Mutual Information for Electromagnetic Information Theory Based on Random Fields

Zhongzhichao Wan, *Student Member, IEEE,* Jieao Zhu, *Student Member, IEEE,*
Zijian Zhang, *Student Member, IEEE,* Linglong Dai, *Fellow, IEEE,* and Chan-Byoung Chae, *Fellow, IEEE*

Abstract—Traditional channel capacity based on the discrete spatial dimensions mismatches the continuous electromagnetic fields. For the wireless communication system in a limited region, the spatial discretization may results in information loss because the continuous field can not be perfectly recovered from the sampling points. Therefore, electromagnetic information theory based on spatially continuous electromagnetic fields becomes necessary to reveal the fundamental theoretical capacity bound of communication systems. In this paper, we propose analyzing schemes for the performance limit between continuous transceivers. Specifically, we model the communication process between two continuous regions by random fields. Then, for the white noise model, we use Mercer expansion to derive the mutual information between the source and the destination. For the close-form expression, an analytic method is introduced based on autocorrelation functions with rational spectrum. Moreover, the Fredholm determinant is used for the general autocorrelation functions to provide the numerical calculation scheme. Further works extend the white noise model to colored noise and discuss the mutual information under it. Finally, we build an ideal model with infinite-length source and destination which shows a strong correpsondence with the time-domain model in classical information theory. The mutual information and the capacity are derived through the spatial spectral density.

Index Terms—Electromagnetic information theory (EIT), mutual information, random field, Fredholm determinant, spatial spectral density (SSD).

I. Introduction

Wireless communication systems employ electromagnetic fields with three continuous spatial dimensions for information exchange. However, the modern multiple-input multiple-output (MIMO) technology, viewed as a discretization of the continuous spatial dimensions, mismatches the continuous nature of electromagnetic fields

Zhongzhichao Wan, Jieao Zhu, Zijian Zhang and Linglong Dai are with the Department of Electronic Engineering, Tsinghua University as well as Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China (E-mails: {wzzc20, zja21, zhangzj20}@mails.tsinghua.edu.cn; daill@tsinghua.edu.cn).

Chan-Byoung Chae is with the School of Integrated Technology, Yonsei University, Seoul 03722, Korea (E-mails: cbchae@yonsei.ac.kr).

in real-world communication systems, thus causing its inability to fully explore the spatial information [1]. Therefore, we should restore to the continuous electromagnetic fields to analyze the fundamental performance limit of an arbitrary communication system, motivating the research of the electromagnetic information theory (EIT) [2].

For EIT, one asymptotic approach to analyze the fundamental limit is called the spatial bandwidth [3], based on which the degrees of freedom (DoF) can be theoretically derived. The spatial bandwidth of scattered fields under a time-harmonic model was rigorously derived in [4]. Further works extended the time-harmonic model, which was focused on a single frequency point, to a more general band-limited model, and analyzed the DoF [5]. Unfortunately, these spatial bandwidth-based procedures rely on the assumption that the occupied region of the information destination tends to infinity. In contrast, communication systems are often confined in a finite-sized space in practice. In this practical scenario, the spatial bandwidth method becomes inaccurate due to the condition of the Nyquist sampling theorem.

For the practical communication system in a limited space, several works utilized the orthogonal expansion techniques to derive the DoF for EIT. For example, inspired by the orthogonal frequency-division multiplexing (OFDM), the authors in [1] designed a set of bases to expand the Green's function, and derived the DoF between two parallel finite-length linear antennas. Moreover, for the scenario where the two finite-length linear antennas bear an intersection angle, a heuristic method was proposed to construct the bases for the derivation of DoF [6]. These works provide us with useful methods for analyzing the DoF in EIT. However, they heavily rely on the assumption of deterministic signals, thus being unable to derive the information-theoretic capacity where stochastic properties should first be modeled.

To derive the capacity of the electromagnetic channels, several works introduced basis expansion to split the continuous electromagnetic channel to almost orthogonal channels. The approximation of the basis based on the radiating term of the Green's function was derived [7]. The orthogonal bases between a pair of concentric spherical source and destination were shown to be spherical

harmonic functions [2]. Moreover, the capacity of the system with spherical source and destination in lossy medium was obtained in [8]. Another important approach using the Kolmogorov ϵ-capacity was utilized in [9] and [10], which was based on the maximum amount of information transmissible through the channel with an uncertainty level.

These works already found the best bases for the electromagnetic channels and derived the capacity with finite bases. The capacity can be viewed as the mutual information under a specific field distribution, where the mutual information is maximized. However, an analytic solution and calculation scheme for the mutual information between the transceivers with arbitrary given field distribution has not been obtained in the literature of EIT. Such analysis requires the analytical analyzing scheme based on infinite bases decomposed from the continuous channel. Moreover, the existing works considered the spatial white Gaussian noise field. Since white noise is only a special case of colored noise field and the practical noise field may include non-white component, further works are necessary to explore the mutual information and capacity under colored noise scenario.

Different from existing works, in this paper, we first build a model considering the non-white component of the noise field. Then, we provide a strict analysis framework of the mutual information and capacity based on the random field theory and operator theory[1]. The electromagnetic waves that carry information are modeled as random fields, which follow the statistical approach of Shannon. We introduce the theory of operator analysis to derive the general expression of the mutual information and provide the numerical calculation scheme of it. Specifically, the contributions of this paper are summarized as follows:

- First, we develop a system model of the wireless communication between two continuous regions. We use random fields to capture the statistical characteristics of the signals and noise in the communication system. The mutual information between the source and the destination is defined by a supremum taken over all the testing functions.
- Then, we consider a simplified model of the communication with finite-length transceivers under white noise field. By exploiting Mercer expansion, we derive the mutual information between the source and the destination. Next, we introduce the analytical scheme of deriving the mutual information for autocorrelation functions with rational spectrum. A special case, which has been adopted in the existing scattering models, is analyzed to show how the analytical scheme works. To generalize the information formulas, we introduce a tool called the Fredholm determinant to obtain a general expression of the mutual information, which enables numerical calcu-

[1]Simulation codes are provided to reproduce the results in this paper: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.
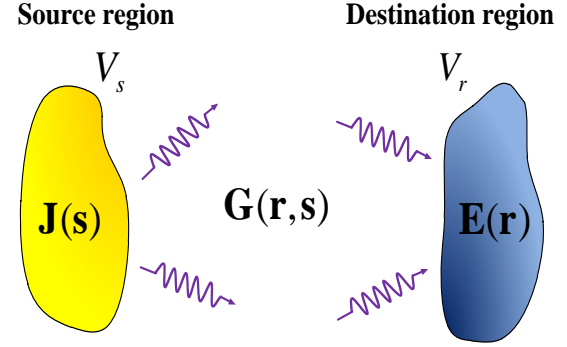

Fig. 1. Electromagnetic model of the communication between two arbitrary continuous regions.

lation and convergence analysis. Some discussions about the extendability of the above schemes are also included.

- Moreover, we adopt the operator analysis schemes to extend the mutual information formula from white noise model to more general cases with colored noise. A discretization scheme is provided for the numerical calculation. Numerical results show the convergence of the calculation scheme and also provide some insights into the performance comparison between continuous and discrete systems.
- Finally, to illustrate the close relationship between the proposed random field-based theory and the classical time-domain information theory, we build an ideal model with parallel infinite-length linear source and destination. In this model, the mutual information between the source and the destination is represented by the spatial spectral density (SSD) of the electric field and the noise field on the destination. The optimal current density distribution on the source is derived by variational calculus to achieve the maximum mutual information, i.e., the capacity between the source and the destination.

Notation: bold uppercase characters denote matrices; bold lowercase characters denote vectors; the dot $\cdot$ denotes the scalar product of two vectors, or the matrix-vector multiplication. $\mathbb{E}[x]$ denotes the mean of random variable x; ϵ_0 is the permittivity of a vacuum, μ_0 is the permeability of a vacuum, and c is the speed of light in a vacuum; $*$ denotes the convolution operation, and $\mathscr{F}[f(x)]$ denotes the Fourier transform of $f(x)$; $(f(x))^+$ is equal to $\frac{f(x)+|f(x)|}{2}$; ∇ is the nabla operator, and $\nabla\times$ is the curl operator; $J_0(x)$ is the Bessel function; $K_0(x)$ is the zeroth-order modified Bessel function of the second kind; $Y_0(x)$ is the zeroth-order Neumann function; $|\phi\rangle$ is the quantum mechanical notation of a function ϕ, where the inner product is denoted by $\langle\psi|\phi\rangle$; the matrix element of an operator T under an orthonormal basis $\{\phi_j\}_{j=1}^{+\infty}$ is represented by the Dirac notation $T_{ij} = \langle i|T|j\rangle$;

$\det(\cdot)$ denotes the matrix determinant or the Fredholm determinant.

II. ELECTROMAGNETIC WAVE COMMUNICATION MODEL

Maxwell's equations, which consist of four differential equations, describe how electromagnetic fields are generated by currents, charges and the change of fields [11]. Here we focus on the two curl equations in Maxwell's equations, which are called the Faraday's law and the Ampère's law, i.e., $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ and $\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$.

These equations are the fundamental physical laws that govern the electromagnetic wave communications. To simplify the analysis, we adopt the common assumption that the electromagnetic wave oscillates on a single frequency point, which is the well-known time-harmonic assumption [2]. In this way, the temporal derivative operator $\partial / \partial t$ can be replaced by $-\mathrm{j}\omega$ in Maxwell's equations. The time-harmonic assumption simplifies the Maxwell's equations into complex-valued partial differential equations involving only spatial derivatives, i.e., $\nabla \times \mathbf{E} = \mathrm{j}\omega \mathbf{B}$ and $\nabla \times \mathbf{H} = \mathbf{J} - \mathrm{j}\omega \mathbf{D}$. From the above equations we can obtain the vector wave equation [12] as

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \kappa_0^2 \mathbf{E}(\mathbf{r}) = \mathrm{j}\omega \mu_0 \mathbf{J}(\mathbf{r}) = \mathrm{j}\kappa_0 Z_0 \mathbf{J}(\mathbf{r}), \quad (1)$$

where $\kappa_0 = \omega \sqrt{\mu_0 \varepsilon_0}$ is the wavenumber and $Z_0 = \mu_0 c = 120\pi \, [\Omega]$ is the free-space intrinsic impedance.

Consider two arbitrary regions V_s and V_r as the source and the destination for wireless communications. The current density at the source is $\mathbf{J}(\mathbf{s}) : \mathbb{R}^3 \to \mathbb{C}^3$, and the induced electric field at the destination is $\mathbf{E}(\mathbf{r})$, where $\mathbf{r} \in \mathbb{R}^3$ is the coordinate of the field observer. The received electric field is $\mathbf{Y}(\mathbf{r}) = \mathbf{E}(\mathbf{r}) + \mathbf{N}(\mathbf{r})$, where $\mathbf{N}(\mathbf{r})$ is the noise field. Exploiting the linear nature of (1), the Green's function $\mathbf{G}(\mathbf{r}, \mathbf{s}) \in \mathbb{C}^{3 \times 3}$ is introduced to solve this equation. Utilizing the Green's function, the electric field $\mathbf{E}(\mathbf{r})$ can be expressed by

$$\mathbf{E}(\mathbf{r}) = \int_{V_s} \mathbf{G}(\mathbf{r}, \mathbf{s}) \mathbf{J}(\mathbf{s}) \mathrm{d}\mathbf{s}, \quad \mathbf{r} \in V_r. \quad (2)$$

The Green's function in unbounded, homogeneous mediums at a fixed frequency point is [13]

$$\begin{aligned}
\mathbf{G}(\mathbf{r}, \mathbf{s}) &= \frac{\mathrm{j}\kappa_0 Z_0}{4\pi} \left(\mathbf{I} + \frac{\nabla_{\mathbf{r}} \nabla_{\mathbf{r}}^{\mathrm{H}}}{\kappa_0^2} \right) \frac{e^{\mathrm{j}\kappa_0 \|\mathbf{r}-\mathbf{s}\|}}{\|\mathbf{r}-\mathbf{s}\|} \\
&= \frac{\mathrm{j}\kappa_0 Z_0}{4\pi} \frac{e^{\mathrm{j}\kappa_0 \|\mathbf{r}-\mathbf{s}\|}}{\|\mathbf{r}-\mathbf{s}\|} \left[(\mathbf{I} - \hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{H}}) \right. \\
&\quad + \frac{\mathrm{j}}{2\pi \|\mathbf{r}-\mathbf{s}\|/\lambda} (\mathbf{I} - 3\hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{H}}) \\
&\quad \left. - \frac{1}{(2\pi \|\mathbf{r}-\mathbf{s}\|/\lambda)^2} (\mathbf{I} - 3\hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{H}}) \right] [\Omega/\mathrm{m}^2],
\end{aligned}$$
$$(3)$$

where $\hat{\mathbf{p}} = \frac{\mathbf{p}}{\|\mathbf{p}\|}$ and $\mathbf{p} = \mathbf{r} - \mathbf{s}$. Some existing works [1] [6] adopted far-field assumption and only preserved the item $\frac{\mathrm{j}\kappa_0 Z_0}{4\pi} \frac{e^{\mathrm{j}\kappa_0 \|\mathbf{r}-\mathbf{s}\|}}{\|\mathbf{r}-\mathbf{s}\|} (\mathbf{I} - \hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathrm{H}})$ in (3). Different from these works, we use the whole expression in (3) to make our analysis and numerical calculation valid for both far-field and near-field scenarios.

