u,s}(t,f_c) = \sum_{m=1}^{M}\sqrt{P_m}\begin{bmatrix}F_s^{\mathrm{V}}\left(\Omega_m^{\mathrm{Tx}}\right)\\ F_s^{\mathrm{H}}\left(\Omega_m^{\mathrm{Tx}}\right)\end{bmatrix}^T\mathbf{A}\begin{bmatrix}F_u^{\mathrm{V}}\left(\Omega_m^{\mathrm{Rx}}\right)\\ F_u^{\mathrm{H}}\left(\Omega_m^{\mathrm{Rx}}\right)\end{bmatrix}\\ \cdot\exp\left(j2\pi\upsilon_m t\right)\exp\left(-j2\pi f_c\tau_m\right).\end{aligned} \tag{9}$$

According to the relationship between the actual radiation pattern of the antenna array and the current coefficient, e.g., (5), the radiation pattern of the transmitting and receiving antennas considering mutual coupling can be represented by the product of their corresponding current coefficient and the open circuit pattern

$$\begin{bmatrix}F_m^{\mathrm{V}}(\Omega)\\ F_m^{\mathrm{H}}(\Omega)\end{bmatrix} = c_m\begin{bmatrix}F_m^{\mathrm{oc,V}}(\Omega)\\ F_m^{\mathrm{oc,H}}(\Omega)\end{bmatrix}. \tag{10}$$

Substituting (10) into (9), we can obtain

$$h_{u,s}(t,f_c) = c_u^{\mathrm{Rx}}h_{u,s}^{\mathrm{oc}}(t,f_c)c_s^{\mathrm{Tx}} \tag{11}$$

where c_s^{Tx} and c_u^{Rx} are the current coefficients of the sth transmitting antenna port and the uth receiving antenna port, respectively, and $h_{u,s}^{\mathrm{oc}}(t,f_c)$ is the channel from the open-circuit radiation pattern, which can be expressed as

$$\begin{aligned}h_{u,s}^{\mathrm{oc}}(t,f_c) = \sum_{m=1}^{M}\sqrt{P_m}\begin{bmatrix}F_s^{\mathrm{oc,V}}\left(\Omega_m^{\mathrm{Tx}}\right)\\ F_s^{\mathrm{oc,H}}\left(\Omega_m^{\mathrm{Tx}}\right)\end{bmatrix}^T\mathbf{A}\begin{bmatrix}F_u^{\mathrm{oc,V}}\left(\Omega_m^{\mathrm{Rx}}\right)\\ F_u^{\mathrm{oc,H}}\left(\Omega_m^{\mathrm{Rx}}\right)\end{bmatrix}\\ \cdot\exp\left(j2\pi\upsilon_m t\right)\exp\left(-j2\pi f_c\tau_m\right).\end{aligned} \tag{12}$$

For simplicity, we denote $h(t,f_c)$ as h. For the $N_{\mathrm{T}}\times N_{\mathrm{R}}$ MIMO system, the total channel matrix can be represented as (13). For simplicity, we

$$\mathbf{H}=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_{\mathrm{T}}} \\ h_{21} & h_{22} & \cdots & h_{2N_{\mathrm{T}}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_{\mathrm{R}}1} & h_{N_{\mathrm{R}}2} & \cdots & h_{N_{\mathrm{R}}N_{\mathrm{T}}} \end{bmatrix}=\begin{bmatrix} c_1^{\mathrm{Rx}}h_{11}^{\mathrm{oc}}c_1^{\mathrm{Tx}} & c_1^{\mathrm{Rx}}h_{12}^{\mathrm{oc}}c_2^{\mathrm{Tx}} & \cdots & c_1^{\mathrm{Rx}}h_{1N_{\mathrm{T}}}^{\mathrm{oc}}c_{N_{\mathrm{T}}}^{\mathrm{Tx}} \\ c_2^{\mathrm{Rx}}h_{21}^{\mathrm{oc}}c_1^{\mathrm{Tx}} & c_2^{\mathrm{Rx}}h_{22}^{\mathrm{oc}}c_2^{\mathrm{Tx}} & \cdots & c_2^{\mathrm{Rx}}h_{2N_{\mathrm{T}}}^{\mathrm{oc}}c_{N_{\mathrm{T}}}^{\mathrm{Tx}} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_{\mathrm{R}}}^{\mathrm{Rx}}h_{N_{\mathrm{R}}1}^{\mathrm{oc}}c_1^{\mathrm{Tx}} & c_{N_{\mathrm{R}}}^{\mathrm{Rx}}h_{N_{\mathrm{R}}2}^{\mathrm{oc}}c_2^{\mathrm{Tx}} & \cdots & c_{N_{\mathrm{R}}}^{\mathrm{Rx}}h_{N_{\mathrm{R}}N_{\mathrm{T}}}^{\mathrm{oc}}c_{N_{\mathrm{T}}}^{\mathrm{Tx}} \end{bmatrix}$$
$$=\begin{bmatrix} c_1^{\mathrm{Rx}} & & & \\ & c_2^{\mathrm{Rx}} & & \\ & & \ddots & \\ & & & c_{N_{\mathrm{R}}}^{\mathrm{Rx}} \end{bmatrix}\begin{bmatrix} h_{11}^{\mathrm{oc}} & h_{12}^{\mathrm{oc}} & \cdots & h_{1N_{\mathrm{T}}}^{\mathrm{oc}} \\ h_{21}^{\mathrm{oc}} & h_{22}^{\mathrm{oc}} & \cdots & h_{2N_{\mathrm{T}}}^{\mathrm{oc}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_{\mathrm{R}}1}^{\mathrm{oc}} & h_{N_{\mathrm{R}}2}^{\mathrm{oc}} & \cdots & h_{N_{\mathrm{R}}N_{\mathrm{T}}}^{\mathrm{oc}} \end{bmatrix}\begin{bmatrix} c_1^{\mathrm{Tx}} & & & \\ & c_2^{\mathrm{Tx}} & & \\ & & \ddots & \\ & & & c_{N_{\mathrm{T}}}^{\mathrm{Tx}} \end{bmatrix} \tag{13}$$

Box I.

denote the open-circuit channel matrix as (see Eq. (13) given in Box I).

$$\mathbf{H}_{\mathrm{oc}}=\begin{bmatrix} h_{11}^{\mathrm{oc}} & h_{12}^{\mathrm{oc}} & \cdots & h_{1N_{\mathrm{T}}}^{\mathrm{oc}} \\ h_{21}^{\mathrm{oc}} & h_{22}^{\mathrm{oc}} & \cdots & h_{2N_{\mathrm{T}}}^{\mathrm{oc}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_{\mathrm{R}}1}^{\mathrm{oc}} & h_{N_{\mathrm{R}}2}^{\mathrm{oc}} & \cdots & h_{N_{\mathrm{R}}N_{\mathrm{T}}}^{\mathrm{oc}} \end{bmatrix} \tag{14}$$

Based on the relationship of the current coefficient with the load impedance and antenna weighting coefficient, i.e., (4), the channel matrix that takes into account the mutual coupling and the actual radiation performance of the antennas can be expressed as

$$\mathbf{H}=\mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}^{\mathrm{Rx}}+\boldsymbol{\Gamma}^{\mathrm{Rx}}\right)^{-1}\boldsymbol{\omega}^{\mathrm{Rx}}\right)\mathbf{H}_{\mathrm{oc}}\mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}^{\mathrm{Tx}}+\boldsymbol{\Gamma}^{\mathrm{Tx}}\right)^{-1}\boldsymbol{\omega}^{\mathrm{Tx}}\right) \tag{15}$$

where $\mathbf{Z}_{\mathrm{L}}^{\mathrm{Tx}}$, $\boldsymbol{\Gamma}^{\mathrm{Tx}}$ and $\boldsymbol{\omega}^{\mathrm{Tx}}$ represent the load impedance matrix, the coupling impedance matrix and the antenna weighting coefficient vector of the transmitting antenna array, respectively. $\mathbf{Z}_{\mathrm{L}}^{\mathrm{Rx}}$, $\boldsymbol{\Gamma}^{\mathrm{Rx}}$ and $\boldsymbol{\omega}^{\mathrm{Rx}}$ are defined as their counterparts at the receive side analogously.

Fig. 3 provides a depiction of the system diagram. According to (2) to (15), it can be found that the current coefficient $\mathbf{c}$ is the weighting coefficient that really acts on the signal and the effect of mutual coupling on the circuit characteristics is included in $\mathbf{c}$. In addition, it is also the current coefficients $\mathbf{c}$ and the open-circuit radiation patterns that together determine the actual radiation performance of the antenna. However, it is challenging to directly control the current coefficient $\mathbf{c}$ in practice. Therefore, we contemplate an indirect approach by manipulating the antenna weighting coefficient of each antenna port or the load impedance of each chain to regulate the current coefficient $\mathbf{c}$.

2.3. System model

Considering a point-to-point MIMO downlink system, where a BS with N_{T} transmitting antennas communicates with a UE with N_{R} receiving antennas. For a flat-fading MIMO channel, the received signal can be expressed as

$$\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{n} \tag{16}$$

where $\mathbf{H}\in C^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$ is the channel matrix containing the actual antenna radiation performance; $\mathbf{x}\in C^{N_{\mathrm{T}}\times 1}$ is the signal vector to be transmitted, and each element of $\mathbf{n}\in C^{N_{\mathrm{R}}\times 1}$ is assumed to be a complex additive white Gaussian noise (AWGN) with zero mean and σ^2 variance.

