

Unified Characterization and Precoding for Non-Stationary Channels

Zhibin Zou, Maqsood Careem, Aveek Dutta

Department of Electrical and Computer Engineering
University at Albany SUNY, Albany, NY 12222 USA
{zzou2, mabdulcareem, adutta}@albany.edu

Abstract—Modern wireless channels are increasingly dense and mobile making the channel highly non-stationary. The time-varying distribution and the existence of joint interference across multiple degrees of freedom (e.g., users, antennas, frequency and symbols) in such channels render conventional precoding sub-optimal in practice, and have led to historically poor characterization of their statistics. The core of our work is the derivation of a high-order generalization of Mercer’s Theorem to decompose the non-stationary channel into constituent fading sub-channels (2-D eigenfunctions) that are jointly orthogonal across its degrees of freedom. Consequently, transmitting these eigenfunctions with optimally derived coefficients eventually mitigates any interference across these dimensions and forms the foundation of the proposed joint spatio-temporal precoding. The precoded symbols directly reconstruct the data symbols at the receiver upon demodulation, thereby significantly reducing its computational burden, by alleviating the need for any complementary decoding. These eigenfunctions are paramount to extracting the second-order channel statistics, and therefore completely characterize the underlying channel. Theory and simulations show that such precoding leads to $>10^4 \times$ BER improvement (at 20dB) over existing methods for non-stationary channels.

I. INTRODUCTION

Precoding at the transmitter is investigated in the literature and is relatively tractable when the wireless channel is stationary, by employing the gamut of linear algebraic and statistical tools to ensure interference-free communication [1], [2], [3]. However, there are many instances, in modern and next Generation propagation environments such as mmWave, V2X, and massive-MIMO networks, where the channel is statistically non-stationary [4], [5], [6] (the distribution is a function of time). This leads to sub-optimal and sometimes catastrophic performance even with state-of-the-art precoding [7] due to two factors: a) the time-dependence of the channel statistics, and b) the existence of interference both jointly and independently across multiple dimensions (space (users/antennas), frequency or time) in communication systems that leverage multiple degrees of freedom (e.g., MU-MIMO, OFDM, OTFS [8]). This necessitates a unified characterization of the statistics of wireless channels that can also incorporate time-varying statistics, and novel precoding algorithms warrant flat-fading over the higher-dimensional interference profiles in non-stationary channels. Our solution to the above addresses a challenging open problem in the literature [9] : “how to decompose non-stationary channels into independently fading sub-channels (along each degree of freedom) and how to precode using them”, which is central to both characterizing

Ngwe Thawdar
US Air Force Research Laboratory
Rome, NY, USA
ngwe.thawdar@us.af.mil

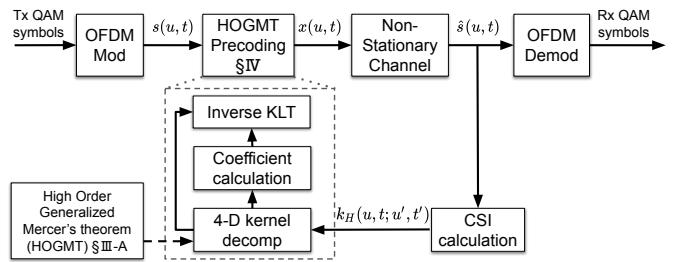


Figure 1: Unified characterization & precoding

channels and minimizing interference.

Unlike stationary channels, the second order statistics of non-stationary channels are 4-dimensional, as they are functions of both time-frequency and delay-Doppler dimensions¹ [10]. The core of our wireless channel characterization method, is the decomposition of this 4-dimensional channel kernel into 2-dimensional eigenfunctions that are jointly orthogonal across these dimensions, using a generalization of Mercer’s Theorem to high-dimensional and asymmetric kernels. Unlike recent literature that only partially characterize the non-stationary channel using a select few local statistics [11], [12], these eigenfunctions are used to extract any second-order statistics of the non-stationary channels that completely characterizes its distribution. Since any wireless channel model (e.g., deterministic, stationary, frequency flat or selective) can be extracted from the general non-stationary channel kernel, the extracted eigenfunctions lead to a unified method to characterize the statistics of any wireless channel.

Figure 1 shows the data flow for joint spatio-temporal precoding at the transmitter. The spatio-temporal CSI obtained from the receivers are used to construct a 4-dimensional channel kernel. In addition to spatial (inter-user or inter-antenna) or temporal (inter-symbol) interference, the time-varying kernel of non-stationary channels, induces joint space-time interference. We design a joint space-time precoding at the transmitter that involves combining the spatio-temporal eigenfunctions obtained by decomposing the space-time channel kernel, with optimal coefficients that minimize the least square error in the transmitted and received symbols. Since the eigenfunctions are independently and jointly orthonormal sub-channels over space and time, precoding using them warrants flat-fading

¹Second order statistics of stationary channels depend only on the delay-Doppler (2-D) and hence can be extracted as a degenerate case of the non-stationary channel model, when its time-frequency dependence is constant.

(interference-free communication) even in the presence of joint space-time interference. Further, these transmitted (precoded) symbols directly reconstruct the data symbols at the receiver when combined with calculated coefficients. Therefore, unlike existing precoding methods that require complementary decoding at the receiver[1], we alleviate any need for complex receiver processing thereby significantly reducing its computational burden. Finally, the precoded symbols are scheduled to each user and are processed through the conventional transmitter signal processing blocks (*e.g.*, CP/ guard insertion) before transmission. To the best of our knowledge, precoding for non-stationary channels is unprecedented in the literature. However, we include a comprehensive comparison with related precoding techniques in Appendix A-A in [13].