In this section, we have reviewed the basic model of electromagnetic wave communication. However, the information transmission process naturally involves the elimination of uncertainty, which requires probabilistic modeling of the EM fields. The amount of information transmitted through the EM field can be viewed as the amount of uncertainty eliminated in the instant the field in V_r is observed. Thus, in the next section, we will use Gaussian random fields to capture the statistical properties of electromagnetic fields, and then derive the mutual information.

III. RANDOM FIELD MODELING

In this section, we analyze the statistical properties of the signals and noises. We use Gaussian random fields to model the signals and noise of wireless communication systems, which follows the statistical approach of Shannon. Based on the random field modeling, we define the mutual information in a supremum form, which provides foundations for further analysis.

A. Random field modeling of the signals

The information-theoretic analysis in this paper is based on stochastic models according to Shannon's approach, which motivates us to model the continuous EM signals by a random field. Each realization of the random field represents a radiating and receiving pattern during one channel use. All the realizations are gathered together to form a probabilistic ensemble, which reflects the statistical characteristics of the wireless communication system, thus leading to mutual information and capacity. Among all kinds of random fields, the Gaussian random field is of theoretical significance, since its Gaussian distribution is the capacity-achieving distribution of an AWGN channel [14]. As a result, we use Gaussian random field to depict the statistical characteristics of both the current density at the source and the electric field at the destination. The Gaussian random field is assumed to be continuous, separable, and fully characterized by its mean and autocorrelation functions [15]. The current density $\mathbf{J}(\mathbf{s})$ at the source and the induced electric field $\mathbf{E}(\mathbf{r})$ at the destination are considered as random fields with autocorrelation functions given by

$$R_{\mathbf{J}}(\mathbf{s}, \mathbf{s}') = \mathbb{E}[\mathbf{J}(\mathbf{s})\mathbf{J}^{\mathrm{H}}(\mathbf{s}')] \, [\mathrm{A}^2/\mathrm{m}^4], \quad (4a)$$
$$R_{\mathbf{E}}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[\mathbf{E}(\mathbf{r})\mathbf{E}^{\mathrm{H}}(\mathbf{r}')] \, [\mathrm{V}^2/\mathrm{m}^2]. \quad (4b)$$

From (2), we can derive the relation between the auto-correlation function of the current density and that of the electric field as

$$\begin{aligned}
R_{\mathbf{E}}(\mathbf{r}, \mathbf{r}') &= \mathbb{E}[\mathbf{E}(\mathbf{r})\mathbf{E}^{\mathrm{H}}(\mathbf{r})] \\
&= \int_{V_s} \int_{V_s} \mathbf{G}(\mathbf{r}, \mathbf{s}) R_{\mathbf{J}}(\mathbf{s}, \mathbf{s}') \mathbf{G}^{\mathrm{H}}(\mathbf{r}, \mathbf{s}) \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{s}'.
\end{aligned}$$
$$(5)$$

In this model, we assume that the channel is deterministic, and it only consists of line-of-sight components, which means that the uncertainty of the field measurements only comes from the additive noise field. The model of the noise field will be discussed in the following subsection.

B. Random field modeling of the noise

In this subsection, we will analyze the model of the noise field. The noise in the communication between a pair of continuous source and destination can be decomposed into two categories: measurement noise and radiation interference.

For the measurement noise, the authors of [2] attributed this kind of noise to the non-ideal factors in communications, including imperfect antenna locations, imprecise field measurements, and numerical errors inevitable during signal processing. Since these defects are usually spatially uncorrelated, the noise is then modeled by a white Gaussian random field, which can be characterized by

$$\mathbb{E}\left[\mathbf{E}_{\text{noise}}(\mathbf{r})\mathbf{E}_{\text{noise}}^{\text{H}}(\mathbf{r}')\right] = \frac{n_0}{2}\mathbf{I}_3\delta(\mathbf{r} - \mathbf{r}'), \qquad (6)$$

where n_0 [V$^2 \cdot$ m] is the single-sided power spectral density in the three-dimensional wavenumber domain. This measurement noise is identical in distribution with the so-called thermal noise, which is widely used in the literature, but it is a wider notion consisting of all the spatially uncorrected undesired signals. The white Gaussian noise (WGN) assumption of this measurement noise helps simplify our analysis. However, the thermal noise model can not depict the interference from other electromagnetic sources which cause the radiation interference.

For the radiation interference, it can be viewed as the superposition of the incident electromagnetic waves which are not generated by the desired source current [1]. The simplest analysis is to assume an isotropic incident wave, i.e., the incident power is uniformly distributed among different spatial directions. Thus, the electromagnetic wave which impinges on the destination from an arbitrary angle can be represented in the spherical coordinates by an elevation angle $\theta \in [0, \pi)$, and an azimuth angle $\phi \in [-\pi, \pi]$. It is derived in [1] that, with this isotropic assumption, the noise field can be characterized by

$$\mathbb{E}\left[\mathbf{N}(\mathbf{r} + \mathbf{r}')\mathbf{N}(\mathbf{r}')^{\text{H}}\right] = \sigma^2\rho(\mathbf{r})\mathbf{I}_3, \qquad (7a)$$

$$\rho(\mathbf{r}) = \text{sinc}\left(\frac{2\,\|\mathbf{r}\|}{\lambda}\right) = \text{sinc}\left(\frac{2r}{\lambda}\right). \qquad (7b)$$

Note that from (7), the noise field should also be completely isotropic, i.e., the field components in orthogonal directions should not be correlated. Unfortunately, this is not the case in reality. This misunderstanding occurs because the physical law that electromagnetic waves are transverse waves was not taken into account in [1], i.e., the direction of the oscillating electric field should be perpendicular to the wave propagation direction. Here we will fix this problem and derive a new noise model based on random fields.

The noise field observed at position $\mathbf{r}$ can be derived by integrating the waves $\mathbf{a} \in \mathbb{C}^3$ incoming $\mathbf{r}$ from all spatial directions, which can be represented by

$$\mathbf{N}(\mathbf{r}) = \int_{-\pi}^{\pi}\int_0^{\pi} \mathbf{a}(\theta, \varphi)e^{\mathbf{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}\theta\mathrm{d}\varphi, \qquad (8)$$

where the wavevector is expressed in spherical coordinates as

$$\boldsymbol{\kappa} = \frac{2\pi}{\lambda}\left[\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta\right] \in \mathbb{R}^3. \qquad (9)$$

Let $\kappa = \|\boldsymbol{\kappa}\|$, and the normalized propagation direction vector $\hat{\boldsymbol{\kappa}} = \boldsymbol{\kappa}/\kappa$, then we have $\mathbf{a}^{\text{H}}(\theta, \varphi)\hat{\boldsymbol{\kappa}} = 0$, i.e., the incident electric field oscillates in the plane perpendicular to the wavevector $\mathbf{k}$. By further assuming an uncorrelated random incident phase of the two distinct polarizations, the correlation matrix of the random vector $\mathbf{a}(\theta, \varphi)$ can be represented by constant multiples of $\mathbf{I} - \hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\text{T}}$. Thus, we can obtain the angular autocorrelation of $\mathbf{a}(\theta, \varphi)$ to be

$$\mathbb{E}\left[\mathbf{a}(\theta, \varphi)\mathbf{a}^{\text{H}}(\theta', \varphi')\right] = \sigma^2 f(\theta, \varphi)(\mathbf{I} - \hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\text{T}})$$
$$\delta(\theta - \theta')\delta(\varphi - \varphi'), \qquad (10)$$

where the dimensionless density $f(\theta, \varphi)$ equals $\frac{\sin\theta}{4\pi}$ because of the isotropic propagation assumption [1], and σ^2 [V^2/m^2] is the average second-order moment of the noise field. With (10), the autocorrelation function of the noise field can then be derived by integrating incident waves from all the spatial directions over the unit spherical shell:

$$\mathbb{E}\left[\mathbf{N}(\mathbf{r} + \mathbf{r}')\mathbf{N}^{\text{H}}(\mathbf{r}')\right]$$
$$= \int_{-\pi}^{\pi}\int_0^{\pi}\int_{-\pi}^{\pi}\int_0^{\pi} \mathbb{E}\left[\mathbf{a}(\theta, \varphi)\mathbf{a}^{\text{H}}(\theta', \varphi')\right] e^{\mathbf{j}\boldsymbol{\kappa}\cdot(\mathbf{r}+\mathbf{r}')}$$
$$e^{-\mathbf{j}\boldsymbol{\kappa}'\cdot\mathbf{r}'}\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}\theta'\mathrm{d}\varphi'$$
$$= \int_{-\pi}^{\pi}\int_0^{\pi} \sigma^2 f(\theta, \varphi)(\mathbf{I} - \hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\text{T}})e^{\mathbf{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}\theta\mathrm{d}\varphi$$
$$= \frac{\sigma^2}{4\pi}\iint\limits_{S_3}(\mathbf{I} - \hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\text{T}})e^{\mathbf{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S, \qquad (11)$$

where $\mathrm{d}S$ denotes the area element on the unit spherical shell. To solve the definite integral in (11), we define two auxiliary functions $f_1(\beta)$ and $f_2(\beta)$ as

$$f_1(\beta) = \int_{-1}^{1} e^{\mathbf{j}\beta x}\mathrm{d}x = 2\frac{\sin\beta}{\beta}, \qquad (12a)$$

$$f_2(\beta) = \int_{-1}^{1} x^2 e^{\mathbf{j}\beta x}\mathrm{d}x = 2\left(\frac{\sin\beta}{\beta} + \frac{2\cos\beta}{\beta^2} - \frac{2\sin\beta}{\beta^3}\right). \qquad (12b)$$

Then, we have

$$\frac{\sigma^2}{4\pi}\iint_{S_3} \mathbf{I}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}\mathbf{S} = \frac{\sigma^2}{4\pi}\int_{-\pi}^{\pi}\int_0^{\pi}\mathbf{I}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\sin\theta\mathrm{d}\theta\mathrm{d}\varphi$$

$$= \frac{\sigma^2}{4\pi}\int_{-\pi}^{\pi}\int_0^{\pi}\mathbf{I}e^{\mathrm{j}\kappa r\cos\theta}\sin\theta\mathrm{d}\theta\mathrm{d}\varphi$$

$$= \frac{\sigma^2}{2}\mathbf{I}f_1(\kappa r). \tag{13}$$

For the term $\frac{\sigma^2}{4\pi}\iint_{S_3}\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S$, we consider its component along $\hat{\mathbf{r}}$ and perpendicular to $\hat{\mathbf{r}}$ sequentially to derive its analytical representation. For the component along $\hat{\mathbf{r}}$, we have

$$\frac{\sigma^2}{4\pi}\iint_{S_3}\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S\cdot\hat{\mathbf{r}} = \frac{\sigma^2}{4\pi}\iint_{S_3}\hat{\boldsymbol{\kappa}}(\hat{\boldsymbol{\kappa}}^{\mathrm{T}}\hat{\mathbf{r}})e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S$$

$$\stackrel{(a)}{=} \frac{\sigma^2}{2}\int_0^{\pi}\sin\theta\cos^2\theta e^{\mathrm{j}\kappa r\cos\theta}\mathrm{d}\theta\hat{\mathbf{r}}$$

$$= \frac{\sigma^2}{2}f_2(\kappa r)\hat{\mathbf{r}}, \tag{14}$$

where (a) comes from the symmetric property of the integral along the $\hat{\mathbf{r}}$ axis. In order to solve the components which are perpendicular to $\hat{\mathbf{r}}$, we assume that $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ are two orthogonal directions on the plane perpendicular to $\hat{\mathbf{r}}$, where the direction cosines of unit vector $\hat{\kappa}$ under the newly defined coordinate system $(\hat{\mathbf{t}}_1, \hat{\mathbf{t}}_2, \hat{\mathbf{r}})$ are $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. Then, we can derive

$$\hat{\mathbf{t}}_1 \cdot \frac{\sigma^2}{4\pi}\iint_{S_3}\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S\cdot\hat{\mathbf{t}}_1$$

$$= \frac{\sigma^2}{4\pi}\iint_{S_3}(\sin\theta\cos\varphi)^2 e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S$$

$$= \frac{\sigma^2}{4\pi}\int_{-\pi}^{\pi}\int_0^{\pi}\sin^3\theta\cos^2\varphi e^{\mathrm{j}\kappa r\cos\theta}\mathrm{d}\theta\mathrm{d}\varphi$$

$$= \frac{\sigma^2}{4}\int_0^{\pi}\sin^3\theta e^{\mathrm{j}\kappa r\cos\theta}\mathrm{d}\theta = \frac{\sigma^2}{4}(f_1(\kappa r)-f_2(\kappa r)), \tag{15}$$

and

$$\hat{\mathbf{t}}_2 \cdot \frac{\sigma^2}{4\pi}\iint_{S_3}\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}}e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S\cdot\hat{\mathbf{t}}_2 = \frac{\sigma^2}{4\pi}\iint_{S_3}(\sin\theta\sin\varphi)^2 e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}\mathrm{d}S$$

$$= \frac{\sigma^2}{4}(f_1(\kappa r)-f_2(\kappa r)). \tag{16}$$

Similarly, we can prove that the components $\hat{\mathbf{t}}_i \cdot \iint_{S_3}\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}}e^{\mathrm{j}\mathbf{k}\cdot\mathbf{r}}\mathrm{d}S\cdot\hat{\mathbf{t}}_j = 0$ for all $i \neq j, i,j \in \{1,2\}$. Combining all the above results into a 3-by-3 autocorrelation matrix, we have

$$\mathbb{E}\left[\mathbf{N}(\mathbf{r}+\mathbf{r}')\mathbf{N}^{\mathrm{H}}(\mathbf{r}')\right]$$

$$= \frac{\sigma^2}{2}f_1(\kappa r)\mathbf{I} - \frac{\sigma^2}{2}f_2(\kappa r)\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{T}}$$

$$\quad - \frac{\sigma^2}{4}(f_1(\kappa r)-f_2(\kappa r))(\mathbf{I}-\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{T}})$$

$$= \frac{\sigma^2}{4}(f_1(\kappa r)+f_2(\kappa r))\mathbf{I} + \frac{\sigma^2}{4}(f_1(\kappa r)-3f_2(\kappa r))\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{T}}. \tag{17}$$