In this paper, we mainly consider the analog–digital precoding at the BS. Assuming the antennas of UE are isotropic, the communication system model can be represented as

$$\mathbf{y}=\mathbf{H}_{\mathrm{oc}}\mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}^{\mathrm{Tx}}+\boldsymbol{\Gamma}^{\mathrm{Tx}}\right)^{-1}\boldsymbol{\omega}^{\mathrm{Tx}}\right)\mathbf{x}+\mathbf{n}. \tag{17}$$

From (16), it can be seen that the load impedance and antenna weighting coefficient provide two additional degrees of freedom for the communication system to adjust the mutual coupling between antennas, and, thereby, to change the channel transfer function. For simplicity, we denote $\mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}^{\mathrm{Tx}}+\boldsymbol{\Gamma}^{\mathrm{Tx}}\right)^{-1}\boldsymbol{\omega}^{\mathrm{Tx}}\right)$ as $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$. It is worth noting that although it is in the form of a diagonal matrix, the mutual impedance describing the mutual coupling characteristics of the circuit in the mutual impedance matrix is already embodied in each (diagonal) element of $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ due to the mutual impedance terms in $\boldsymbol{\Gamma}$ and the determinant term of the matrix inversion $(\mathbf{Z}_{\mathrm{L}}+\boldsymbol{\Gamma})^{-1}$. This is essentially consistent with the traditional precoding schemes that considering mutual coupling [17,18].

For closed-loop MIMO systems, assuming the transmitter has perfect CSI, joint analog–digital precoding can be performed, and the communication system can be represented as

$$\mathbf{y}=\frac{1}{f}\mathbf{H}_{\mathrm{oc}}\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}\mathbf{P}\mathbf{s}+\mathbf{n} \tag{18}$$

where $\mathbf{s}\in C^{N_{\mathrm{R}}\times 1}$ is the data symbol vector, satisfying $\mathrm{E}\{\mathbf{s}\mathbf{s}^H\}=\mathbf{I}$; $\mathbf{P}\in C^{N_{\mathrm{T}}\times N_{\mathrm{R}}}$ is the precoding matrix; $\mathbf{x}=\frac{1}{f}\mathbf{P}\mathbf{s}$ is the pre-encoded signal to be transmitted, and f is the scaling factor that prevents non-physical amplification of the signal after passing through the precoder.

3. The proposed mutual coupling precoding scheme

3.1. Joint analog–digital precoding based on mutual coupling

In the proposed mutual coupling precoding scheme, each antenna element is assumed to be equipped with a tunable varactor diode as the load impedance in order to control the mutual coupling effect of the antennas. The scheme aims at maximizing the received SINR by optimizing the values of antenna weighting coefficients and load impedances to improve the performance of the communication system. The analysis of (5)–(15) shows that the antenna weighting coefficients and load impedances affect the actual radiation patterns of the MIMO antenna array, which in turn affect the overall communication channel. To exploit the mutual coupling effect, based on (18), the precoding matrix $\mathbf{P}$ can be constructed as

$$\mathbf{P}=\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}^{-1}\mathbf{W}_{\mathrm{oc}} \tag{19}$$

where $\mathbf{W}_{\mathrm{oc}}$ is the precoding matrix based on the open-circuit channel matrix $\mathbf{H}_{\mathrm{oc}}$. For simplicity, $\mathbf{W}_{\mathrm{oc}}$ is denoted as the open-circuit precoding matrix.

When considering the zero-forcing (ZF) precoding, $\mathbf{W}_{\mathrm{oc}}$ can be expressed as

$$\mathbf{W}_{\mathrm{oc}}^{\mathrm{ZF}}=\mathbf{H}_{\mathrm{oc}}^{H}\left(\mathbf{H}_{\mathrm{oc}}\mathbf{H}_{\mathrm{oc}}^{H}\right)^{-1}. \tag{20}$$

For other digital precoding schemes, this can be easily achieved by varying $\mathbf{W}_{\mathrm{oc}}$. In order to prevent the baseband signals from being unphysically amplified due to the precoding matrix, the scaling factor can be expressed as

$$f=\|\mathbf{P}\|_F=\sqrt{\mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right)}. \tag{21}$$

In this case, the received SINR for the kth data stream of the system can be obtained as

$$\gamma_k=\frac{\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_k\right|^2}{\sum_{i\neq k}\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_i\right|^2+f^2\sigma^2} \tag{22}$$

where $\mathbf{w}_k$ denotes the kth column vector of the open-circuit precoding $\mathbf{W}_{\mathrm{oc}}$ and $\mathbf{H}_{\mathrm{oc}} = \left[\mathbf{h}_{\mathrm{oc},1}^T, \mathbf{h}_{\mathrm{oc},2}^T, \ldots, \mathbf{h}_{\mathrm{oc},N_{\mathrm{R}}}^T\right]^T$. From the above equation, it can be seen that f is essentially the noise amplification factor. In particular, when ZF precoding is considered, the received SINR of the kth data stream of the system can be simplified as

$$\gamma_k = \frac{1}{f^2\sigma^2}. \tag{23}$$

From (22) and (23), it can be found that the precoding matrix $\mathbf{P}$ amplifies the noise while equalizing the inter-stream interference. Therefore, optimizing the antenna weighting coefficients and load impedances is equivalent to minimizing the noise amplification factor f in order to enhance the received SINR. However, the adjustments of the load impedances and antenna weighting coefficients are subject to certain practical constraints. Firstly, to ensure the proper radiation of the antenna, the adjustable load must satisfy the constraint

$$\Re\left(z_{\mathrm{L},i}\right) \geq 0, \forall i = 1, 2, \ldots, N_{\mathrm{T}}. \tag{24}$$

Additionally, in tuning the load impedances and antenna weighting coefficients, the total apparent power (which is a vector sum of the active power and reactive power) of the transmitter system cannot be amplified, hence the constraint

$$\left|\boldsymbol{\omega}^T\right| \left|(\mathbf{Z}_{\mathrm{L}} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\omega}\right| \leq P_0 \tag{25}$$

should be adhered to. It is worth noting that this study introduces a constraint on the total apparent power of all chains within the transmitter system. In essence, this constraint represents a more stringent form of the constraint applied to the active power (the sum of the radiated power and the dissipated power due to the ohmic loss) of the transmitter system [30,31].

Consequently, the precoding based on the antenna weighting coefficient and load impedance can be transformed into the following optimization problem

$$\begin{aligned} &\mathcal{P}_1 : \max\nolimits_{\boldsymbol{\omega},\mathbf{z}_{\mathrm{L}}} \gamma_k = \frac{\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_k\right|^2}{\sum_{i\neq k}\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_i\right|^2 + f^2\sigma^2} \\ &\text{s.t.} \quad \Re\left(z_{\mathrm{L},i}\right) \geq 0, \forall i = 1, 2, \ldots, N_{\mathrm{T}} \\ &\left|\boldsymbol{\omega}^T\right| \left|(\mathbf{Z}_{\mathrm{L}} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\omega}\right| \leq P_0 \end{aligned} \tag{26}$$

In this optimization problem, the received SINR is primarily determined by the open-circuit channel matrix $\mathbf{H}_{\mathrm{oc}}$, open-circuit precoding matrix $\mathbf{W}_{\mathrm{oc}}$, noise variance σ^2, and noise amplification factor f. However, once the open-circuit precoder is selected, the only variable that can be adjusted is the noise amplification factor f. Under these conditions, the maximization of the system's received SINR can be equivalently expressed as the minimization of the noise amplification factor f. In this context, (26) can be equivalently represented as

$$\begin{aligned} &\mathcal{P}_2 : \min_{\boldsymbol{\omega},\mathbf{z}_{\mathrm{L}}} f^2 = \left\|\mathrm{diag}\left(\left(\mathbf{Z}_L + \boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\omega}\right)^{-1}\mathbf{W}_{\mathrm{oc}}\right\|_F^2 \\ &\text{s.t.} \quad \Re\left(z_{\mathrm{L},i}\right) \geq 0, \forall i = 1, 2, \ldots, N_{\mathrm{T}} \\ &\left|\boldsymbol{\omega}^T\right| \left|(\mathbf{Z}_{\mathrm{L}} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\omega}\right| \leq P_0 \end{aligned} \tag{27}$$

After obtaining the optimal antenna weighting coefficient $\boldsymbol{\omega}_{\mathrm{opt}}$ and the optimal load impedance $\mathbf{Z}_{\mathrm{L,opt}}$ through optimization, the digital domain precoder can be expressed as

$$\mathbf{P}_{\mathrm{opt}} = \mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L,opt}} + \boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\omega}_{\mathrm{opt}}\right)^{-1}\mathbf{W}_{\mathrm{oc}}. \tag{28}$$

Under normal circumstances, for typical antenna weighting coefficients and load impedance, it is known from (4) to (15) that $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ is non-singular. However, for different precoding schemes, such as the effective channel singular value decomposition based on the water-filling algorithm, the optimized results may lead to some subchannel antenna current coefficients being zero, thus making $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ singular. In this case, since $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ is a diagonal matrix, when $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ is singular, the $\mathbf{D}^{\dagger}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ can be used instead of $\mathbf{D}^{-1}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$. In this situation, the proposed method in this work remains effective. Based on this, the optimal noise amplification factor can be expressed as

$$f_{\mathrm{opt}} = \left\|\mathbf{P}_{\mathrm{opt}}\right\|_F. \tag{29}$$

The received SINR of the system after joint analog–digital precoding can be obtained

$$\gamma_{k,\mathrm{opt}} = \frac{\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_k\right|^2}{\sum_{i\neq k}\left|\mathbf{h}_{\mathrm{oc},k}\mathbf{w}_i\right|^2 + {f_{\mathrm{opt}}}^2\sigma^2}. \tag{30}$$

In the transceiver architecture proposed in this work, we do not further design or impose restrictions on the digital precoder and the receiver. While this may result in a loss of some optimality, it brings many advantages. Firstly, the separation of digital and analog domain transceiver designs facilitates transforming the optimization problem into a convex form, thereby reducing the complexity of precoder design and improving model stability. Additionally, in a non-joint design context, the proposed precoding architecture is a universal model that can be applied to almost all linear transceiver architectures [1, 32]. On the one hand, we can enhance system performance through mutual coupling in the analog domain without disrupting the traditional transceiver architecture, greatly improving the feasibility of the proposed solution. On the other hand, various mature transceiver models already developed can be easily applied to the proposed model for different communication scenarios and application requirements, significantly enhancing the model's flexibility and scalability. Finally, the considerable performance improvement as shown in the numerical result section also highlights the research value of this work.