Specifically, the noise field correlation measured along the polarization direction $\hat{\mathbf{r}}$ is given by

$$\mathbb{E}\left[N(r+r')N(r')\right] = \hat{\mathbf{r}}^{\mathrm{H}}\mathbf{R_N}\hat{\mathbf{r}} = \frac{\sigma^2}{2}f_1(\kappa r) - \frac{\sigma^2}{2}f_2(\kappa r)$$

$$= 2\sigma^2\left(\frac{\sin(\kappa r)}{(\kappa r)^3} - \frac{\cos(\kappa r)}{(\kappa r)^2}\right). \tag{18}$$

The result can also be applied to a uniform linear antenna array placed on the x-axis, where the polarization orientation is parallel to the y-axis. With this assumption, the noise field correlation is given by

$$\mathbb{E}\left[N(r+r')N(r')\right] = \frac{\sigma^2}{4}(f_1(\kappa r)+f_2(\kappa r)). \tag{19}$$

This equation reveals that different from our previous understanding of the noise that its autocorrelation is of the form $\mathrm{sinc}(\cdot)$. The precise noise autocorrelation contains more high-spatial-frequency components, which come from the term $f_2(\kappa r)$.

Remark 1. *Equation (17) fully depicts the autocorrelation of the noise field of three polarization orientations under the assumption of isotropic incidence. This form of autocorrelation is also employed to describe the channel autocorrelation [16, eqn (19)], which is reasonable because the noise discussed here can be considered as unwanted signals received through a random channel.*

Remark 2. *This noise model is based on the isotropic scattering assumption, which can be easily extended to the non-isotropic case. If we assume that the radiation interference is not uniform in spatial angle but concentrated near a certain angle, we can use von Mises-Fisher (vMF) distribution [17, 18] to depict such angular concentration. A similar idea was proposed in [19] to analyze the angular selectivity of the random channel modeling but the analytic solution was not given. For the 3-dimensional von Mises-Fisher distribution, which is a close approximation to the spherical analogue of the Gaussian distribution, the probability density function is given by*

$$f(\boldsymbol{x}|\boldsymbol{\mu}) = C(\|\boldsymbol{\mu}\|)e^{\hat{\boldsymbol{x}}\cdot\boldsymbol{\mu}}, \tag{20}$$

where $C(\|\boldsymbol{\mu}\|) = \frac{\|\boldsymbol{\mu}\|}{2\pi(e^{\|\boldsymbol{\mu}\|}-e^{-\|\boldsymbol{\mu}\|})}$ is the normalization constant. Instead of (11) we have

$$\mathbb{E}\left[\mathbf{N}(\mathbf{r}+\mathbf{r}')\mathbf{N}^{\mathrm{H}}(\mathbf{r}')\right] = \frac{\sigma^2}{4\pi}\iint_{S_3}(\mathbf{I}-\hat{\boldsymbol{\kappa}}\hat{\boldsymbol{\kappa}}^{\mathrm{T}})e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}e^{\hat{\boldsymbol{\kappa}}\cdot\boldsymbol{\mu}}\mathrm{d}S. \tag{21}$$

Then, by replacing $\mathbf{r}\in\mathbb{R}^3$ in (11) by $\mathbf{r}=\mathbf{r}_R-\mathrm{j}\boldsymbol{\mu}/\kappa\in\mathbb{C}^3$. The imaginary part $\mathrm{j}\boldsymbol{\mu}/k$ can be combined into the spatial harmonic factor $e^{\mathrm{j}\boldsymbol{\kappa}\cdot\mathbf{r}}$ of the planar wave, in order to describe a vMF distribution of the incident angles θ

and φ. Equation (17) can then be extended to the non-isotropic case accordingly by analytic continuation [20] techniques, leading to

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{N}(\mathbf{r} + \mathbf{r}')\mathbf{N}^{\mathrm{H}}(\mathbf{r}')\right] = {} & \frac{\sigma^2}{4}(f_1(\kappa r_1) + f_2(\kappa r_1))\mathbf{I} \\
& + \frac{\sigma^2}{4}(f_1(\kappa r_1) - 3f_2(\kappa r_1))\hat{\mathbf{r}}_1\hat{\mathbf{r}}_1^{\mathrm{T}},
\end{aligned}
\tag{22}
$$

where $r_1 = \sqrt{(\boldsymbol{r} - \mathrm{j}\boldsymbol{\mu}/\kappa)^{\mathrm{T}}(\boldsymbol{r} - \mathrm{j}\boldsymbol{\mu}/\kappa)}$ and $\hat{\mathbf{r}}_1 = \frac{\boldsymbol{r} - \mathrm{j}\boldsymbol{\mu}/\kappa}{r_1}$.

C. Mutual information based on random field modeling

The mutual information was originally defined by Shannon in his seminal paper [14] to be the amount of uncertainty reduced by observations of the channel outputs. The channel usually admits discrete random input symbols, and gives out discrete output symbols. However, in our study on electromagnetic information theory, the channel output $\mathbf{E}(\mathbf{r})$ is a spatially continuous electromagnetic field, to which we assign uncertainty by modeling it as a Gaussian random field. Similar to the standard definition of mutual information, upon obtaining noisy measurements $\mathbf{Y}(\mathbf{r})$ of this field, we can also evaluate the amount of information received through the uncertainty reduction mechanism, but the field measurements should be well-defined to avoid the "continuum" difficulties encountered when analyzing a spatially continuous field.

As introduced in the above subsection, the noise is modeled as a Gaussian random field with autocorrelation function $R_{\mathbf{N}}(\mathbf{r}, \mathbf{r}')$ [1] and the autocorrelation function of the noisy electric field is denoted by $R_{\mathbf{Y}}(\mathbf{r}, \mathbf{r}')$. Following the definition in [15], we use testing functions to define the mutual information between the random fields $\mathbf{J}$ and $\mathbf{Y}$ by

$$
I(\mathbf{J}; \mathbf{Y}) = \sup\{I(\mathbf{J}(\phi_1, \cdots, \phi_m), \mathbf{Y}(\psi_1, \cdots, \psi_n))\},
\tag{23}
$$

where $\mathbf{J}(\phi_1, \ldots, \phi_m) = \{\langle \phi_1|\mathbf{J}\rangle, \cdots, \langle \phi_m|\mathbf{J}\rangle\}$ and $\mathbf{Y}(\psi_1, \ldots, \psi_n) = \{\langle \psi_1|\mathbf{Y}\rangle, \cdots, \langle \psi_n|\mathbf{Y}\rangle\}$ are random vectors, and the inner product $\langle \phi|\mathbf{J}\rangle$ equals to $\int_{V_s} \phi^*(\mathbf{s})\mathbf{J}(\mathbf{s})\mathrm{d}\mathbf{s}$. The supremum is taken over all the possible integers m, n, and testing functions $\phi, \psi \in \Phi$, where Φ is the set of all the smooth functions that vanish outside V_s.

According to [15, **Theorem 1.3**], we can select a sequence of ϕ_i^k which converges to $\delta(\mathbf{s} - \mathbf{s}_i)$ when $k \to \infty$. Then, the mutual information $I(\mathbf{J}, \mathbf{Y})$ can be calculated by the mutual information between two groups of points, expressed as $\sup\{I(\mathbf{J}(\mathbf{s}_1, \cdots, \mathbf{s}_m), \mathbf{Y}(\mathbf{r}_1, \cdots, \mathbf{r}_n))\}$. This means that we can use the random variables defined on the spatial sampling points $\{\mathbf{s}_1, \cdots, \mathbf{s}_m\}$ and $\{\mathbf{r}_1, \cdots, \mathbf{r}_n\}$ to approximate the mutual information between the two random continuous electromagnetic fields. It is worth noting that this claim only shows the theoretical existence of the sampling points, which does not involve the way to
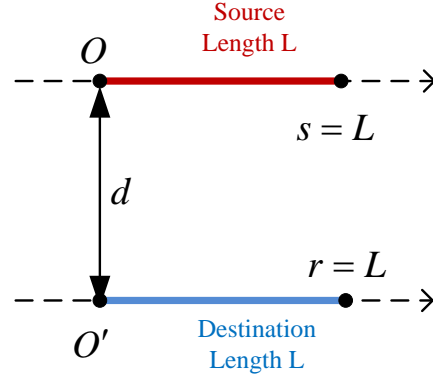

Fig. 2. Analysis of mutual information between a source and destination, both of finite length.

choose them. In this paper, we will show a discretization scheme of the continuous field and the convergence is provided according to numerical analysis.

In this section, the mutual information between continuous regions is defined in a supremum form using random fields. Then, in the next section, we will simplify the model with continuous regions to obtain an analytical solution to the mutual information.

IV. MUTUAL INFORMATION WITH FINITE-LENGTH TRANSCEIVERS UNDER WHITE NOISE

In this section, to obtain more insightful results, we consider the simplified case where a pair of parallel finite-length linear source and destination are employed as transceivers. White noise model is used in this section for simplicity and non-white noise model will be discussed in the next section for generality. Since the simplified model here is similar to the classical stochastic process in the one-dimensional time domain, most of the existing analyzing schemes and tools can be transfered to the analysis in this simplified scenario. To make the schemes more general, we also show how the adopted tools and results from stochastic process can be extended to random field with at least two dimensions. Specifically, to evaluate the mutual information, we apply Mercer expansion [21] as a basic mathematical tool. Furthermore, we show how to derive analytical solutions of the mutual information and provide an example. Finally, we introduce Fredholm determinant to provide a general expression of the mutual information. The expression based on Fredholm determinant has numerical evaluation scheme whose convergence is guaranteed in the literature.

A. Mercer expansion of a Gaussian random field

Assume that the source and the destination are two linear parallel length-L continuous antenna arrays placed along the z-axis and separated by a distance of d. Without loss of generality, we only care about the fields along the

z-axis, i.e., the vector field $\mathbf{E}(\mathbf{r})$ is only observed at the z-direction. In the following analysis, the scalars $J(s)$ and $E(r)$ represent $\mathbf{J}(\mathbf{s}) \cdot \hat{\mathbf{e}}_{\mathbf{z}}$ and $\mathbf{E}(\mathbf{r}) \cdot \hat{\mathbf{e}}_{\mathbf{z}}$, respectively. The second moments (autocorrelation) of $J(s)$ are denoted by $R_J(s, s'), s, s' \in [-L/2, +L/2]$.

Through this approach, we have built a simplified model which is highly similar to the classical model of random processes in information theory. Then it will be convenient for us to use the tools in classical information theory to solve and analyze the performance of EIT systems, with the transformation from the time-frequency domain to the space-wavenumber domain. Furthermore, we show that the tools for random processes can be extended to multi-dimensional random fields, which can solve other scenarios like rectangular transceivers in EIT.

The relationship between $J(s)$ and $E(r)$ can be described using the element in the upper left corner of the matrix $\mathbf{G}$ in (3), i.e., $G_{1,1}$, which can be derived as

$$g(r,s) = \frac{\mathrm{j} Z_0 e^{\mathrm{j} 2\pi \sqrt{x^2 + d^2}/\lambda}}{2\lambda \sqrt{x^2 + d^2}} \left[\frac{\mathrm{j}}{2\pi \sqrt{x^2 + d^2}/\lambda} \frac{d^2 - 2x^2}{x^2 + d^2} \right.$$
$$\left. + \frac{d^2}{x^2 + d^2} - \frac{1}{(2\pi/\lambda)^2 (x^2 + d^2)} \frac{d^2 - 2x^2}{x^2 + d^2} \right], \tag{24}$$

where $x = r - s$ and d is the distance between the parallel source and destination. Therefore, we have $E(r) = \int_0^L g(r,s) J(s) \mathrm{d}s$.

Since the noiseless received field is uniquely determined by the source and the deterministic channel, the autocorrelation function of the receiving electric field is expressed by the source autocorrelation $R_J(s, s')$ and the Green's function $g(r, s)$, written as

$$R_E(r, r') = \int_0^L \int_0^L g(r, s) R_J(s, s') g^*(r', s') \mathrm{d}s \mathrm{d}s'. \tag{25}$$

In order to evaluate $I(\mathbf{J}; \mathbf{Y})$, according to (23), a set of supremum-achieving testing functions $\{\phi_k\}_{k=1}^{+\infty}$ are used to convert the random field into a sequence of pairs of random variables $(\langle \phi_k | \mathbf{J} \rangle, \langle \phi_k | \mathbf{Y} \rangle)_{k=1}^{+\infty}$ for further information-theoretic analysis. Fortunately, the Mercer expansion provides us with a powerful tool to decompose random fields into random variables, as long as the autocorrelation function is given. The supremum-achieving of this expansion is based on the [15], as long as $I(\mathbf{J}; \mathbf{Y}) < \infty$. According to Mercer's theorem [21], the Mercer expansion of the electric field is

$$R_E(r, r') = \sum_{k=1}^{+\infty} \lambda_k \phi_k(r) \phi_k^*(r'), \tag{26}$$

where $\phi_k(r)$ are the solution functions to the integral eigen problem

$$\lambda_k \phi_k(r') = \int_0^L R_E(r, r') \phi_k(r) \mathrm{d}r; k > 0, k \in \mathbb{N}. \tag{27}$$

The eigenfunctions are orthonormal, satisfying the following equation

$$\int_0^L \phi_{k_1}(r) \phi_{k_2}^*(r) \mathrm{d}r = \mathbb{1}_{k_1 = k_2}, \tag{28}$$

where the eigenvalues $\{\lambda_k\}_{k=1}^{+\infty}$ are arranged in descending order. The field $E(r)$ can be decomposed by $E(r) = \sum_{k=1}^{\infty} \xi_k \phi_k(r)$, where $\mathbb{E}[\xi_{k_i} \xi_{k_j}^*] = \lambda_{k_i} \mathbb{1}_{i=j}$. In fact, other orthogonal bases can be chosen instead of $\phi_k(r)$, but only Mercer expansion can guarantee that ξ_{k_i} and ξ_{k_j} are uncorrelated, which makes the mutual information of the decomposed subchannels additive.

Similarly, for the noise field, the Mercer expansion also exists and we denote the eigenfunctions by $\phi_k'(r)$ and the eigenvalues by λ_k'. Now we further assume that the noise field is white Gaussian[2] with the noise power spectral density $n_0/(2\sqrt{2\pi})$. This assumption greatly simplifies the analysis, because the Gaussian white noise has a simple autocorrelation function $R_N(r, r') = \frac{n_0}{2}\delta(r - r')$, which means that the integral equation $\frac{n_0}{2}\phi_k(r') = \int_0^L R_N(r, r')\phi_k(r)\mathrm{d}r; k > 0, k \in \mathbb{N}$ holds for arbitrary integrable function $\phi_k(r)$. Therefore, the received electric field and the noise[3] can be expanded on the same orthogonal bases $\phi_k(r)$, and the mutual information between the electric field E and the received field Y can be expressed by the eigenvalues of (27)

$$I(E; Y) = \sum_{k=1}^{+\infty} \log \left(1 + \frac{\lambda_k}{n_0/2} \right). \tag{29}$$

The rigorous proof is provided by [23, **Theorem 1**].

Remark 3. *For simplicity, we focus on the linear transceivers in the paper, so as to use the tools and conclusions in the classical theory of stochastic process. However, the analyzing schemes here are not limited to the scenario with linear transceivers, but can be extended to other scenarios like two-dimensional surfaces. The conclusions in stochastic processes should also be extended to those in random fields. For example, if the source is a rectangular surface with current density $J(\mathbf{s}) = J(s_x, s_y)$ and the destination receives electric field $E(\mathbf{r}) = E(r_x, r_y)$ in a rectangular surface. We can also derive the autocorrelation function $R_E(\mathbf{r}, \mathbf{r}')$ of the electric field. Since the rectangular surface is a compact separable metric space if we define the area as the measure on the space, we can perform Mercer expansion on random field $E(r_x, r_y)$ according to [21]. The expansion is similar to (26) as $R_E(\mathbf{r}, \mathbf{r}') = \sum_k \lambda_k \phi_k(\mathbf{r}) \phi_k(\mathbf{r}')$. Therefore, mutual information expression and the corresponding derivation can also be done under the scenario with two-dimensional*

---

[2]In the next section, we will deal with the much more complicated case where the noise is Gaussian but not necessarily white.

[3]The Gaussian white noise field is no longer a "proper" random field [22]. It cannot be specified by the finite-dimensional distribution function. Instead, it can only be defined by its projections onto a complete orthonormal set of basis functions.

transceivers. For three-dimensional transceivers, the extension is similar to the two-dimensional cases. Furthermore, if we want to consider the vector-valued fields as $\mathbf{J}(\mathbf{s})$ *and* $\mathbf{E}(\mathbf{r})$*, we can decompose* $\mathbf{E}(\mathbf{r})$ *according to the extension of the Mercer theorem [24].*

From (29) we have derived the mutual information between the induced electric field and the received field. One interesting phenomenon is that although J and E may not obey a one-to-one correspondence, the equation that $I(E;Y) = I(J;Y)$ still exists. The reason that J and E may not be one-to-one correspondent is when we do Fourier transform on both sides, we have $\mathcal{F}[E] = \mathcal{F}[g]\mathcal{F}[J]$. Since $\mathcal{F}[g]$ has two zeros on $\kappa = \pm\kappa_0$, the corresponding two points in $\mathcal{F}[J]$ can not be solved. Physically this means that the electromagnetic wave with $\kappa = \pm\kappa_0$ can only transmit along the line of the source and can not be received by the destination.

Next we will show why $I(J;Y) = I(E;Y)$. This is from the equality that $I(J;Y) + I(Y;E|J) = I(Y;E) + I(Y;J|E)$. On the one hand, we can conclude that $I(J;Y) \leqslant I(E;Y)$ by the data processing inequality [25], because $J \to E \to Y$ forms a Markov chain and $I(Y;J|E) = 0$. On the other hand, since $E(r) = \int_0^L g(r,s)J(s)\mathrm{d}s$ is a deterministic function, we have $I(Y;E|J) = 0$, which leads to $I(J;Y) \geqslant I(E;Y)$. Therefore, we have $I(J;Y) = I(E;Y)$.

Remark 4. *Our analysis here from the mutual information perspective is compatible with existing works about the capacity. The existing works decompose the current density and electric field on a set of bases, which is similar to the SVD decomposition on* $\mathbf{H}\mathbf{H}^\dagger$*, where* $\mathbf{H}$ *is the channel matrix in discrete MIMO systems [26]. Our work decomposes the current density and electric field according to their autocorrelation function, which finds the bases for the field with given autocorrelation function. This scheme is similar to the decomposition of the covariance matrix to obtain the mutual information between random vectors, since* $I(X;X+N) = \log\det\frac{K_X + K_N}{K_N} = \sum_{i=1}^m (1 + \frac{\lambda_m}{n_0/2})$*, where* λ_k *is the eigenvalue of* K_X *and* $K_N = (n_0/2)\mathbf{I}$*.*

It is obvious that when fixing the field distribution which maximizes the mutual information, the capacity between continuous transceivers can be obtained. This conclusion is an extension from the discrete matrices to continuous operators. For discrete matrices we have $C = \max_{K_J}\log\det\frac{HK_JH^\mathrm{H} + K_N}{K_N}$ *with respect to* $\mathrm{Tr}(K_J) \leqslant P$*. The optimal solution is obtained by using singular value decomposition and Karush–Kuhn–Tucker (KKT) theorem. For operators, Mercer expansion is utilized instead of the singular value decomposition. Through this approach, an operator can be viewed as an infinitely large matrix [2], and the optimization problem can be solved by extended KKT theorem [27] for countably infinite variables. The result is in the form of water-filling method.*

The added values of the mutual information obtained in the paper can be listed as follows: 1) we can derive the mutual information under arbitrary given field distribution. If the field distribution can only be chosen from several given modes instead of the best one when capacity is achieved, the mutual information analyzing scheme can help to find the performance bound. 2) closed-form expression and numerical calculation scheme whose convergent rate is guaranteed can be obtained from the scheme we adopted, which will be introduced in the following subsections. 3) mutual information under colored noise field can be obtained and the corresponding numerical calculation scheme is proposed, which will be presented in **Section. V**.

B. Analytic expression based on autocorrelation function with rational spectrum

From the above subsection, we have

$$I(J;Y) = I(E;Y) = \sum_{k=1}^{+\infty} \log\left(1 + \frac{\lambda_k}{n_0/2}\right). \quad (30)$$

This formula is hard to evaluate, since there are infinite number of λ_K to calculate. In this part, we will first show how to derive an analytic expression without calculating λ_k in [15]. Then, we will propose a specific scenario and derive the corresponding analytic expression of mutual information.

For $R_E(r,r') = \int_{-\infty}^{\infty} e^{\mathrm{j}(r-r')x}\frac{|Q(\mathrm{j}x)|^2}{|P(\mathrm{j}x)|^2}\mathrm{d}x$, where $Q(\mathrm{j}x) = \sum_{i=0}^n b_i(\mathrm{j}x)^i$ and $P(\mathrm{j}x) = \sum_{i=0}^n a_i(\mathrm{j}x)^i$, the eigen-problem can be expressed by $\lambda\phi(r) = \int_0^L \int_{-\infty}^{\infty} e^{\mathrm{j}(r-r')x}\frac{|Q(\mathrm{j}x)|^2}{|P(\mathrm{j}x)|^2}\mathrm{d}x\phi(r')\mathrm{d}r'$. By performing $P(\frac{\mathrm{d}}{\mathrm{d}r})P(-\frac{\mathrm{d}}{\mathrm{d}r})$ on the integral equation, we have

$$P(\frac{\mathrm{d}}{\mathrm{d}r})P(-\frac{\mathrm{d}}{\mathrm{d}r})\phi(r) = \frac{2\pi}{\lambda}Q(\frac{\mathrm{d}}{\mathrm{d}r})Q(-\frac{\mathrm{d}}{\mathrm{d}r})\phi(r). \quad (31)$$

By performing $\frac{\mathrm{d}^k}{\mathrm{d}r^k}P(\frac{\mathrm{d}}{\mathrm{d}r})$ $(k = 0, \cdots, n-1)$ on the integral equation, we have boundary conditions

$$\frac{\mathrm{d}^k}{\mathrm{d}r^k}P(\frac{\mathrm{d}}{\mathrm{d}r})\phi(r) = \frac{1}{\lambda}\int_0^L \int_{-\infty}^{\infty} e^{\mathrm{j}(r-r')x}\frac{|Q(\mathrm{j}x)|^2(\mathrm{j}x)^k}{P(-\mathrm{j}x)}\mathrm{d}x$$
$$\phi(r')\mathrm{d}r', \quad (32)$$

where r is set on 0 and L to obtain the boundary conditions.

By solving (31) and substituting the solutions into the boundary conditions in (32), we have $g(z) = 0$, which contains $1/\lambda_k$ as its roots [15]. By deleting useless solutions in $g(z) = 0$, we can obtain $f(z) = f(0)\prod_{k=1}^{\infty}(1 - z\lambda_k)$ as a direct consequence of the Hadamard's factorization theorem [20]. Then, from (30), we can directly obtain the mutual information without solving every λ_k. The mutual information can be expressed by

$$I = \log\prod_{k=1}^{+\infty}\left(1 + \frac{\lambda_k}{n_0/2}\right) = \log\frac{f(-\frac{1}{n_0/2})}{f(0)}. \quad (33)$$

For simple autocorrelation functions like the exponential autocorrelation function we have discussed in the paper, such analysis is much easier than deriving all λ to obtain the mutual information. Since any correlation function can be approximated by a function with rational spectrum, this analyzing procedure is extendable to a wide range of scenarios.

Next, we will introduce a special case to show how this analyzing procedure works. The special case is when the electric field has exponential autocorrelation function as $R_E(r, r') = Pe^{-\alpha|r-r'|}$. We choose this special case not only because it is a very simple case where a stochastic process has a rational spectrum, but also because it has some practical meanings. For example, it is a widely used model when analyzing the scattering field from a surface [28]. If some surfaces in the channel scatter the field, the autocorrelation function may behave like exponential functions. The parameter $P\,[\mathrm{V}^2/\mathrm{m}^2]$ determines the power of the received field, and the parameter $\alpha\,[1/\mathrm{m}]$ controls the correlation of the received field. If we want to solve the eigenvalues of the integral equation (27), we can obtain the solutions from [23] as $\lambda_k = \frac{2\alpha P}{\alpha^2 + \omega_k^2}$ and $\phi_k(r) = \frac{1}{Z_k}(\omega_k \cos(\omega_k r) + \alpha \sin(\omega_k r))$, where Z_k are the normalization constants to ensure the orthonormality (28), and the resonant frequencies ω_k [23] are the positive solutions to the transcendental equation

$$2\arctan(\omega_k/\alpha) = k\pi - \omega_k L, \ k \geq 1, \ k \in \mathbb{Z}. \quad (34)$$

Although we can obtain λ_k, we will follow the approach in [15] to derive the mutual information and use the λ_k derived to verify the correctness. According to (31) we have

$$\phi''(r) - (\alpha^2 - \frac{2P\alpha}{\lambda})\phi(r) = 0. \quad (35)$$

According to (32) we have the boundary conditions $\phi'(0) - \alpha\phi(0) = 0$ and $\phi'(L) + \alpha\phi(L) = 0$. By solving (35) and substituting it into the boundary conditions, we have $g(z) = 2\alpha\omega\cos(\omega L) - (\omega^2 - \alpha^2)\sin(\omega L) = 0$, where $\omega^2 = \frac{2\alpha P}{\lambda} - \alpha^2 = 2\alpha Pz - \alpha^2$.