3.2. Tuning of load impedances and antenna weighting coefficients

The formulation of the problem from (19) to (27) reveals that the proposed scheme requires adaptive tuning of load impedances or antenna weighting coefficients based on channel variations. Therefore, adaptive tuning methods are essential. Firstly, adaptive tuning of antenna weighting coefficients can be achieved through high-speed control by programmable digital chips in the RF front-end. Commercial chips, such as Texas Instruments' AFE80xx RF sampling transceiver chip, are capable of autonomous automatic weighting control over a wide bandwidth. Thus, rapid adaptive tuning of antenna weighting coefficients can be effectively achievable.

Secondly, adaptive tuning of load impedances can be accomplished using variable capacitors. Existing literature indicates that semiconductor-based variable capacitors and ferroelectric-based variable capacitors, which support adaptive impedance tuning, can achieve tuning speeds of up to 1–100 ns [33]. Additionally, the adaptive matching networks based on automatic impedance tuning units proposed in [33,34] can facilitate the implementation of the precoding scheme presented in this study. Furthermore, the application of electronically steerable parasitic array radiators (ESPARs), which require changing the value of each load impedance on a per-symbol basis to form radiation patterns, further supports the implementation of the proposed precoding scheme [35–37].

3.3. Precoding based on load impedances

From the analysis in the first subsection, it can be seen that the mutual coupling precoding scheme proposed in this paper can be translated into an optimization problem for both the antenna weighting coefficients and the load impedances in the analog domain. However, in $\mathcal{P}_2$, due to the constraint $\left|\boldsymbol{\omega}^T\right| \left|(\mathbf{Z}_{\mathrm{L}} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\omega}\right| \leq P_0$, the problem becomes a typical joint non-convex constraint regarding $\boldsymbol{\omega}$ and $\mathbf{Z}_{\mathrm{L}}$ [38]. Therefore, it is challenging to use effective convex optimization toolboxes to jointly optimize the above problem. Therefore, in this subsection and the next subsection we consider the cases of optimizing only the load impedances with the antenna weighting coefficients fixed

and optimizing only the antenna weighting coefficients with the load impedances fixed, respectively. The comparison of these two schemes and the analysis of alternate optimization are discussed in Section 3.5.

In this subsection, we mainly analyze the case with the antenna weighting coefficients fixed and optimizes the load impedances alone. First, we can further simplify the objective function in $\mathcal{P}_2$, which can be expressed as

$$f_0(\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}) = \left\| \mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}+\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\omega}\right)^{-1}\mathbf{W}_{\mathrm{oc}} \right\|_F^2 \tag{31}$$

Combining (4), the above equation can be equated as

$$\begin{aligned} f_0(\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}) &= f_0(\mathbf{c}) = \left\| \mathrm{diag}(\mathbf{c})^{-1}\mathbf{W}_{\mathrm{oc}} \right\|_F^2 \\ &= \mathrm{tr}\left(\mathrm{diag}(\mathbf{c})^{-1}\mathbf{W}_{\mathrm{oc}}\mathbf{W}_{\mathrm{oc}}^H \mathrm{diag}(\mathbf{c})^{-H} \right) \\ &= \mathrm{tr}\left(\mathrm{diag}(\mathbf{c})^{-1}\mathrm{diag}(\mathbf{c})^{-H}\mathbf{W}_{\mathrm{oc}}\mathbf{W}_{\mathrm{oc}}^H \right) \end{aligned} \tag{32}$$

For simplicity, we denote

$$\mathbf{B} = \mathbf{W}_{\mathrm{oc}}\mathbf{W}_{\mathrm{oc}}^H \tag{33a}$$

$$\mathbf{b} = \left[b_1, b_2, \ldots, b_{N_{\mathrm{T}}} \right]^T = \left[\mathbf{B}_{1,1}, \mathbf{B}_{2,2}, \ldots, \mathbf{B}_{N_{\mathrm{T}},N_{\mathrm{T}}} \right]^T \tag{33b}$$

where $\mathbf{B}_{i,i}$ represents the ith diagonal element in matrix $\mathbf{B}$. Substitute (33) into (32), the objective function can be further simplified to

$$f_0(\mathbf{c}) = \sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{c_k c_k^*} = \sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{|c_k|^2}. \tag{34}$$

When the antenna weighting coefficient $\boldsymbol{\omega}_0$ is fixed, the original optimization problem can be simplified as

$$\begin{aligned} \mathcal{P}_3:\ & \min_{\mathbf{Z}_{\mathrm{L}}} \textstyle\sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{|c_k|^2} \\ \text{s.t.}\ & \Re\left(z_{\mathrm{L},k}\right) \geq 0, \forall k = 1,2,\ldots,N_{\mathrm{T}} \\ & \left(\mathbf{Z}_{\mathrm{L}}+\boldsymbol{\Gamma}\right)\mathbf{c} = \boldsymbol{\omega}_0 \\ & \left|\boldsymbol{\omega}_0^T\right| |\mathbf{c}| \leq P_0 \end{aligned} \tag{35}$$

From the above equation, we can see that it is an optimization of the current coefficient $\mathbf{c}$. Once the optimal current coefficient $\mathbf{c}_{\mathrm{opt}}$ is obtained, the load impedance can be solved by $(\mathbf{Z}_{\mathrm{L}}+\boldsymbol{\Gamma})\mathbf{c}_{\mathrm{opt}} = \boldsymbol{\omega}_0$. Based on this, $\mathcal{P}_3$ can be further simplified as

$$\begin{aligned} \mathcal{P}_4:\ & \min_{\mathbf{c}} \textstyle\sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{|c_k|^2} \\ \text{s.t.}\ & \left|\boldsymbol{\omega}_0^T\right| |\mathbf{c}| \leq P_0 \end{aligned} \tag{36}$$

For simplicity, we denote

$$\mathbf{c}_{\mathrm{abs}} = |\mathbf{c}| = \left[c_1^{\mathrm{abs}}, c_2^{\mathrm{abs}}, \ldots, c_{N_{\mathrm{T}}}^{\mathrm{abs}} \right]^T, \tag{37a}$$

$$\boldsymbol{\omega}_{\mathrm{abs}} = |\boldsymbol{\omega}| = \left[\omega_1^{\mathrm{abs}}, \omega_2^{\mathrm{abs}}, \ldots, \omega_{N_{\mathrm{T}}}^{\mathrm{abs}} \right]^T. \tag{37b}$$

At this point, $\mathcal{P}_4$ can be converted to

$$\begin{aligned} \mathcal{P}_5:\ & \min_{\mathbf{c}_{\mathrm{abs}}} \textstyle\sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{\left(c_k^{\mathrm{abs}}\right)^2} \\ \text{s.t.}\ & \boldsymbol{\omega}_{\mathrm{abs}}^T\mathbf{c}_{\mathrm{abs}} \leq P_0 \end{aligned} \tag{38}$$

$\mathcal{P}_5$ is a typical convex optimization problem, and the corresponding Lagrange function can be written as [38]

$$\mathcal{L}(\lambda,\mathbf{c}_{\mathrm{abs}}) = \sum_{k=1}^{N_{\mathrm{T}}} \frac{b_k}{\left(c_k^{\mathrm{abs}}\right)^2} + \lambda\boldsymbol{\omega}_{\mathrm{abs}}^T\mathbf{c}_{\mathrm{abs}} - \lambda P_0 \tag{39}$$

where λ is a dual variable and satisfies $\lambda \geq 0$. Based on (35), the Karush-Kuhn–Tucker (KKT) condition corresponding to the optimization problem can be expressed as

$$\boldsymbol{\omega}_{\mathrm{abs}}^T\mathbf{c}_{\mathrm{abs}} \leq P_0, \tag{40a}$$

$$\lambda \geq 0, \tag{40b}$$

$$\lambda(\boldsymbol{\omega}_{\mathrm{abs}}^T\mathbf{c}_{\mathrm{abs}} - P_0) = 0, \tag{40c}$$

$$\frac{\partial\mathcal{L}(\lambda,\mathbf{c}_{\mathrm{abs}})}{\partial\mathbf{c}_{\mathrm{abs}}} = 0. \tag{40d}$$

Eq. (40d) can be further expanded as

$$\begin{bmatrix} -2b_1/(c_1^{\mathrm{abs}})^3 \\ -2b_2/(c_2^{\mathrm{abs}})^3 \\ \vdots \\ -2b_{N_{\mathrm{T}}}/(c_{N_{\mathrm{T}}}^{\mathrm{abs}})^3 \end{bmatrix} + \lambda \begin{bmatrix} \omega_1^{\mathrm{abs}} \\ \omega_2^{\mathrm{abs}} \\ \vdots \\ \omega_{N_{\mathrm{T}}}^{\mathrm{abs}} \end{bmatrix} = \mathbf{0}. \tag{41}$$

Based on (41), we can obtain the relationship between $\mathbf{c}_{\mathrm{opt}}$ and λ as

$$\lambda\omega_k^{\mathrm{abs}} = \frac{2b_k}{\left(c_k^{\mathrm{abs}}\right)^3} \quad \forall k = 1,2,\ldots,N_{\mathrm{T}}. \tag{42}$$