Since $z = \alpha/(2P)$ is a root of $g(z)$ but an exceptional root of the boundary conditions, we remove it from $g(z)$ to construct another entire function $f(z)$ directly from $g(z)$. The f function has the following form $f(z) = g(\sqrt{2\alpha Pz - \alpha^2})/\sqrt{2\alpha Pz - \alpha^2}$. It can be proved that all the zeros of $g(\omega)$ lie on the real axis, so all zeros of the entire function $f(z)$ are given by $1/\lambda_k$, arranged in the ascending order on the positive real axis.

Therefore, from (33), the mutual information is then explicitly expressed as

$$
\begin{aligned}
I = \log\Bigg(&\cosh(\alpha L\sqrt{1 + 4P/n_0\alpha}) + \frac{1 + 2P/n_0\alpha}{\sqrt{1 + 4P/n_0\alpha}} \\
&\sinh(\alpha L\sqrt{1 + 4P/n_0\alpha})\Bigg) - \alpha L.
\end{aligned}
$$
$$(36)$$

This analyzing scheme for the linear transceivers can be extended to other scenarios, e.g., circular transceivers. The integral problem can be transferred from (27) to

$$\lambda_k\phi_k(\theta') = \int_0^{\theta_0} R_E(\theta, \theta')\phi_k(\theta)\mathrm{d}\theta; k > 0, k \in \mathbb{N}, \quad (37)$$
$$\theta \in [0, \theta_0], \theta' \in [0, \theta_0],$$

where $\theta_0 \in (0, 2\pi)$. The assumption that $\theta_0 < 2\pi$ is to avoid the boundary condition $\phi^{(n)}(0) = \phi^{(n)}(2\pi)$. When $\theta_0 < 2\pi$, the analyzing procedure of this model has no difference from the procedures mentioned above. The mutual information of the scenario when $\theta_0 = 2\pi$ can be viewed as the limitation $\lim_{\theta \to 2\pi} I_\theta$.

C. General expression based on the Fredholm determinant

In the previous section, fortunately, we have found an explicit expression of the mutual information between a pair of correlated random fields. The restriction on the random fields is that they should be the observation of stationary random fields in a spatial region, which ensures the existence of the spectrum of the random fields. However, in general, the condition that $R_E(r, r')$ only relies on $r - r'$ does not exist. Therefore, we cannot always find a simple closed-form analytic function whose zeros are exactly all the eigenvalues of a general autocorrelation operator. But such an analytic function always exists mathematically, which can be explicitly constructed from the *Fredholm determinant* [29] of a bounded invertible operator in the form $\mathbf{1} + T$. In this section, we will employ the Fredholm determinant to provide an analytic solution to the mutual information between the fields at the source and the destination, which equals the result from Mercer expansion but has good analytical properties and reliable numerical calculation schemes.

Let G be the group of bounded operators of the form $\mathbf{1} + T$ on a Hilbert space H, where T is a trace-class operator [30], and $\mathbf{1}$ is the identity operator. The Fredholm determinant of $\mathbf{1} + T \in G$ is defined by the following infinite series:

$$\det(\mathbf{1} + T) := \sum_{k=0}^{+\infty} \mathrm{Tr}(\Lambda^k T), \quad (38)$$

where $\Lambda^k T$ is the k-th exterior power [31] of the bounded operator T on $\Lambda^k H$, and $\Lambda^k H$ denotes the k-th exterior product of H. Thus, an analytic function $f(z) := \det(\mathbf{1} + zT)$, $z \in \mathbb{C}$ can be naturally constructed, from which the mutual information can be evaluated.

Remark 5. *The Fredholm determinant* $\det(\cdot)$ *is a homomorphism of* G *into the multiplicative group of the complex numbers* $\mathbb{C}$, *i.e.,* $\det(AB) = \det(A)\det(B)$, $\forall A, B \in G$, *which is similar to the determinant of a square matrix. Furthermore, suppose the trace-class operator* T *has eigenvalues* λ_k, *then* $\det(\mathbf{1} + T) = \prod_k(1 + \lambda_k)$.

From the above remark, we can conclude that the Fredholm determinant is a nearly-perfect parallel of the matrix determinant. Similar to the fact that the ergodic MIMO capacity can be expressed in a determinant form, the mutual information between continuous regions can also be expressed by the Fredholm determinant. Specifically, starting from (29) which is obtained by using Mercer expansion, the mutual information with finite-length destination has already been expressed in the infinite product form $I = \log \prod_{k=1}^{+\infty} \left(1 + \frac{\lambda_k}{n_0/2}\right)$. Now we try to express I by the Fredholm determinant of some operator. Since the autocorrelation function $R_E(r, r')$ can be considered as an integral operator T_E:

$$(T_E\phi)(r) = \int_0^a R_E(r, r')\phi(r')\mathrm{d}r', \qquad (39)$$

according to **Remark 5**, we can take the Fredholm determinant of the operator $(\mathbf{1} + T_E/(n_0/2))$ to compute the infinite product $\prod_k(1 + \lambda_k/(n_0/2))$, where $\{\lambda_k\}$ are the eigenvalues of T_E:

$$I = \prod_k \left(1 + \frac{\lambda_k}{n_0/2}\right) = \log \det \left(\mathbf{1} + \frac{T_E}{n_0/2}\right). \quad (40)$$

Note that the eigenvalues λ_k of T_E are equivalent to the eigenvalues of the eigen problem (27), and the mutual information I is exactly the function value of the induced analytic function $f(z) := \det(\mathbf{1} + zT_E)$ evaluated at $z = 1/(n_0/2)$.

Our analysis in the following parts is based on Fredholm determinant for operator T with kernels $K(r, r')$, which correspondes to the linear transceivers. However, the expression of Fredholm determinant and the related analyzing schemes are not restricted to this scenario. The linear transceivers can be naturally extended to rectangular transceivers or other shapes, since the Fredholm determinant is well defined for any trace class operator T [29].

D. Numerical calculation method of Fredholm determinant

Since the Fredholm determinant is a widely-used tool in physics [32], its numerical properties have been thoroughly studied. As a result, we can provide a numerical calculation method for the mutual information I and the corresponding convergence analysis according to [33]. Inspired by the numerical integral method, the operator T_E can be approximated by discrete summation

$$(T_E\phi)(r) := \int_0^a K(r, r')\phi(r')\mathrm{d}r' \approx \sum_{j=1}^m w_j K(r, r'_j)\phi(r'_j),$$
$$(41)$$

where $K = R_E : [0, L]^2 \to \mathbb{C}$ is called the kernel of the integral operator T_E, and the integral equation can then be discretized into

$$\sum_{j=1}^m w_j K(r_i, r_j)\phi_k(r_j) = \lambda_k \phi_k(r_i). \qquad (42)$$

When $K(r_i, r_j) = \delta(r_i, r_j)$, the identity operator $\mathbf{1}$ is discretized into an identity matrix $\mathbf{1} \sim \sum_{j=1}^m \mathbb{1}_{i=j}$. The Fredholm determinant can then be approximated by

$$I \approx \mathrm{logdet} \left(\mathbb{1}_{i=j} + \frac{w_j K(r_i, r_j)}{n_0/2}\right)_{i,j=1}^m$$
$$= \mathrm{logdet} \left(\mathbb{1}_{i=j} + \frac{w_j^{1/2} K(r_i, r_j) w_j^{1/2}}{n_0/2}\right)_{i,j=1}^m. \qquad (43)$$

The convergence of this discretization method is proven in [33], which is dependent on the way of discretizing the integral operator T_E. For example, Gauss-Legendre quadrature has been proved to have satisfactory convergence properties, and achieves the highest degrees of precision. A Gauss-Legendre quadrature with n points has degrees of precision $2n - 1$ and order $2n$. The Gauss-Legendre quadrature on the integration interval $[-1, 1]$ is based on the Legendre orthogonal polynomials $P_0(x) := 1$ and $P_n(x) := \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[(x^2 - 1)^n\right]$.

The Gauss points $x_1, \ldots, x_n$ are the zeros of $P_n(x)$. The weights $A_1, \cdots, A_n$ can be derived as follows

$$A_k = \frac{2}{n} \frac{1}{P_{n-1}(x_k) P'_n(x_k)}. \qquad (44)$$

The convergence of (43) can be derived as

$$\left| Q(f) - \int_0^a f(x)\mathrm{d}x \right| \leqslant c_k a^{k+1} v^{-k} \left\| f^{(k)} \right\|_{L^\infty(0,a)}, \qquad (45)$$

for $f \in C^{k-1,1}([0, a])$, quadrature rule Q of order $v \geqslant k$ with positive weights. The reference [33, **Theorem 6.2**] shows that

$$|d_Q(z) - d(z)| \leqslant c_k 2^k a^k v^{-k} \Phi(|z| \, a \, \|K\|_k), \qquad (46)$$

where $d(z) := \det(I + zT)$, $K \in C^{k-1,1}([0, a]^2)$, quadrature rule Q of order $v \geqslant k$ with positive weights, $\|K\|_k = \max_{i+j \leqslant k} \left\| \partial_1^i \partial_2^j K \right\|_{L^\infty}$ and $\Phi(z) = \sum_{n=1}^{+\infty} \frac{n^{(n+2)/2}}{n!} z^n$.

The convergence of the numerical calculation scheme guarantees the reliability of the scheme. Therefore, the Fredholm determinant is not only a mathematical representation of the mutual information, but also provides a useful tool to calculate the mutual information that can be obtained from the received noisy field.

V. MUTUAL INFORMATION WITH FINITE-LENGTH TRANSCEIVERS UNDER COLORED NOISE FIELD

In the above section, we have discussed the mutual information between finite-length source and destination under the assumption of a white noise field, which is a common model adopted by recent works. However, the white noise model is only a simplified model that comes from the assumption of traditional MIMO modeling. In this section, we consider other mechanisms which have been analyzed in Section III that may contribute to the noise, and analyze the mutual information of the proposed model. A general numerical calculation procedure is proposed of the mutual information, which can be viewed as the extension of the numerical scheme in Section IV.

A. Mutual information based on operator analysis

In Section. III-B, we have derived a non-white noise model, which means that the noise field has non-zero spatial correlation coefficients. Following the same procedure, we assume that the covariance kernel of the noise field is $K_N(r, r') \neq P\delta(r - r')$, and the corresponding operator is $T_N : \mathscr{L}^2(V_R) \to \mathscr{L}^2(V_R)$. The operator of the received noisy field $Y = E + N$ is denoted by T_Y. Here we can perform Mercer expansion on the electric field $E(r)$ and the noise field $N(r)$ similar to Section IV, which leads to $R_E(r, r') = \sum_k \lambda_k \phi_k(r) \phi_k^*(r')$, $E(r) = \sum_k \xi_{E,k} \phi_k(r)$, $R_N(r, r') = \sum_k \lambda_k' \psi_k(r) \psi_k^*(r')$ and $N(r) = \sum_k \xi_{N,k} \psi_k(r)$. The received field $Y(r)$ can also be decomposed as $Y(r) = \sum_k \xi_{Y,k} \varphi_k(r)$. However, different from the white noise case where $\psi(r)$ can be arbitrary orthogonal bases, the expansion of non-white noise is confirmed, which means that the decomposition of $E(r)$, $N(r)$ and $Y(r)$ will lead to different sets of eigenfunctions and the mutual information can not be directly obtained like (29).

Inspired by [15], we can construct two spaces H_1 and H_2 which are extended by $\xi_{E,k}$ and $\xi_{Y,k}$ separately. The inner product in these spaces is defined as the cross-correlation between two random variables in the space. The mutual information between $E(r)$ and $Y(r)$ is actually the difference between the space H_1 and H_2 and is determined by the angle between them according to [15]. Specifically, the mutual information can be evaluated by finding new sets of bases $\hat{\xi}_{E,k}$ and $\hat{\xi}_{Y,k}$ in H_1 and H_2 which satisfy $\mathbb{E}[\hat{\xi}_{E,k_1} \hat{\xi}_{Y,k_2}^*] = 0$, $\mathbb{E}[\hat{\xi}_{E,k_1} \hat{\xi}_{E,k_2}^*] = 0$ and $\mathbb{E}[\hat{\xi}_{Y,k_1} \hat{\xi}_{Y,k_2}^*] = 0$ when $k_1 \neq k_2$. These $\hat{\xi}$ can be viewed as the linear combinations from ξ decomposed from $E(r)$ and $Y(r)$. Moreover, they correspond to the projection operators $B_1 = P_1 P_2$ and $B_2 = P_2 P_1$, where P_1 is the projection onto the space H_1 and P_2 is the projection onto the space H_2. We have

$$I = -\sum_k \log(1 - \frac{\mathbb{E}[\hat{\xi}_{E,k} \hat{\xi}_{Y,k}^*]^2}{\mathbb{E}[\hat{\xi}_{E,k} \hat{\xi}_{E,k}^*] \mathbb{E}[\hat{\xi}_{Y,k} \hat{\xi}_{Y,k}^*]})$$
$$= -\mathrm{logdet}(\mathbf{1} - P_1 P_2) = -\mathrm{logdet}(\mathbf{1} - P_2 P_1). \quad (47)$$

Since P_1 is the projection operator, we have $\mathbb{E}[P_1 \xi_E * \xi_E^*] = \mathbb{E}[\xi_Y \xi_E^*]$. By defining the operator T_D and $T_{D'}$ which represent the mutual correlation between $E(r)$ and $Y(r)$, we can obtain P_1 and P_2, thus obtaining the mutual information. Here we define T_D by $(T_D \psi)(r) := \int_{V_R} \mathbb{E}[Y(r) E^*(r')] \psi(r') \mathrm{d}r'$ and $T_{D'}$ by $(T_{D'} \psi)(r) := \int_{V_R} \mathbb{E}[Y^*(r) E(r')] \psi(r') \mathrm{d}r'$. They can also be represented by infinite-size matrix to be easier to understand, as $T_D \sim \|\mathbb{E}[\langle \phi_j | Y \rangle \langle E | \phi_i \rangle]\|_{i,j=1}^m$. Through this approach we have $P_1 T_E = T_D$ and $P_2 T_Y = T_{D'}$, which means $P_1 = T_D T_E^{-1}$ and $P_2 = T_{D'} T_Y^{-1}$. Therefore we have

$$I(J; Y) = I(E; Y) = -\mathrm{logdet}(\mathbf{1} - T_D T_E^{-1} T_{D'} T_Y^{-1})$$
$$= -\mathrm{logdet}(\mathbf{1} - T_{D'} T_Y^{-1} T_D T_E^{-1}), \quad (48)$$

This result inspired by [15] is about stochastic process over the real line and the sampling of it is a single variable, which coincides with the model of parallel linear transceivers that we used. It is necessary to mention that such expression based on operator is not restricted to this simplified scenario, but can be extended to random field over arbitrary manifolds and the sampling of it can be a random vector, which has theoretical bases from [34]. By using random field on two-dimensional surfaces we can extend the model from linear transceivers to rectangular or circular transceivers. By using random vector as the sampling result we can extend the scalar wave field to three-dimensional wave field.

In our assumption, the noise field N is independent of the noiseless electric field E, which means $T_D = T_{D'} = T_E$. The mutual information in (48) is then simplified to

$$I(E; Y) = -\mathrm{logdet}(\mathbf{1} - T_D T_E^{-1} T_{D'} T_Y^{-1})$$
$$= -\mathrm{logdet}(\mathbf{1} - T_E (T_E + T_N)^{-1}). \quad (49)$$

In the special case when the noise field is assumed to be a white Gaussian random field, the kernel corresponding to the noise operator is $K_N(r, r') = \frac{n_0}{2} \delta(r - r')$. Since for arbitrary $\phi(r)$ the following integral equation holds $\int_0^a K_N(r, r') \phi(r') \mathrm{d}r' = \frac{n_0}{2} \phi(r)$, we have $T_N |\phi\rangle = \frac{n_0}{2} |\phi\rangle$. Therefore, the difference between T_N and the identity operator $\mathbf{1}$ is only a constant factor, i.e., we can represent T_N by $T_N = \frac{n_0}{2} \mathbf{1}$. The mutual information in (49) can be further simplified to

$$I(E; Y) = -\mathrm{logdet}(\mathbf{1} - T_E (T_E + T_N)^{-1})$$
$$= \mathrm{logdet}\left(\mathbf{1} + \frac{T_E}{n_0/2}\right), \quad (50)$$

which coincides with the formula derived in (40). Therefore, the mutual information under white noise model can be viewed as a special case of the result in this section.

B. Numerical calculation of the mutual information

As shown in the previous subsection, we can express the mutual information between the electric field and the received noisy field by (49). Now we want to find a numerical algorithm to calculate the mutual information, and show how the noise field affects the information we can obtain from the received field.

In Section IV, we have clarified that the numerical approximation of the mutual information converges under the white-noise scenario. For the colored noise scenario, if we view the operator $T_E (T_E + T_N)^{-1}$ as a whole, we can directly utilize the conclusion of convergence in the case of white noise. However, we can not directly discretize $T_E (T_E + T_N)^{-1}$ because the inverse operator $(T_E + T_N)^{-1}$ is hard to obtain. Therefore, we discretize T_E and $T_E + T_N$ separately instead of directly discretizing $T_E (T_E + T_N)^{-1}$. The algorithm of the discretization is as **Algorithm 1**, where the kernel of T_N equals $R_{N,\mathrm{non-white}}(r, r') + \sigma_{\mathrm{white}}^2 \delta(r - r')$.

Algorithm 1 Numerical calculation scheme for the mutual information

Input:

    $R_E(r, r')$ % the kernel of the operator T_E

    $R_{N,\text{non-white}}(r, r')$

    σ_{white}^2

    a % the length of the destination

    m % the size of the discretized matrix

Output:

    I_{approx} % approximation of the mutual information

1: $(w_1^m, r_1^m) \leftarrow \text{Quadrature}(0, a, m)$

2: $a_0 \leftarrow 0$

3: **for** $1 \leq i \leq m$ **do**

4:     $a_i \leftarrow a_{i-1} + w_i$

5: **for** $1 \leq i \leq m$ **do**

6:     **for** $1 \leq j \leq m$ **do**

7:         $\mathbf{K}_{E,i,j} \leftarrow \int_{a_{i-1}}^{a_i} \int_{a_{j-1}}^{a_j} K_E(x,y)\mathrm{d}y\mathrm{d}x \approx w_i w_j R_E(r_i, r_j)$

8:         $\mathbf{K}_{N,i,j} \leftarrow \int_{a_{i-1}}^{a_i} \int_{a_{j-1}}^{a_j} K_N(x,y)\mathrm{d}y\mathrm{d}x \approx w_i w_j R_{N,\text{non-white}}(r_i, r_j) + w_i \sigma_{\text{white}}^2 \mathbb{1}_{i=j}$

9: $\mathbf{K}_{\text{sep}} \leftarrow \mathbf{K}_E(\mathbf{K}_E + \mathbf{K}_N)^{-1}$

10: $I_{\text{approx}} \leftarrow -\text{logdet}(\mathbf{I} - \mathbf{K}_{\text{sep}})$

11: **return** I_{approx}

Remark 6. *The discretization algorithm of the mutual information not only has the mathematical meaning but also has physical correspondence. Since* $-\text{logdet}(\mathbf{I} - \mathbf{K}_{\text{sep}})$ *in Algorithm 1 equals* $\log\det(\mathbf{K}_E + \mathbf{K}_N)/\det(\mathbf{K}_N)$, *which is the form of the mutual information of multiple Gaussian channel model. If we express* $\mathbf{K}_{E,i,j}$ *and* $\mathbf{K}_{N,i,j}$ *by* $\mathbb{E}\left[\int_{a_{i-1}}^{a_i} E(x)\mathrm{d}x \int_{a_{j-1}}^{a_j} E^*(y)\mathrm{d}y\right]$ *and* $\mathbb{E}\left[\int_{a_{i-1}}^{a_i} N(x)\mathrm{d}x \int_{a_{j-1}}^{a_j} N^*(y)\mathrm{d}y\right]$ *respectively, it is easy to find that the mutual information after discretization can be viewed as using several patch antennas in the destination region which receives the integral of the electric field. Therefore, the convergence of the numerical approximation of the mutual information shows how the discrete destinations approach the continuous one. Here we use Gauss-Legendre quadrature rule because it converges faster than the equally-spaced trapezoid or rectangular quadrature rules.*

The algorithm above that separately discretize T_E and $T_E + T_N$ is used to numerically calculate the information obtained from the received electric field under non-white noise. The accuracy is guaranteed by the quadrature rule. Based on the numerical scheme, we will provide some numerical results of the mutual information. Here we assume that the source and destination both have length L, the current density $J(s)$ on the source has the autocorrelation function $R_J(r, r') = \delta(r - r')$, which means that no channel state information at the transmitter (CSIT) is available. Then, the electric field on the destination can

be derived by

$$
\begin{aligned}
R_E(r, r') &= \int_0^L \int_0^L g(r, s)\delta(s - s')g^*(r', s')\mathrm{d}s\mathrm{d}s' \\
&= \int_0^L g(r, s)g^*(r', s)\mathrm{d}s. \quad (51)
\end{aligned}
$$

The noise field is considered a mixture of the measurement noise and the radiation interference. We denote the mixed noise field[4] by $N(r) = \sigma_1 N_{\text{mea}}(r) + \sigma_2 N_{\text{int}}(r)$ and its autocorrelation function by $R_N(r, r')$. According to (18), we have

$$
\begin{aligned}
&R_N(r, r') \\
&= \sigma_1^2 \delta(r - r') + 2\sigma_2^2 \left(\frac{\sin(\kappa r - \kappa r')}{(\kappa r - \kappa r')^3} - \frac{\cos(\kappa r - \kappa r')}{(\kappa r - \kappa r')^2} \right).
\end{aligned}
\quad (52)
$$

Now we set the length L of the source and destination to $2\,\text{m}$ and the wavelength λ of the electromagnetic wave to $0.25\,\text{m}$. The distance between the source and the destination is fixed to $1\,\text{m}$. Therefore, the half-wavelength sampling on the destination has 16 sampling points. The mutual information obtained from the discretization is expressed in Fig. 3. In this figure, the power of the radiation interference is fixed to $\sigma_2^2 = 0.5\,[\text{V}^2/\text{m}^2]$, and the measurement noise is flexible. It is shown that when the power of the measurement noise is large, the mutual information converges quickly with the increasing number of discretization points. For example, when $\sigma_1^2 = 1\,[\text{V}^2/\text{m}]$ or $\sigma_1^2 = 10^{-1}\,[\text{V}^2/\text{m}]$, the information obtained from the continuous destination is almost the same as the information obtained from the 16-points discretization. However, when $\sigma_1^2\,[\text{V}^2/\text{m}]$ is very small, which means that the radiation interference plays a dominant role in the noise model, the gap between the continuous case and the discretized case becomes obvious. When $\sigma_1^2 = 10^{-5}\,[\text{V}^2/\text{m}]$, the information obtained from the continuous destination is 9% larger than the information obtained from the 16-points discretized destination. When $\sigma_1^2 = 10^{-10}\,[\text{V}^2/\text{m}]$, the information obtained from the continuous destination is 20.6% more than that obtained from the 16-points discretized destination.

Fig. 3 also demonstrates that the half-wavelength sampling is strictly suboptimal, since 16 discretization points fail to harvest all the mutual information provided by (49). Furthermore, the extent of such suboptimality depends on the ratio σ_1^2/σ_2^2 of the white noise power with respect to the non-white noise power. This is because the colored interference noise is more "structural", and lives in a lower-dimensional noisy subspace, while the white noise projects equally into any subspace. Thus, a smaller portion of white noise leads to more clean signal subspaces, which further results in more mutual information.

---

[4]Here the unit of σ_1^2 is $[\text{V}^2/\text{m}]$ and the unit of σ_2^2 is $[\text{V}^2/\text{m}^2]$. These parameters are used to keep the unit of $N(r)$ to be $[\text{V}/\text{m}]$. The units of σ_1^2 and σ_2^2 are different because the measurement noise comes from unideal measurement or data processing steps, while the radiated interference is strongly correlated to the wavelength of the electromagnetic field.
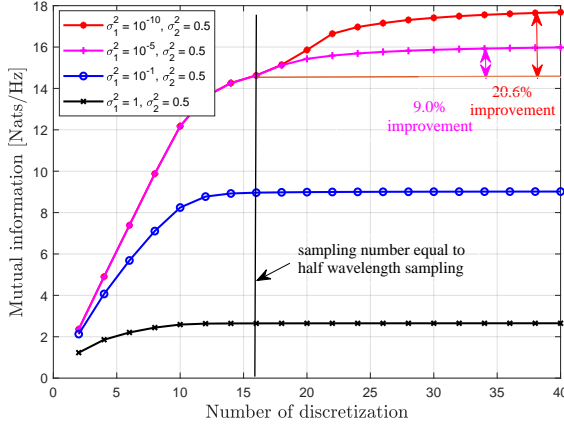
Fig. 3. Mutual information $I(L)$ in (49) calculated by applying **Algorithm 1**. The x-axis represents the number of discretization points, and the y-axis is the mutual information measured in nats. The mutual information value increases as the number of discretization points increases, and finally converges to the continuous-space mutual information defined by the supremum in (23).

VI. MUTUAL INFORMATION WITH INFINITE-LENGTH TRANSCEIVERS

In the above sections, we have analyzed the mutual information between finite-length transceivers and provided its expression based on the Fredholm determinant. To further strengthen the connection between the model of parallel linear transceivers and classical model in information theory, we introduce an ideal model in which the transceivers are of infinite length. A strong correspondence between this model in the spatial dimension and the classical signal model in the time dimension is built. Based on this correspondence, we utilize the analyzing methods in classical information theory to provide the mutual information and the capacity for this scenario.