In the case of realistic constraints, c_k^{abs}, ω_k^{abs} and b_k in (42) are all greater than 0, so we can get $\lambda > 0$. At this point, we can obtain

$$c_k^{\mathrm{abs}} = \left(\frac{2b_k}{\lambda\omega_k^{\mathrm{abs}}} \right)^{1/3}. \tag{43}$$

For the convex optimization $\mathcal{P}_5$ in (38), it is easy to verify that the Slater's condition is satisfied [38], which means that the dual gap is zero. Therefore, we can solve $\mathcal{P}_5$ by solving its corresponding dual problem, which can be obtained

$$\begin{aligned} \mathcal{P}_6:\ & \max \lambda^{2/3}\left[\left(2^{-2/3}+2^{1/3}\right) \sum_{k=1}^{N_{\mathrm{T}}} b_k^{1/3}\left(\omega_k^{\mathrm{abs}}\right)^{2/3} \right] - \lambda P_0 \\ \text{s.t.}\ & \lambda > 0 \end{aligned} \tag{44}$$

For simplicity, we denote $a = \left(2^{-2/3}+2^{1/3}\right)\sum_{k=1}^{N_{\mathrm{T}}} b_k^{1/3}\left(\omega_k^{\mathrm{abs}}\right)^{2/3}$. At this point, $\mathcal{P}_6$ can be expressed as

$$\begin{aligned} \mathcal{P}_7:\ & \max L(\lambda) = a\lambda^{2/3} - P_0\lambda \\ \text{s.t.}\ & \lambda > 0 \end{aligned} \tag{45}$$

It can be found that $\mathcal{P}_7$ is a very simple convex optimization problem, which can be easily solved by solving the point where the gradient of the objective function w.r.t. λ is equal to 0

$$\begin{aligned} & \tfrac{\partial L(\lambda)}{\partial\lambda} = \tfrac{2}{3}a\lambda^{-1/3} - P_0 = 0 \\ & \Rightarrow \lambda = \left(\tfrac{2a}{3P_0}\right)^3 \end{aligned} \tag{46}$$

By substituting (46) into (43), the closed-form solution of the optimal current coefficient vector for problem $\mathcal{P}_5$ can be obtained

$$c_k^{\mathrm{abs}} = \frac{3P_0}{2a}\left(\frac{2b_k}{\omega_k^{\mathrm{abs}}} \right)^{1/3} \quad \forall k = 1,2,\ldots,N_{\mathrm{T}}. \tag{47}$$

It can be seen that the minimum value of the objective function is only related to the modal values of the current coefficients when the antenna weighting coefficients (modal value) are given. Furthermore, (47) gives a closed-form solution for the optimal current coefficient modulus value at a fixed antenna weighting coefficient, which is independent of the phase of the current coefficient. This result provides a large degree of freedom for the actual system, which is free to change the phase according to the demand, and the performance of the proposed mutual coupling precoding scheme is not affected by it.

3.4. Precoding based on antenna weighting coefficients

In this subsection, we mainly consider the case where the fixed load impedances only optimizes the antenna weighting coefficients. Based on the analysis in the third subsection, the optimization problem when the fixed load impedances only optimizes the antenna weighting coefficients can be expressed as

$$\begin{aligned} \mathcal{P}_8:\ & \min_{\boldsymbol{\omega}} \left\| \mathrm{diag}\left(\left(\mathbf{Z}_{\mathrm{L}}^0+\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\omega}\right)^{-1}\mathbf{W}_{\mathrm{oc}} \right\|_F^2 \\ \text{s.t.}\ & \left|\boldsymbol{\omega}^T\right| \left| \left(\mathbf{Z}_{\mathrm{L}}^0+\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\omega} \right| \leq P_0 \end{aligned} \tag{48}$$

where $\mathbf{Z}_{\mathrm{L}}^0$ is the fixed initial load impedance diagonal matrix. In $\mathcal{P}_8$, it can be easily verified that when the objective function is minimized, the inequality constraints will become equality constraints. Based on

Algorithm 1 Newton method for Solving $\mathcal{P}_9$.

INPUT: $\mathbf{W}_{\rm oc}$, $\left(\mathbf{Z}_{\rm L}^0+\mathbf{\Gamma}\right)^{-1}$, $\boldsymbol{\omega}^0$, λ^0, $N_{\max}$, δ_{th}, α_{th}, $infeb_{th}$
OUTPUT: $\boldsymbol{\omega}_{\rm opt}$
initialization : $k=0, \boldsymbol{\omega}^{(0)}=\boldsymbol{\omega}^0$
while $k \le N_{\max}$, $\delta \ge \delta_{th}$ and $\alpha > \alpha_{th}$ or $infeb > infeb_{th}$ **do**
obtain $\boldsymbol{\omega}^{(k+1)}$ and $\lambda^{(k+1)}$ by (52)-(53)
$\delta = \left\|\boldsymbol{\omega}^{(k+1)}-\boldsymbol{\omega}^{(k)}\right\|$
$\alpha = \left\|\nabla f(\boldsymbol{\omega}^{(k+1)}) - (\nabla h(\boldsymbol{\omega}^{(k+1)}))^T \lambda^{(k+1)}\right\|_\infty$
$infeb = \left\|h(\boldsymbol{\omega}^{(k+1)})\right\|_\infty$
$k = k+1$
end while
$\boldsymbol{\omega}_{\rm opt} = \boldsymbol{\omega}^k$

this, $\mathcal{P}_8$ can be equivalently rewritten as

$$\begin{aligned} \mathcal{P}_9 :\quad & \min_{\boldsymbol{\omega}} \left\|\mathrm{diag}\left(\left(\mathbf{Z}_{\rm L}^0+\mathbf{\Gamma}\right)^{-1}\boldsymbol{\omega}\right)^{-1}\mathbf{W}_{\rm oc}\right\|_F^2 \\ \text{s.t.}\quad & \left|\boldsymbol{\omega}^T\right|\left|\left(\mathbf{Z}_{\rm L}^0+\mathbf{\Gamma}\right)^{-1}\boldsymbol{\omega}\right| = P_0 \end{aligned} \tag{49}$$

It can be found that $\mathcal{P}_9$ is a relatively complex optimization problem, nevertheless, we can use the well-established Newton method [38–40] in optimization theory to solve this problem. For simplicity, we note that the objective function in $\mathcal{P}_9$ is

$$f_{0,\mathcal{P}_9}(\boldsymbol{\omega}) = \left\|\mathrm{diag}\left(\left(\mathbf{Z}_{\rm L}^0+\mathbf{\Gamma}\right)^{-1}\boldsymbol{\omega}\right)^{-1}\mathbf{W}_{\rm oc}\right\|_F^2$$

and the equality constraints is $h(\boldsymbol{\omega}) = \left|\boldsymbol{\omega}^T\right|\left|\left(\mathbf{Z}_{\rm L}^0+\mathbf{\Gamma}\right)^{-1}\boldsymbol{\omega}\right| - P_0 = 0$. Then, $\mathcal{P}_9$ can be rewritten as

$$\begin{aligned} \mathcal{P}_{10} :\quad & \min_{\boldsymbol{\omega}} f_{0,\mathcal{P}_9}(\boldsymbol{\omega}) \\ \text{s.t. } & h(\boldsymbol{\omega}) = 0 \end{aligned} \tag{50}$$

At this point, $\mathcal{P}_{10}$ can be transformed into the unconstrained optimization problem

$$\mathcal{P}_{11} : \min_{\boldsymbol{\omega},\lambda} \mathcal{L}(\boldsymbol{\omega},\lambda) = f_{0,\mathcal{P}_9}(\boldsymbol{\omega}) - \lambda h(\boldsymbol{\omega}) \tag{51}$$

where $\mathcal{L}(\boldsymbol{\omega},\lambda)$ is the Lagrangian objective function and λ is the Lagrangian multiplier. Based on this, the Newton equation for problem $\mathcal{P}_{11}$ can be obtained as [40]

$$\begin{bmatrix} -\mathbf{H}(\boldsymbol{\omega},\lambda) & (\nabla h(\boldsymbol{\omega}))^T \\ \nabla h(\boldsymbol{\omega}) & 0 \end{bmatrix}\begin{bmatrix} \Delta\boldsymbol{\omega} \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} \nabla f_{0,\mathcal{P}_9}(\boldsymbol{\omega}) - (\nabla h(\boldsymbol{\omega}))^T\lambda \\ -h(\boldsymbol{\omega}) \end{bmatrix} \tag{52}$$

where ∇ denotes the gradient operator; $\mathbf{H}(\boldsymbol{\omega},\lambda) = \nabla^2 f_{0,\mathcal{P}_9}(\boldsymbol{\omega}) - \lambda\nabla^2 h(\boldsymbol{\omega})$; $\Delta\boldsymbol{\omega}$ is the optimization direction; $\Delta\lambda$ is the feasibility direction. Denote $\alpha = \left\|\nabla f(\boldsymbol{\omega}) - (\nabla h(\boldsymbol{\omega}))^T\lambda\right\|_\infty$ and $infeb = \|h(\boldsymbol{\omega})\|_\infty$, which measure the first-order optimality and original infeasibility of the optimization problem $\mathcal{P}_9$, respectively. $(\Delta\boldsymbol{\omega},\Delta\lambda)$ constitutes the updated Newton step of the Newton method, i.e., the search direction.