A. Connection between the infinite-length transceivers and classical models in information theory

Since we extend the finite-length transceivers to infinite-length transceivers in our new ideal model, we have $E(r) = \int_{-\infty}^{+\infty} g(r,s)J(s)\mathrm{d}s$. The main difference between this model and the model with finite-length transceivers is that the current density and the electric field become random processes on the infinite-length spatial dimension, which is similar to the channel model $Y(t) = \int_{-\infty}^{+\infty} h(t,\tau)X(\tau)\mathrm{d}\tau$ in classical information theory. Similar to the classical information theory which often assumes that $X(\tau)$ and $Y(t)$ are stationary random processes, we assume that, the current density J and electric field E are also stationary stochastic processes in the space domain. The spatial stationary means that, $\mathbb{E}[J(s)J^*(s')]$ only depends on $\Delta s = s' - s$, and we can introduce the autocorrelation function as $R_J(\Delta s) = \mathbb{E}[J(s)J^*(s')]$ and $R_E(\Delta r) = \mathbb{E}[E(r)E^*(r')] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(r,s)R_J(\Delta s)g^*(r',s')\mathrm{d}s\mathrm{d}s'$.

The SSD can then be derived as

$$S_J(\kappa) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} R_J(\Delta s)e^{-\mathrm{j}\kappa\Delta s}\mathrm{d}\Delta s. \qquad (53)$$

The SSD of the electric field can then be derived by the SSD of the current density via the Fourier transform of the Green's function $G(\kappa)$:

$$S_E(\kappa) = 2\pi S_J(\kappa)|G(\kappa)|^2. \qquad (54)$$

B. Spatial spectrum analysis of the communication between transceivers

In this subsection, we will reveal how some important parameters affect the communication quality of the system with continuous source and destination from $G(\kappa)$, which is the Fourier transform of the Green function. We introduce **Lemma 1** to derive a closed-form solution for $G(\kappa)$.

Lemma 1. *The Fourier transform of the Green function in (24) can be expressed by $G(\kappa) = F_1(\kappa) * F_2(\kappa)$, where*

$$F_1(\kappa) = \frac{-\mathrm{j}Z_0 d^2}{4\pi\lambda}\left\{ \begin{array}{ll} \frac{\pi}{2}\mathrm{j}J_0(dm) - \frac{\pi}{2}Y_0(dm), & [0 < \kappa < \frac{2\pi}{\lambda}] \\ K_0(dm), & [\kappa > \frac{2\pi}{\lambda}] \end{array} \right. \qquad (55)$$

with $m = \sqrt{|(\frac{2\pi}{\lambda})^2 - \kappa^2|}$, and

$$\begin{aligned} F_2(\kappa) = &\ d\sqrt{\frac{\pi}{2}}e^{-d|\kappa|} + \frac{\mathrm{j}d\lambda}{2\pi}\sqrt{\frac{2}{\pi}}|\kappa|\,K_1(d\,|\kappa|) \\ &- \frac{\mathrm{j}\lambda}{\pi}\Big[\sqrt{\frac{2}{\pi}}K_0(d\,|\kappa|) - d\sqrt{\frac{2}{\pi}}|\kappa|\,K_1(d\,|\kappa|)\Big] \\ &- (\frac{\lambda}{2\pi})^2\frac{1}{2d}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|} \\ &+ (\frac{\lambda}{2\pi})^2\Big[\sqrt{\frac{\pi}{2}}\frac{2e^{-d|\kappa|}}{d} - \frac{1}{d}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|}\Big]. \end{aligned} \qquad (56)$$

Proof: See Appendix A. ∎

From the Fourier transform of the Green function in **Lemma 1**, we can find that the distance between the source and destination d deeply affects the behavior of $|G(\kappa)|$. Since $|G(\kappa)|$ depicts how the spectrum of the current on the source affects the spectrum of the electric field on the destination, larger values of $|G(\kappa)|$ will cause higher channel gain in the corresponding wavenumber bands.

In Fig. 4, we plot $|G(\kappa)|$ with different distances d while λ is fixed to $5\,\mathrm{m}$. We find that $|G(\kappa)|$ has a main lobe in the band $[-\kappa_0, \kappa_0]$, which means that the electromagnetic wave is a plane wave. When κ equals 0, the wavefront of the plane wave is parallel to the linear destination. The figure of $G(\kappa)$ has side lobes when κ falls out of the band $[-\kappa_0, \kappa_0]$, which means that the corresponding component of the electromagnetic wave is an evanescent wave and vanishes when the distance is larger than the wavelength. The channel with obvious side lobes can support broad bandwidth in the wavenumber domain, thus the DoF can be increased. We find that,
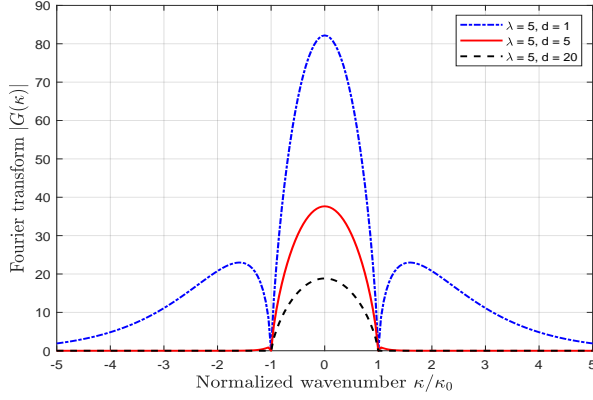
Fig. 4. Fourier transform of the Green function with different distances d.

when d decreases, the main lobe and the side lobes both have more energy. Therefore, the DoF and the gain per degree both increase. From this phenomenon, we conclude that with small d which is comparable with the wavelength, we can get more DoFs besides the power gain per degree, thus improving the channel capacity. This phenomenon is in the reactive near-field region, which only has theoretical meanings and is acceptable to be neglected in the current wireless communication scenarios.

C. Capacity between the transceivers

We suppose that the noise obeys the spatial additive white Gaussian noise (AWGN) model, which means that the autocorrelation function of the noise $R_N(\Delta s) = \sigma^2 \delta(\Delta s)$ and the SSD of the noise is $S_N(\kappa) = \frac{\sigma^2}{\sqrt{2\pi}}$. Here we follow [7] to assume that the power constraint[5] is similar to that of the traditional MIMO model, which constrains the sum of the squares of current excitations on the source antennas. Note that the integral of the square of current in the spatial domain can be transferred to the integral of its spectrum in the wavenumber domain based on the Parseval's theorem. Thus, the power consumption for the source can be bounded by

$$\int_{-\infty}^{+\infty} S_J(\kappa) \mathrm{d}\kappa \leqslant P. \qquad (57)$$

Inspired by the integral form of the Shannon capacity for colored AWGN noise channel [35, 36], the channel can be split to infinite narrow-band subchannels in the wavenumber domain. For each subchannel, the capacity is

---

[5]The radiation power which can be calculated from the Poynting vector is physically more meaningful, but the optimal field distribution is very hard to derive under this power constraint, so here we impose the power constraint on the SSD.

$\mathrm{d}C_{\kappa_1} = \frac{1}{2\pi}\log\left(1 + \frac{S_E(\kappa_1)}{S_N(\kappa_1)}\right)\mathrm{d}\kappa_1$. Therefore, the overall capacity is

$$C = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\log\left(1 + \frac{2\pi S_J(\kappa)}{S_N(\kappa)/|G(\kappa)|^2}\right)\mathrm{d}\kappa. \qquad (58)$$

To simplify the analysis, we introduce the equivalent noise N' with SSD $S_{N'}(\kappa) = S_N(\kappa)/(2\pi|G(\kappa)|^2)$. To derive the best SSD of the current density $S_J(\kappa)$, we obtain the following **Theorem 1**.

Theorem 1. *For the overall capacity in (58) and the power constraint for the current density in (57), the SSD of the best current density obeys $S_J(\kappa) = \left(\frac{1}{2\pi\mu} - S_{N'}(\kappa)\right)^+$, where μ is obtained by the power constraint.*

Proof: To obtain the optimal SSD $S_J(\kappa)$ that maximizes (58), we apply the variational calculus [37]. The Lagrange multiplier μ can be introduced to take into consideration the power constraint $\int S_J \mathrm{d}\kappa \leqslant P_J$:

$$\begin{aligned}\mathcal{L}(S_J, \mu) = &\frac{1}{2\pi}\int_{-\infty}^{+\infty}\log\left(1 + \frac{S_J(\kappa)}{S_{N'}(\kappa)}\right)\mathrm{d}\kappa \\ &- \mu\left(\int_{-\infty}^{+\infty} S_J(\kappa)\mathrm{d}\kappa - P_J\right).\end{aligned} \qquad (59)$$

Taking the variation of (59) we obtain

$$\delta C = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{1}{1 + \frac{S_J(\kappa)}{S_{N'}(\kappa)}}\frac{\delta S_J(\kappa)}{S_{N'}(\kappa)}\mathrm{d}\kappa - \mu\int_{-\infty}^{+\infty}\delta S_J(\kappa)\mathrm{d}\kappa. \qquad (60)$$

From (60) we obtain that the optimal solution for S_J should satisfy $\frac{1}{2\pi(S_J(\kappa) + S_{N'}(\kappa))} - \mu \equiv 0$. Taking the non-negative condition on S_J into consideration, we obtain $S_J(\kappa) = \left(\frac{1}{2\pi\mu} - S_{N'}(\kappa)\right)^+$. ∎

For example, we consider a set of parameters where the wavelength λ is $5\,\mathrm{m}$ and the distance d between the two parallel lines is $1\,\mathrm{m}$. We use AWGN model with the noise SSD $S_N(\kappa) = 90\,[\mathrm{V}^2/\mathrm{m}]$. The power constraint for the current density P is equal to $3\,[\mathrm{A}^2/\mathrm{m}^4]$. From **Theorem 1** we can derive the best SSD of the source current, as shown in Fig. 5. It is worth noticing that the scenario considered here is in the reactive near field, when the wavelength is comparable to the distance between the transceivers. The best spectrum density in the $[-\kappa_0, \kappa_0]$ band is nearly flat. This phenomenon coincides with Fig. 6 in [1], where discrete Fourier bases are chosen to decompose the channel instead of the continuous Fourier transform we adopted in this section.

VII. CONCLUSIONS

In this paper, we analyzed the mutual information and the capacity of electromagnetic fields based on random field theory. We first developed the system model of communication between two continuous regions based on random fields. The model based on random fields provided a general framework for solving the problems
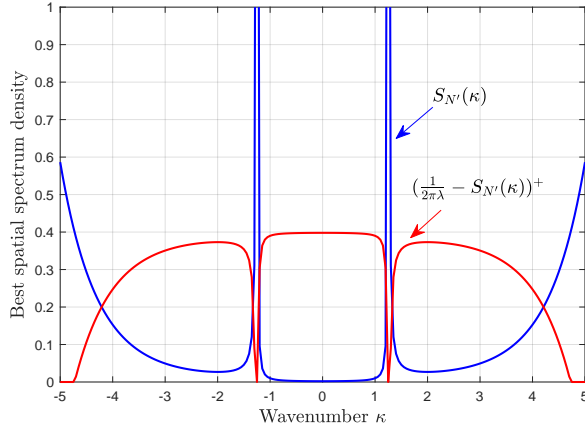
Fig. 5. Optimal SSD of the current density on the source based on variational calculus.

like mutual information and capacity for EIT. Then, we considered a simplified model with parallel finite-length linear source and destination and analyze the mutual information under white and colored noise fields. Numerical results confirmed the suboptimality of the half-wavelength sampling. Finally, we considered an ideal model with an infinite-length destination and obtained the mutual information and capacity through SSD. Moreover, the gain of DoF and mutual information in the reactive near-field region is shown using SSD.

Further work will focus on the analytical solutions of more general cases like non-linear sources and destinations. The correlation between the current density distribution on the source and the radiated power in the space is also necessary to be explored.

APPENDIX A
PROOF OF LEMMA 1

For the Green function in (24), we split it into the product of two functions $g(x) = f_1(x)f_2(x)$, where

$$f_1(x) = \frac{-\mathrm{j}Z_0\eta e^{\mathrm{j}2\pi\sqrt{x^2+d^2}/\lambda}}{2\lambda\sqrt{x^2+d^2}}, \tag{61a}$$

$$f_2(x) = \frac{d^2}{x^2+d^2} + \frac{\mathrm{j}}{2\pi\sqrt{x^2+d^2}/\lambda}\frac{d^2-2x^2}{x^2+d^2} \tag{61b}$$
$$- \frac{1}{(2\pi/\lambda)^2(x^2+d^2)}\frac{d^2-2x^2}{x^2+d^2}.$$

Utilizing the convolution theory, we can express $G(\kappa)$ by $G(\kappa) = F_1(\kappa) * F_2(\kappa)$, where $F_1(\kappa) = \mathscr{F}[f_1(x)]$ and $F_2(\kappa) = \mathscr{F}[f_2(x)]$.