The Newton equation can be solved iteratively by preprocessing the conjugate gradient method [41] in a relatively simple way, and then the Newton step can be obtained. In this case, the feasible points can be iteratively updated by [42]

$$\left.\begin{aligned} \boldsymbol{\omega}^{(k+1)} &= \boldsymbol{\omega}^{(k)} + \eta\Delta\boldsymbol{\omega}^{(k)} \\ \lambda^{(k+1)} &= \lambda^{(k)} + \eta\Delta\lambda^{(k)} \end{aligned}\right\} \tag{53}$$

where η is the update step. The Newton method for solving $\mathcal{P}_9$ is summarized in Algorithm 1. In Algorithm 1, $\boldsymbol{\omega}^0$ is the initial antenna weighting coefficients; $N_{\max}$ is the maximum number of iterations; α_{th} is the first-order optimality accuracy; $infeb_{th}$ is the infeasibility accuracy and δ_{th} is a variable representing convergence accuracy, defined as $\delta_{th} = \left\|\boldsymbol{\omega}^{(k+1)}-\boldsymbol{\omega}^{(k)}\right\|$.

Fig. 4 shows the convergence of Algorithm 1 in a channel drop, where the initial antenna weighting coefficient is set to $\omega_i^{\rm opt} = 10$ volts $\forall i = 1,2,\ldots,N_{\rm T}$. It can be seen from Fig. 4 that Algorithm 1 has a good convergence speed, and f has basically converged within 12 iterations.

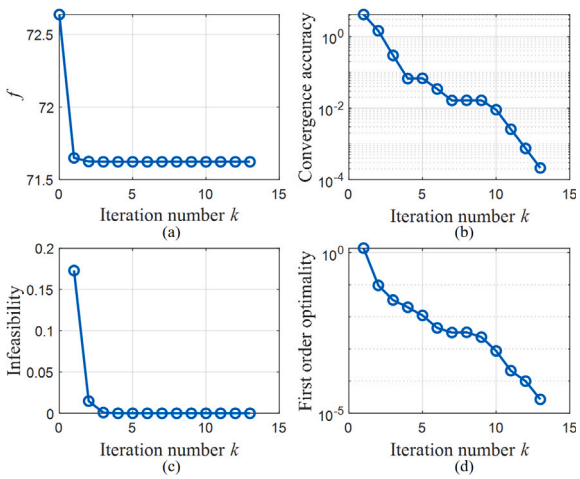

Fig. 4. Convergence speed of the Newton method in Algorithm 1, $N_{\rm T} = N_{\rm R} = 4$, $N_{\max} = 20$, $\delta_{th} = 10^{-3}$, $infeb_{th} = 0$, $\alpha_{th} = 10^{-4}$. (a) Value of f. (b) Value of δ. (c) Value of $infeb$. (d) Value of α.

3.5. Relationship between precoding based on antenna weighting coefficients and load impedances

Although the optimization problem $\mathcal{P}_2$ has two optimization variables: load impedance $\mathbf{Z}_{\rm L}$ and antenna weighting coefficient $\boldsymbol{\omega}$, $\mathbf{Z}_{\rm L}$ and $\boldsymbol{\omega}$ satisfy the linear relationship shown in (4). For the optimal current coefficient solution $\mathbf{c}_{\rm opt}$ of $\mathcal{P}_2$, $\mathbf{Z}_{\rm L}$ and $\boldsymbol{\omega}$ can be expressed as each other

$$\omega_m = z_{{\rm L},m}c_{{\rm opt},m} + \sum_{n=1}^{N_{\rm T}}\mathbf{\Gamma}_{m,n}c_{{\rm opt},n} \quad m = 1,2,\ldots N_{\rm T} \tag{54a}$$

$$z_{{\rm L},m} = \frac{\omega_m - \sum_{n=1}^{N_{\rm T}}\mathbf{\Gamma}_{m,n}c_{{\rm opt},n}}{c_{{\rm opt},m}} \quad m = 1,2,\ldots N_{\rm T} \tag{54b}$$

where $\mathbf{\Gamma}_{m,n}$ denotes the mth row and nth column element of the mutual impedance matrix. According to (34), it can be found that the minimum value of the objective function of $\mathcal{P}_2$ is mainly related to the mode value $\mathbf{c}_{\rm opt}^{\rm abs}$ of the optimal current coefficient. On this basis, (54) implies that when there are no constraints on the range of load impedance and antenna weighting coefficient, theoretically, optimizing any one of them alone can yield the optimal current coefficient solution. However, in practice, the real part of all the adjustable loads should satisfy the constraint $\Re\left(z_{{\rm L},i}\right) \ge 0$, $\forall i = 1,2,\ldots,N_{\rm T}$. This constraint significantly limits the feasible domain of the magnitude of the optimal current coefficient when only optimizing the load impedances with fixed antenna weighting coefficients

$$\mathbf{c}_{\rm opt}^{\rm abs} \in \left(0, \left|\mathbf{\Gamma}^{-1}\boldsymbol{\omega}_0\right|\right], \text{when } \Re\left(\mathbf{Z}_{\rm L}\right) \in [0,+\infty). \tag{55}$$

For the case of fixed load impedance $\mathbf{Z}_{\rm L}^0$ optimizing only the antenna weighting coefficient $\boldsymbol{\omega}$, the feasible domain of the magnitude of the optimal current coefficient is not restricted

$$\mathbf{c}_{\rm opt}^{\rm abs} \in [0,+\infty), \text{ when } |\boldsymbol{\omega}| \in [0,+\infty). \tag{56}$$

It can be seen from (55) and (56) that because there are no constraints on the range of antenna weighting coefficient $\boldsymbol{\omega}$, the feasible domain of the corresponding magnitude of the current coefficient is also unconstrained. Theoretically, optimizing only the antenna weighting coefficient $\boldsymbol{\omega}$ while keeping the load impedance $\mathbf{Z}_{\rm L}$ fixed can yield the optimal current coefficient solution without the need for iterative optimization. However, in the case of load impedance, as it is subject to certain limitations, fixing the antenna weighting coefficient and optimizing only the load impedance may not necessarily lead to a globally optimal current coefficient solution. It may only result in

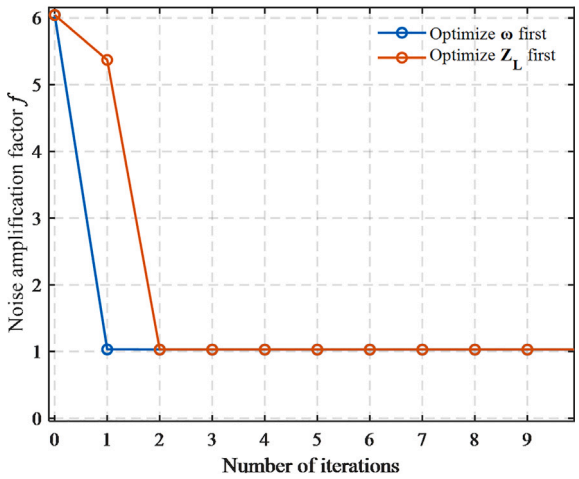

Fig. 5. Curve of noise amplification factor versus number of iterations for alternate optimization of load impedance and antenna weighting coefficient.

a locally optimal solution within the corresponding feasible interval. Thus optimizing only the load impedance with fixed antenna weighting coefficient is inferior to optimizing only the antenna weighting coefficient with fixed load impedance.

Fig. 5 illustrates the variation of the noise amplification factor f with the number of iterations during the iterative optimization of load impedance $\mathbf{Z}_{\mathrm{L}}$ and antenna weighting coefficient $\boldsymbol{\omega}$. In Fig. 5, the blue line represents the iterative process of optimizing $\boldsymbol{\omega}$ first and then $\mathbf{Z}_{\mathrm{L}}$, while the red line represents the iterative process of optimizing $\mathbf{Z}_{\mathrm{L}}$ first and then $\boldsymbol{\omega}$.

From Fig. 5, it can be observed that optimizing only the antenna weighting coefficient $\boldsymbol{\omega}$ yields the optimal results. In the proposed scheme, joint optimization is unnecessary. Furthermore, it can also be seen from Fig. 5 that optimizing only the load impedance $\mathbf{Z}_{\mathrm{L}}$ may not necessarily result in the optimal outcomes due to the constraints in its current coefficient feasibility domain. However, studying the optimization of load impedance is still necessary and meaningful. On the one hand, for scenarios with fixed antenna weighting coefficients, optimizing only the load impedance can yield a closed-form solution, which is crucial in limited computational resources environments. On the other hand, as shown in the numerical results section, when the antenna spacing is less than half a wavelength, the BER performance of the system with only optimized load impedance is still better than that of a system without mutual coupling, indicating that significant performance improvement can still be achieved by optimizing the load impedance alone.

3.6. Extraction of $\mathbf{H}_{\mathrm{oc}}$

From (19)–(27), we can find that the transmitter needs to obtain the open-circuit channel matrix $\mathbf{H}_{\mathrm{oc}}$ for equalization. However, based on the pilot we can only obtain the $\mathbf{H}_{\mathrm{oc}}\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$, so we need to extract $\mathbf{H}_{\mathrm{oc}}$ from $\mathbf{H}_{\mathrm{oc}}\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$. Note that the load impedance $\mathbf{Z}_{\mathrm{L}}$ and the antenna weighting coefficient $\boldsymbol{\omega}$ in $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ are known to the transmitter. The mutual impedance matrix $\boldsymbol{\Gamma}$ is usually known to the transmitter by the induced electric potential method or other experimental measurements (e.g., edge value method and transmission line method [22]), and it depends only on the array structure and does not vary with the load impedance $\mathbf{Z}_{\mathrm{L}}$ and antenna weighting coefficient $\boldsymbol{\omega}$. Therefore, for a closed-loop MIMO system, the coupling matrix $\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$ is known to the transmitter, at which point the transmitter can easily extract $\mathbf{H}_{\mathrm{oc}}$ from the $\mathbf{H}_{\mathrm{oc}}\mathbf{D}_{\mathbf{Z}_{\mathrm{L}},\boldsymbol{\omega}}$. In this case the scheme proposed in this section can be easily applied.

Table 1
Abbreviations of different schemes.

Abbreviations	Schemes
ZF x λ	Zero-forcing (ZF) precoding with x wavelength transmit antenna spacing
ZF with MC	ZF precoding with mutual coupling
ZF no MC	ZF precoding without mutual coupling
ZF AD-ω	The proposed precoding based on antenna weighting coefficient with ZF
ZF AD-$\mathbf{Z}_{\mathrm{L}}$	The proposed precoding based on load impedance with ZF
MMSE x λ	Minimum mean square error (MMSE) precoding with x wavelength transmit antenna spacing
MMSE with MC	MMSE precoding with mutual coupling
MMSE no MC	MMSE precoding without mutual coupling
MMSE AD-ω	The proposed precoding based on antenna weighting coefficient with MMSE
MMSE AD-$\mathbf{Z}_{\mathrm{L}}$	The proposed precoding based on load impedance with MMSE

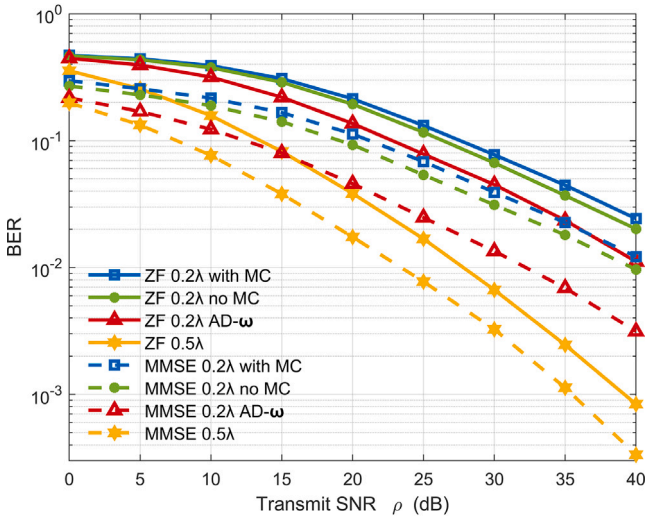

Fig. 6. BER performances of 4 × 4 MIMO system with different schemes as a function of transmit SNR.

4. Numerical results

To evaluate the performance of the proposed mutual coupling precoding scheme, this section conducts Monte Carlo experiments based on the aforementioned theoretical model and presents the numerical simulation results. In the simulation process, we assume that the system operates at 2.6 GHz with QPSK modulation, and the channel model considering the actual antenna radiation performance follows (6)–(15). In order to simulate a more realistic channel environment, the standardized channel model 3GPP 38.901 with non-line-of-sight (NLOS) urban microcell (Umi) is chosen as the specific channel simulation scenario, and its specific channel parameters can be found in [43].

Since this work focuses on the radiation performance of the BS, the UE is assumed to be isotropic antennas in the simulation. The BS is a ULA consisting of half-wavelength dipoles, and the height of the BS is set to 25 m. The height of the UE is set to 1.5 m, and the UE is randomly distributed in a circle centered on the BS with the distance from the BS as [25, 200] m. In the simulation, initial load impedance $z_{\mathrm{L},i}^{0} = 50\ \Omega$, initial antenna weighting coefficient $\omega_i^0 = 10$ V, $\forall i = 1, 2, \ldots, N_{\mathrm{T}}$, and the apparent power of the baseline group is equally distributed among all RF chains. $\rho = 1/\sigma^2$ is denoted as the transmit SNR and the total apparent power of BS system is $P_0 = 10N_{\mathrm{s}}$ VA, where N_{s} is the number of data streams transmitted. 20,000 channel drops were sampled in the numerical simulation. For the sake of clarity, the abbreviations used throughout the simulation are summarized in Table 1.

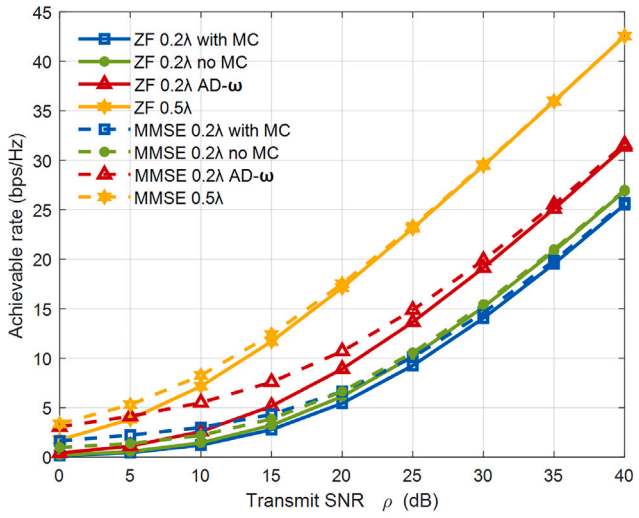

Fig. 7. Achievable rate performances of 4 × 4 MIMO system for different schemes.

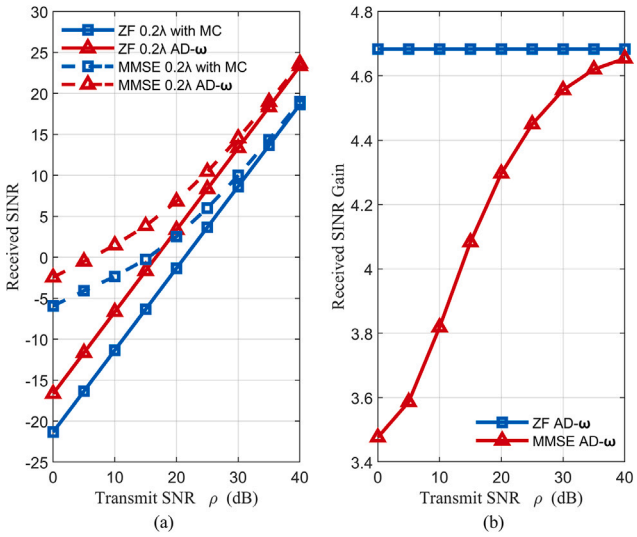

Fig. 8. (a) Received SINR and corresponding, (b) SINR gain for ZF-based and MMSE-based mutual coupling precoders.

Fig. 6 shows the BER performances of the system with different schemes. As can be seen in Fig. 6, for ZF and minimum mean square error (MMSE) precoding, the conventional scheme with fixed mutual coupling has the worst BER performance, whereas the proposed mutual coupling precoding scheme achieves the best BER performance when the antenna spacing is $0.2\lambda_c$. Because the proposed scheme exploits mutual coupling, it can be seen from Fig. 6 that the BER performance of the scheme is better than the case without mutual coupling at the same antenna spacing. In addition, the simulation results in Fig. 6 show that the BER performance of the system at antenna spacing of $0.2\lambda_c$ cannot exceed that of the conventional method at an antenna spacing of $0.5\lambda_c$ for the same number of antennas, even if mutual coupling between antennas is utilized. This is mainly because the antenna array with large spacing has a larger aperture (and thus larger spatial degrees of freedom) when the antenna number is the same. However, with the same antenna spacing, the proposed mutual coupling precoding still achieves the optimal performance over the precoding schemes without exploiting mutual coupling, and it improves the SNR by more than 7 dB over the conventional method under the MMSE scheme when the BER is at the 10^{-2} level.

Besides, the achievable rate of the system can be expressed as

$$R = \sum_{i=1}^{N_T} \log_2\left(1+\gamma_i\right) \tag{57}$$

where γ_i is the received SINR of the ith data stream. γ_i can be obtained by (22) for different precoding schemes.

Fig. 7 shows the achievable rate performances of the system with different schemes. We can draw a similar conclusion that the proposed mutual coupling precoding achieves the optimal achievable rate over the precoding schemes without exploiting mutual coupling for the same antenna spacing, since the method exploits the mutual coupling between the antennas. At a transmit SNR of 30 dB, the proposed scheme achieves 5 bps/Hz improvement in the achievable rate over the conventional scheme.

Furthermore, as can be seen in Figs. 6 and 7, the BER performance of the MMSE-based mutual coupling precoding is better than that of the ZF-based mutual coupling precoding under the same circumstances, especially at low SNR level. This is mainly related to the relationship between the received SINR gain and the transmitted SNR gain of different schemes. Fig. 8 present the received SINR and the corresponding SINR gain for the ZF-based and MMSE-based mutual coupling precoders. As depicted, for MMSE precoding, according to Eq. (22), at low transmit SNR levels, the SINR gain from the MMSE-based mutual coupling precoding scheme is lower than that achieved with the ZF-based precoder due to the presence of inter-stream interference. As the transmit SNR increases, the SINR gains of these two approaches gradually converge. It is noteworthy that Figs. 6 and 7 show the BER and achievable rate as functions of the transmit SNR. However, as observed from Fig. 8(a), unlike the ZF precoding, the improvement in received SINR is not linearly equivalent to the increase in transmit SNR for MMSE precoding. For instance, at low transmit SNR, although the actual SINR improvement of the MMSE scheme is less than that of the ZF scheme, the corresponding increase in transmit SNR is greater than the MMSE scheme. This implies that although the MMSE scheme's improvement in received SINR is less than that of the ZF scheme, due to the non-linear relationship between received SINR and transmit SNR, the performance gain exhibited by the MMSE scheme is still better than that of the ZF scheme when plotting BER and achievable rate curves with transmit SNR as the horizontal axis.