For the Fourier transform of $f_1(x)$, we have

$$F_1(\kappa) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{-\mathrm{j}Z_0 e^{\mathrm{j}2\pi\sqrt{x^2+d^2}/\lambda}}{2\lambda\sqrt{x^2+d^2}}e^{\mathrm{j}\kappa x}\mathrm{d}x = A_1 + A_2, \tag{62}$$

where $A_1 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{-\mathrm{j}Z_0 e^{\mathrm{j}2\pi\sqrt{x^2+d^2}/\lambda}}{2\lambda\sqrt{x^2+d^2}}\cos\kappa x\mathrm{d}x$ and $A_2 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{Z_0 e^{\mathrm{j}2\pi\sqrt{x^2+d^2}/\lambda}}{2\lambda\sqrt{x^2+d^2}}\sin\kappa x\mathrm{d}x$. Since

$\frac{Z_0 e^{\mathrm{j}2\pi\sqrt{x^2+d^2}/\lambda}}{2\lambda\sqrt{x^2+d^2}}$ is an even function, it is obvious that A_2 equals 0. For A_1, we utilize [38, Eq. (3.876)] to get

$$A_1 = \frac{-\mathrm{j}Z_0 d^2}{4\pi\lambda}\left\{\begin{array}{ll}\frac{\pi}{2}\mathrm{j}J_0(dm) - \frac{\pi}{2}Y_0(dm), & [0<|\kappa|<\frac{2\pi}{\lambda}]\\ K_0(dm), & [|\kappa|>\frac{2\pi}{\lambda}]\end{array}\right. \tag{63}$$

where m denotes $\sqrt{\left|\left(\frac{2\pi}{\lambda}\right)^2 - \kappa^2\right|}$. Then we obtain the closed-form solution of $F_1(\kappa) = A_1$.

For the Fourier transform of $f_2(x)$, we need to derive the Fourier transform of $\frac{1}{(d^2+x^2)^{3/2}}$, $\frac{x^2}{(d^2+x^2)^{3/2}}$, $\frac{1}{(d^2+x^2)^2}$ and $\frac{x^2}{(d^2+x^2)^2}$. From [38, Eq. (3.961)], we can get $\mathscr{F}\left[\frac{1}{(d^2+x^2)^{3/2}}\right] = \frac{1}{d}\sqrt{\frac{2}{\pi}}|\kappa|K_1(d|\kappa|)$ and $\mathscr{F}\left[\frac{x^2}{(d^2+x^2)^{3/2}}\right] = \sqrt{\frac{2}{\pi}}K_0(d|\kappa|) - d\sqrt{\frac{2}{\pi}}|\kappa|K_1(d|\kappa|)$. For $\frac{1}{(d^2+x^2)^2}$, we use residue theorem to get

$$\mathscr{F}\left[\frac{1}{(d^2+x^2)^2}\right] = 2\pi\mathrm{j}\operatorname*{Res}_{x=\mathrm{j}d}\left[\frac{e^{\mathrm{j}\kappa x}}{(d^2+x^2)^2}\right]$$
$$= \frac{1}{2d^3}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|}. \tag{64}$$

For $\frac{x^2}{(d^2+x^2)^2}$, we can get

$$\mathscr{F}\left[\frac{x^2}{(d^2+x^2)^2}\right] = \mathscr{F}\left[\frac{1}{d^2+x^2}\right] - d^2\mathscr{F}\left[\frac{1}{(d^2+x^2)^2}\right]$$
$$= \frac{\sqrt{\frac{\pi}{2}}e^{-d|\kappa|}}{d} - \frac{1}{2d}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|} \tag{65}$$

Finally, we can get

$$F_2(\kappa) = d\sqrt{\frac{\pi}{2}}e^{-d|\kappa|} + \frac{\mathrm{j}d\lambda}{2\pi}\sqrt{\frac{2}{\pi}}|\kappa|K_1(d|\kappa|)$$
$$- \frac{\mathrm{j}\lambda}{\pi}\left[\sqrt{\frac{2}{\pi}}K_0(d|\kappa|) - d\sqrt{\frac{2}{\pi}}|\kappa|K_1(d|\kappa|)\right]$$
$$- \left(\frac{\lambda}{2\pi}\right)^2\frac{1}{2d}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|}$$
$$+ \left(\frac{\lambda}{2\pi}\right)^2\left[\sqrt{\frac{\pi}{2}}\frac{2e^{-d|\kappa|}}{d} - \frac{1}{d}\sqrt{\frac{\pi}{2}}(1+|d\kappa|)e^{-|d\kappa|}\right]. \tag{66}$$

REFERENCES

[1] L. Sanguinetti, A. A. D'Amico, and M. Debbah, "Wavenumber-division multiplexing in line-of-sight holographic MIMO communications," *IEEE Trans. Wireless Commun.*, early access, Sep. 2022, doi: 10.1109/TWC.2022.3208961.
[2] F. K. Gruber and E. A. Marengo, "New aspects of electromagnetic information theory for wireless and antenna systems," *IEEE Trans. Antennas Propag.*, vol. 56, no. 11, pp. 3470–3484, Nov. 2008.
[3] O. Bucci, G. Franceschetti, and G. d'Elia, "Fast analysis of large antennas–a new computational philosophy," *IEEE Trans. Antennas Propag.*, vol. 28, no. 3, pp. 306–310, Dec. 1980.
[4] O. Bucci and G. Franceschetti, "On the spatial bandwidth of scattered fields," *IEEE Trans. Antennas Propag.*, vol. 35, no. 12, pp. 1445–1455, Dec. 1987.
[5] M. Franceschetti, "On Landau's eigenvalue theorem and information cut-sets," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 5042–5051, Jul. 2015.

[6] N. Decarli and D. Dardari, "Communication modes with large intelligent surfaces in the near field," *IEEE Access*, vol. 9, pp. 165 648–165 666, Dec. 2021.

[7] M. A. Jensen and J. W. Wallace, "Capacity of the continuous-space electromagnetic channel," *IEEE Trans. Antennas Propag.*, vol. 56, no. 2, pp. 524–531, Feb. 2008.

[8] W. Jeon and S.-Y. Chung, "Capacity of continuous-space electromagnetic channels with lossy transceivers," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1977–1991, Dec. 2017.

[9] M. D. Migliore, "On electromagnetics and information theory," *IEEE Trans. Antennas Propag.*, vol. 56, no. 10, pp. 3188–3200, Oct. 2008.

[10] ——, "Horse (electromagnetics) is more important than horseman (information) for wireless transmission," *IEEE Trans. Antennas Propag.*, vol. 67, no. 4, pp. 2046–2055, Apr. 2018.

[11] D. J. Griffiths, *Introduction to Electrodynamics*. American Association of Physics Teachers, 2005.

[12] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.

[13] A. S. Poon, R. W. Brodersen, and D. N. Tse, "Degrees of freedom in multiple-antenna channels: A signal space approach," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 523–536, Jan. 2005.

[14] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[15] I. M. Gelfand and A. IAglom, *Calculation of the amount of information about a random function contained in another such function*. American Mathematical Society Providence, 1959.

[16] M. Shafi, M. Zhang, A. L. Moustakas, P. J. Smith, A. F. Molisch, F. Tufvesson, and S. H. Simon, "Polarized MIMO channels in 3-d: models, measurements and mutual information," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 514–527, Mar. 2006.

[17] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550–560, Apr. 2002.

[18] R. Gatto and S. R. Jammalamadaka, "The generalized von Mises distribution," *Statistical Methodology*, vol. 4, no. 3, pp. 341–353, Jul. 2007.

[19] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Spatial characterization of electromagnetic random channels," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 847–866, Apr. 2022.

[20] L. V. Ahlfors, "Complex analysis: an introduction to the theory of analytic functions of one complex variable," *New York, London*, vol. 177, 1953.

[21] J. Mercer, "Functions of positive and negative type, and their connection with the theory of integral equations," *Philos. Trans. Roy. Soc. London*, vol. 209, no. 441-458, pp. 415–446, Jan. 1909.

[22] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.

[23] J. Zhu, Z. Zhang, Z. Wan, and L. Dai, "On finite-time mutual information," in *Proc. 2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 2910–2915.

[24] E. De Vito, V. Umanità, and S. Villa, "An extension of mercer theorem to matrix-valued measurable kernels," *Applied and Computational Harmonic Analysis*, vol. 34, no. 3, pp. 339–351, May 2013.

[25] Y. Polyanskiy and Y. Wu, "Lecture notes on information theory," *Lecture Notes for 6.441 (MIT), ECE563 (University of Illinois Urbana-Champaign), and STAT 664 (Yale)*, vol. 6, no. 2012-2016, p. 7, 2014.

[26] D. A. Miller, "Waves, modes, communications, and optics: a tutorial," *Advances in Optics and Photonics*, vol. 11, no. 3, pp. 679–825, 2019.

[27] M. Bachir, A. Fabre, and S. Tapia-Garcia, "Finitely determined functions," *Advances in Operator Theory*, vol. 6, no. 2, pp. 1–24, 2021.

[28] G. Franceschetti and D. Riccio, *Scattering, natural surfaces, and fractals*. Elsevier, 2006.

[29] B. Simon, *Trace ideals and their applications*. American Mathematical Soc., 2005, no. 120.

[30] C. Brislawn, "Kernels of trace class operators," *American Math. Soc.*, vol. 104, no. 4, 1988.

[31] B. Simon, "Notes on infinite determinants of Hilbert space operators," *Advances in Mathematics*, vol. 24, no. 3, pp. 244–273, 1977.

[32] J. A. Wheeler, "On the mathematical description of light nuclei by the method of resonating group structure," *Physical Review*, vol. 52, no. 11, p. 1107, 1937.

[33] F. Bornemann, "On the numerical evaluation of Fredholm determinants," *Mathematics of Computation*, vol. 79, no. 270, pp. 871–915, Sep. 2009.

[34] R. Gangolli, "Wide-sense stationary sequences of distributions on hilbert space and the factorization of operator valued functions," *Journal of Mathematics and Mechanics*, pp. 893–910, 1963.

[35] C. E. Shannon, "Communication in the presence of noise," *Proc. of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[36] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[37] I. M. Gelfand, R. A. Silverman *et al.*, *Calculus of variations*. Courier Corporation, 2000.

[38] D. Zwillinger, V. Moll, I. Gradshteyn, and I. Ryzhik, Eds., *Table of Integrals, Series, and Products (Eighth Edition)*. Boston: Academic Press, 2014.

Zhongzhichao Wan received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2020. He is currently pursuing the Ph.D. degree in Department of Electronic Engineering at Tsinghua University, Beijing, China. His research interests include electromagnetic information theory (EIT), coding theory, and channel modeling. He has received the Excellent Graduates of Tsinghua University in 2020.

Jieao Zhu received the B.E. degree in electronic engineering, and the second B.S. degree in applied mathematics from Tsinghua University in 2021. Currently, he is pursuing the Ph.D. degree in Department of Electronic Engineering at Tsinghua University, Beijing, China. His research interests include electromagnetic information theory (EIT), coding theory, and quantum computing. He has received the National Scholarship in 2018 and 2020, and the Excellent Graduates of Beijing in 2021.

Zijian Zhang (Student Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2020. He is currently working toward the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China. His research interests include physical-layer algorithms for massive MIMO and reconfigurable intelligent surfaces (RIS). He has received the National Scholarship in 2019 and the Excellent Thesis Award of Tsinghua University in 2020.

Linglong Dai (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree from the China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree from Tsinghua University, Beijing, in 2011. From 2011 to 2013, he was a Post-Doctoral Researcher with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016, an Associate Professor from 2016 to 2022, and has been a Professor since 2022. His current research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and Terahertz communications, wireless AI, and electromagnetic information theory.

He has coauthored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or coauthored over 70 IEEE journal papers and over 40 IEEE conference papers. He also holds 19 granted patents. He has received five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, and the IEEE ICC 2018. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ComSoc Stephen O. Rice Prize in 2022, and the IEEE ICC Best Demo Award in 2022. He was listed as a Highly Cited Researcher by Clarivate Analytics in 2020-2022. He was elevated as an IEEE Fellow in 2021. Particularly, he is dedicated to reproducible research and has made a large amount of simulation codes publicly available.

Chan-Byoung Chae (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin (UT) in 2008.

Prior to joining UT, he was a Research Engineer at the Telecommunications Research and Development Center, Samsung Electronics, Suwon, South Korea, from 2001 to 2005. He is currently an Underwood Distinguished Professor with the School of Integrated Technology, Yonsei University, South Korea. Before joining Yonsei University, he was with Bell Labs, Alcatel-Lucent, Murray Hill, NJ, USA, from 2009 to 2011, as a Member of Technical Staff, and Harvard University, Cambridge, MA, USA, from 2008 to 2009, as a Post-Doctoral Research Fellow.

Dr. Chae was a recipient/co-recipient of the CES Innovation Award in 2023, the IEEE ICC Best Demo Award in 2022, the IEEE WCNC Best Demo Award in 2020, the Best Young Engineer Award from the National Academy of Engineering of Korea (NAEK) in 2019, the IEEE DySPAN Best Demo Award in 2018, the IEEE/KICS Journal of Communications and Networks Best Paper Award in 2018, the IEEE INFOCOM Best Demo Award in 2015, the IEIE/IEEE Joint Award for Young IT Engineer of the Year in 2014, the KICS Haedong Young Scholar Award in 2013, the *IEEE Signal Processing Magazine* Best Paper Award in 2013, the IEEE ComSoc AP Outstanding Young Researcher Award in 2012, and the IEEE VTS Dan. E. Noble Fellowship Award in 2008. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON MOLECULAR, BIOLOGICAL, AND MULTI-SCALE COMMUNICATIONS and a Senior Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS. He has served/serves as an Editor for the *IEEE Communications Magazine* since 2016, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2012 to 2017, and the IEEE WIRELESS COMMUNICATIONS LETTERS since 2016. He is an IEEE ComSoc Distinguished Lecturer from 2020 to 2023. He is an IEEE Fellow and NAEK Fellow.