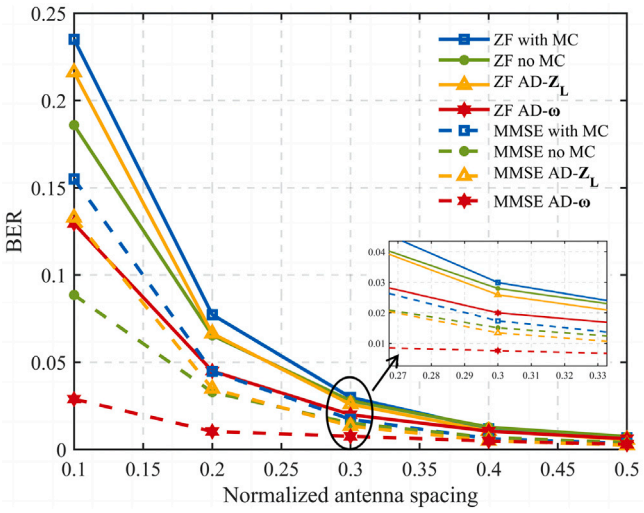

Fig. 9. BER performances for different antenna spacings in 4 × 4 MIMO system with 30-dB transmit SNR.

To evaluate the performance of the proposed precoding scheme for different mutual coupling cases, Fig. 9 shows the BER performance of the proposed precoding for different antenna spacings when the transmit SNR is 30 dB.

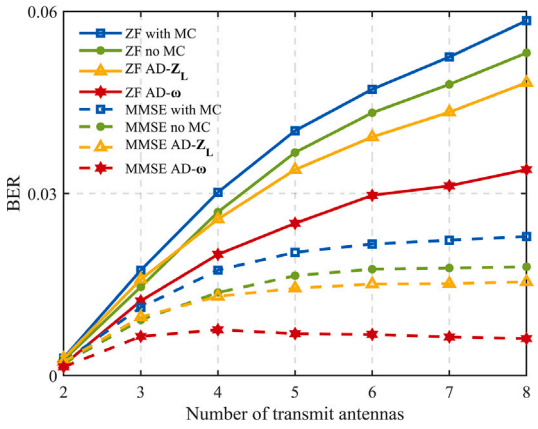

Fig. 10. BER performances for different number of transmit antennas in $N_T \times N_T$ MIMO system with $0.3\lambda_c$ antenna spacing and 30-dB transmit SNR.

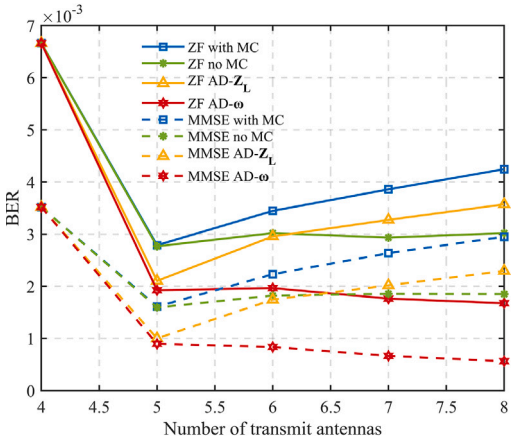

Fig. 11. BER for different number of transmit antennas in $N_T \times 4$ MIMO system with constant 1.5 wavelength aperture, QPSK modulation, transmit SNR $\rho = 30$ dB.

As can be seen in Fig. 9, when the antenna spacing is small, the mutual coupling between the antennas becomes more severe, and the BER performance of the conventional method with mutual coupling is much worse than the other schemes at this time. For the proposed scheme, it can be observed that the best BER performance is achieved by the antenna weighting coefficient-based precoding, which is significantly better than the conventional scheme with mutual coupling or even without mutual coupling, because it exploits the mutual coupling effect to further improve the system performance. The BER performance of the load impedance-based precoding is significantly better than that of the conventional scheme with mutual coupling, and it also outperforms the scheme without mutual coupling when the antenna spacing is greater than $0.2\lambda_c$. While, for antenna spacing less than $0.2\lambda_c$, the load impedance-based precoding performance is worse than that of the scheme without MC. This is mainly because the antenna efficiency degrades drastically at such small antenna separation due to the absorption loss caused by the mutual coupling [10,11], and this adverse effect surpasses the benefit of exploiting the mutual coupling. Furthermore, it is observed that the BER performance of the proposed precoding based on antenna weighting coefficient-based improves as the antenna spacing decreases, which indicates that the proposed scheme is particularly suitable in compact arrays where mutual coupling effects are severe.

Fig. 10 shows the BER performance of the proposed precoding for different numbers of transmit antennas when the transmit SNR is 30 dB, the transmit antenna spacing is $0.3\lambda_c$ and the number of receive antennas is equal to the number of transmit antennas. The system is assumed to operate in full spatial multiplexing mode.

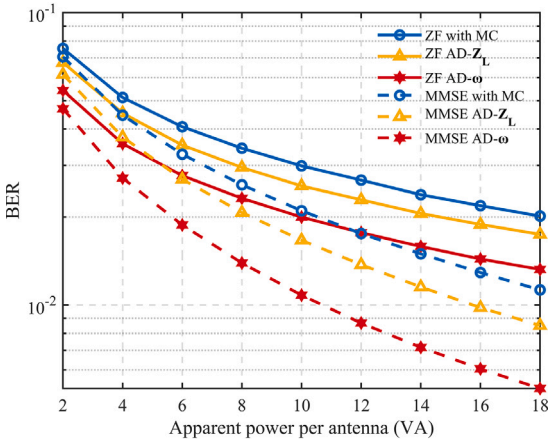

Fig. 12. BER for different apparent power (per antenna) in 4 × 4 MIMO system with QPSK modulation, antenna spacing 0.3 wavelength, transmit SNR $\rho = 30$ dB.

The simulation results in Fig. 10 show that BER performance of each scheme degrades as the number of antennas increases, but consistent performance gains can be observed in all the proposed schemes compared with traditional schemes with or without mutual coupling. Especially, the proposed precoding scheme based on antenna weighting coefficients achieves the optimal BER performance over all simulated precoding schemes. Moreover, the BER performance of the proposed scheme improves with the increase of the number of antennas, which means that the scheme proposed in this paper is especially suitable for higher dimensional MIMO systems.

Fig. 11 shows the BER performances corresponding to different numbers of transmitting antennas when the aperture of the ULA is fixed at 1.5 wavelength. It can be observed that in the case of fixed aperture, BER of the traditional ZF/MMSE precoding will first decrease and then increase with the increase of the number of transmitting antennas. This is mainly due to the fact that the mutual coupling between the antennas also becomes severe due to the increase in the number of transmitting antennas. The proposed antenna weighting coefficient-based mutual coupling precoding scheme still achieves the optimal performance over all simulated precoding schemes.

Fig. 12 shows the BER performance for a 4 × 4 MIMO system with different apparent power (per antenna) when the antenna spacing is $0.3\lambda_c$ and the transmit SNR is 30 dB. As can be seen from Fig. 12, BER performances of all schemes will be better with the increase of apparent power, and it can be obviously observed that the proposed mutual coupling precoding scheme still achieves the optimal BER performance for all apparent power cases. Moreover, Fig. 12 shows that the BER performance gain of the proposed scheme is greater as the apparent power increases, which implies that the proposed scheme will have better performance for BS with large apparent power.

5. Conclusion

In this paper, an analog–digital precoding scheme was proposed to exploit the mutual coupling effect between antennas with consideration of the actual radiation performance of the MIMO antenna array. Open-circuit radiation patterns and coupling matrix are used to accurately evaluate the mutual coupling effects, so that the antenna model is no longer required to satisfy the minimum scattering condition, but is applicable for arbitrary antenna. In the digital domain the proposed scheme can use conventional linear precoding, and in the analog domain it uses convex communication system. The differences and relations between load impedance-based precoding and antenna

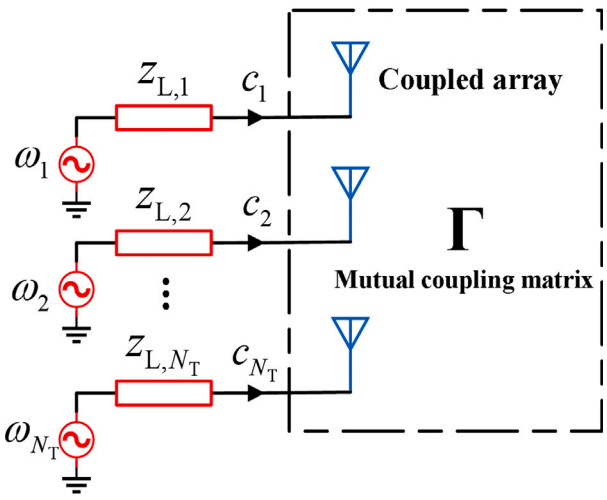

Fig. A.13. Equivalent impedance network of N_T-port antenna system.

weighting coefficient-based precoding were discussed, and the closed-form solution of excitation current coefficient for load impedance-based precoding was given. The mathematical analysis finds that the load impedance-based precoding is a suboptimal case. Simulation results show that the proposed antenna weighting coefficient-based precoding achieves the optimal performance, and the performance gain of this scheme improves as the antenna spacing decreases, the number of antennas increases, or the apparent power increases.

CRediT authorship contribution statement

Jianchuan Wei: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation. **Xiaoming Chen:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Ruihai Chen:** Writing – review & editing, Validation, Formal analysis. **Chongwen Huang:** Writing – review & editing, Supervision, Conceptualization. **Xiaoyu Huang:** Writing – review & editing, Methodology, Formal analysis. **Wei E.I. Sha:** Writing – review & editing, Methodology, Conceptualization. **Mérouane Debbah:** Writing – review & editing, Methodology, Conceptualization.

Acknowledgment

This work was supported by Key Research and Development Program of Shaanxi (Program No. 2024GX-ZDCYL-01-26).

Appendix. Derivation of current coefficient

For an antenna system with N_T elements, its circuit characteristics can be equivalently represented by an impedance network, as shown in Fig. A.13. In this impedance network, ω_i denotes the equivalent antenna weighting coefficient of the ith port, $z_{L,i}$ represents the adjustable load impedance at this port, and $\mathbf{\Gamma}$ is the mutual impedance matrix of the network. The relationship between the antenna weighting coefficient ω_i and the current coefficient c_i in the impedance network is given by [25–27]

$$\omega_i = z_{L,i}c_i + \sum_{j=1}^{N_T} [\mathbf{\Gamma}]_{i,j}c_j. \tag{A.1}$$

The equivalent matrix form is

$$\boldsymbol{\omega} = (\mathbf{Z}_L + \mathbf{\Gamma})\,\mathbf{c} \tag{A.2}$$

On this basis, the current coefficient can be expressed as

$$\mathbf{c} = (\mathbf{Z}_L + \mathbf{\Gamma})^{-1}\boldsymbol{\omega}. \tag{A.3}$$

Data availability

No data was used for the research described in the article.

References

[1] A.J. Paulraj, R.U. Nabar, D.A. Gore, Introduction to Space-Time Wireless Communications, Cambridge University Press, 2003.
[2] R. He, B. Ai, G.L. Stuber, et al., Geometrical based modeling for millimeter wave MIMO mobile-to-mobile channels, IEEE Trans. Veh. Technol. 67 (4) (2018) 2848–2863.
[3] A. Zaidi, F. Athley, J. Medbo, et al., 5G Physical Layer: Principles, Models and Technology Components, Academic Press, 2018.
[4] J. Hoydis, S. Ten Brink, M. Debbah, Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? IEEE J. Sel. Areas Commun. 31 (2) (2013) 160–171.
[5] L. Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin, R. Zhang, An overview of massive MIMO: Benefits and challenges, IEEE J. Sel. Top. Signal Process. 8 (5) (2014) 742–758.
[6] A. Pizzo, L. Sanguinetti, T.L. Marzetta, Holographic MIMO communications, 2021, arXiv preprint arXiv:2105.01535.
[7] A. Pizzo, T.L. Marzetta, L. Sanguinetti, Spatially-stationary model for holographic MIMO small-scale fading, IEEE J. Sel. Areas Commun. 38 (9) (2020) 1964–1979.
[8] C. Huang, S. Hu, G.C. Alexandropoulos, A. Zappone, M. Debbah, Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends, IEEE Wirel. Commun. 27 (5) (2020) 118–125.
[9] Christos Masouros, Mathini Sellathurai, Tharm Ratnarajah, Large-scale MIMO transmitters in fixed physical spaces: The effect of transmit correlation and mutual coupling, IEEE Trans. Commun. 61 (7) (2013) 2794–2804.
[10] X. Chen, S. Zhang, Q. Li, A review of mutual coupling in MIMO systems, IEEE Access 6 (2018) 24706–24719.
[11] S.S.A. Yuan, X. Chen, C. Huang, W.E.I. Sha, Effects of mutual coupling on degree of freedom and antenna efficiency in holographic MIMO communications, IEEE Open J. Antennas Propag. 4 (2023) 237–244.
[12] D. Gao, Z.X. Cao, S.D. Fu, X. Quan, P. Chen, A novel slot-array defected ground structure for decoupling microstrip antenna array, IEEE Trans. Antennas and Propagation 68 (10) (2020) 7027–7038.
[13] Y. Da, X. Chen, A.A. Kishk, In-band mutual coupling suppression in dual-band shared-aperture base station arrays using dielectric block loading, IEEE Trans. Antennas and Propagation 70 (10) (2022) 9270–9281.
[14] L. Zhao, Y. He, G. Zhao, et al., Scanning angle extension of a millimeter wave antenna array using electromagnetic band gap ground, IEEE Trans. Antennas and Propagation 70 (8) (2022) 7264–7269.
[15] L. Sun, Y. Li, Z. Zhang, Z. Feng, Wideband 5G MIMO antenna with integrated orthogonal-mode dual-antenna pairs for metal-rimmed smartphones, IEEE Trans. Antennas and Propagation 68 (4) (2019) 2494–2503.
[16] A. Li, C. Masouros, Exploiting mutual coupling by means of analog-digital zero forcing, in: IEEE International Conference on Acoustics, 2017, pp. 3564–3568.
[17] A. Li, C. Masouros, Mutual coupling exploitation for point-to-point MIMO by constructive interference, in: IEEE International Conference on Communications, 2017, pp. 1–6.
[18] A. Li, C. Masouros, Exploiting constructive mutual coupling in P2P MIMO by analog-digital phase alignment, IEEE Trans. Wirel. Commun. 16 (3) (2017) 1948–1962.
[19] A. Li, C. Masouros, M. Sellathurai, Analog-digital beamforming in the MU-MISO downlink by use of tunable antenna loads, IEEE Trans. Veh. Technol. 67 (4) (2018) 3114–3129.
[20] Clerckx, et al., Impact of antenna coupling on 2 ×2 MIMO communications, IEEE Trans. Veh. Technol. 56 (3) (2007) 1009–1018.
[21] Y. Ji, W. Fan, P. Kyosti, J. Li, G.F. Pedersen, Antenna correlation under geometry-based stochastic channel models, IEEE Antennas Wirel. Propag. Lett. 18 (12) (2019) 2567–2571.
[22] C.A. Balanis, Antenna Theory: Analysis and Design, fourth ed., Wiley, Hoboken, NJ, USA, 2016.
[23] D.R. Fuhrmann, J.P. Browning, M. Rangaswamy, Signaling strategies for the hybrid MIMO phased-array radar, IEEE J. Sel. Top. Signal Process. 4 (1) (2010) 66–78.
[24] W.Q. Wang, Z. Zheng, Hybrid MIMO and phased-array directional modulation for physical layer security in mmwave wireless communications, IEEE J. Sel. Areas Commun. 36 (7) (2018) 1383–1396.
[25] Z. Han, S. Shen, Y. Zhang, C.-Y. Chiu, R. Murch, A pattern correlation decomposition method for analysis of ESPAR in single-RF MIMO systems, IEEE Trans. Wirel. Commun. 21 (7) (2022) 4654–4668.
[26] Z. Han, Y. Zhang, S. Shen, Y. Li, C.-Y. Chiu, R. Murch, Characteristic mode analysis of ESPAR for single-RF MIMO systems, IEEE Trans. Wirel. Commun. 20 (4) (2021) 2353–2367.
[27] M.A. Sedaghat, R. Müller, G. Fischer, A. Ali, Discrete load-modulated single-RF MIMO transmitters, in: WSA 2016; 20th International ITG Workshop on Smart Antennas, 2016, pp. 1–7.

[28] A.K. Bhattacharyya, Phased Array Antennas: Floquet Analysis, Synthesis, BFNs, and Active Array Systems, John Wiley and Sons, Inc. Press, 2006.
[29] B.H. Fleury, First- and second-order characterization of direction dispersion and space selectivity in the radio channel, IEEE Trans. Inform. Theory 46 (6) (2000) 2027–2044.
[30] G. Nobile, M. Cacciatoand, E. Vasta, Measuring active power as the difference between the peak value of instantaneous power and the apparent power, Sensors 22 (9) (2022) 3517.
[31] T. Zhang, S. Cialdea, A.E. Emanuel, J.A. Orr, The implementation of correct reactive power measurement is long overdue, in: 2014 16th International Conference on Harmonics and Quality of Power, ICHQP, Bucharest, Romania, 2014, pp. 467–473.
[32] Shayesteh, H. Behroozi, Z. Ding, Digital/hybrid beamforming via KLMS algorithm in presence of mutual coupling with experimental evaluation for 5G and V2V communications, IEEE Trans. Veh. Technol. 73 (4) (2024) 5229–5242.
[33] J.-S. Fu, Adaptive Impedance Matching Circuits Based on Ferroelectric and Semiconductor Varactors (Ph.D. dissertation), Dept. Elect. Eng. Univ. Michigan, Ann Arbor, MI, USA, 2009.
[34] P. Sjoblom, H. Sjoland, An adaptive impedance tuning CMOS circuit for ISM 2.4-GHz band, IEEE Trans. Circuits Syst. I. Regul. Pap. 52 (6) (2005) 1115–1124.
[35] H. Kawakami, T. Ohira, Electrically steerable passive array radiator (ESPAR) antennas, IEEE Antennas Propag. Mag. 47 (2) (2005) 43–50.
[36] A. Li, C. Masouros, C.B. Papadias, MIMO transmission for single-fed ESPAR with quantized loads, IEEE Trans. Commun. 65 (7) (2017) 2863–2876.
[37] J. Lee, Y. Lee, K. Lee, T.J. Kim, λ/64-spaced compact ESPAR antenna via analog RF switches for a single RF chain MIMO system, ETRI J. 41 (4) (2019) 536–548.
[38] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
[39] C. Roos, A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization, Siam J. Optim. 16 (4) (2006) 1110–1136.
[40] G. Cornuejols, R. Tutuncu, Optimization Methods in Finance, Cambridge University Press, 2006.
[41] T.F. Coleman, A. Verma, A preconditioned conjugate gradient approach to linear equality constrained minimization, Comput. Optim. Appl. 20 (1) (2001) 61–72.
[42] X. Zhang, Matrix Analysis and Applications, Tsinghua University Press, 2014.
[43] 3GPP TR 38901, Study on channel model for frequencies from 0.5 to 100 ghzv16.1.0, 2020.