II. BACKGROUND

The wireless channel is typically expressed by a linear operator H , and the received signal $r(t)$ is given by $r(t)=Hs(t)$, where $s(t)$ is the transmitted signal. The physics of the impact of H on $s(t)$ is described using the delays and Doppler shifts in the multipath propagation [10] given by (1),

$$r(t) = \sum_{p=1}^P h_p s(t - \tau_p) e^{j2\pi\nu_p t} \quad (1)$$

where h_p , τ_p and ν_p are the path attenuation factor, time delay and Doppler shift for path p , respectively. (1) is expressed in terms of the overall delay τ and Doppler shift ν [10] in (2),

$$r(t) = \iint S_H(\tau, \nu) s(t - \tau) e^{j2\pi\nu t} d\tau d\nu \quad (2)$$

$$= \int L_H(t, f) S(f) e^{j2\pi t f} df \quad (3)$$

$$= \int h(t, \tau) s(t - \tau) d\tau \quad (4)$$

where $S_H(\tau, \nu)$ is the (*delay-Doppler*) *spreading function* of channel H , which describes the combined attenuation factor for all paths in the delay-Doppler domain. $S(f)$ is the Fourier transform of $s(t)$ and the time-frequency (TF) domain representation of H is characterized by its *TF transfer function*, $L_H(t, f)$, which is obtained by the 2-D Fourier transform of $S_H(\tau, \nu)$ as in (5). The time-varying impulse response $h(t, \tau)$ is obtained as the Inverse Fourier transform of $S_H(\tau, \nu)$ from the Doppler domain to the time domain as in (6).

$$L_H(t, f) = \iint S_H(\tau, \nu) e^{j2\pi(t\nu - f\tau)} d\tau d\nu \quad (5)$$

$$h(t, \tau) = \int S_H(\tau, \nu) e^{j2\pi t \nu} d\nu \quad (6)$$

Figures 2a and 2b show the time-varying response and TF transfer function for an example of a time-varying channel. For stationary channels, the TF transfer function is a stationary process with $\mathbb{E}\{L_H(t, f)L_H^*(t', f')\}=R_H(t-t', f-f')$, and the spreading function is a white process (uncorrelated scattering), i.e., $\mathbb{E}\{S_H(\tau, \nu)S_H^*(\tau', \nu')\}=C_H(\tau, \nu)\delta(\tau-\tau')\delta(\nu-\nu')$, where $\delta(\cdot)$ is the Dirac delta function. $C_H(\tau, \nu)$ and $R_H(t-t', f-f')$ are the *scattering function* and *TF correlation function*,

, respectively, which are related via 2-D Fourier transform,

$$C_H(\tau, \nu) = \iint R_H(\Delta t, \Delta f) e^{-j2\pi(\nu\Delta t - \tau\Delta f)} d\Delta t d\Delta f \quad (7)$$

In contrast, for non-stationary channels, the TF transfer function is non-stationary process and the spreading function is a non-white process. Therefore, a *local scattering function* (LSF) $C_H(t, f; \tau, \nu)$ [10] is defined to extend $C_H(\tau, \nu)$ to the non-stationary channels in (8). Similarly, the *channel correlation function* (CCF) $\mathcal{R}(\Delta t, \Delta f; \Delta\tau, \Delta\nu)$ generalizes $R_H(\Delta t, \Delta f)$ to the non-stationary case in (9).

$$\begin{aligned} & C_H(t, f; \tau, \nu) \\ &= \iint R_L(t, f; \Delta t, \Delta f) e^{-j2\pi(\nu\Delta t - \tau\Delta f)} d\Delta t d\Delta f \end{aligned} \quad (8)$$

$$\begin{aligned} &= \iint R_S(\tau, \nu; \Delta\tau, \Delta\nu) e^{-j2\pi(\Delta\tau\nu - f\Delta\tau)} d\Delta\tau d\Delta\nu \\ & \mathcal{R}(\Delta t, \Delta f; \Delta\tau, \Delta\nu) \\ &= \iint R_L(t, f; \Delta t, \Delta f) e^{-j2\pi(\Delta\nu t - \Delta f t)} dt df \\ &= \iint R_S(\tau, \nu; \Delta\tau, \Delta\nu) e^{-j2\pi(\Delta t\nu - \Delta f\tau)} d\tau d\nu \end{aligned} \quad (9)$$

where $R_L(t, f; \Delta t, \Delta f) = \mathbb{E}\{L_H(t, f + \Delta f)L_H^*(t - \Delta t, f)\}$ and $R_S(\tau, \nu; \Delta\tau, \Delta\nu) = \mathbb{E}\{S_H(\tau, \nu + \Delta\nu)S_H^*(\tau - \Delta\tau, \nu)\}$. For stationary channels, CCF reduces to TF correlation function $\mathcal{R}(\Delta t, \Delta f; \Delta\tau, \Delta\nu) = R_H(\Delta t, \Delta f)\delta(\Delta t)\delta(\Delta f)$. Due to space constraint, the complete proofs of Theorems and Lemmas are provided in an external document [13].

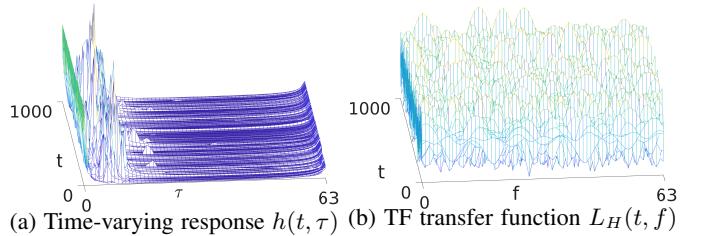


Figure 2: Illustration of a non-stationary channel

III. NON-STATIONARY CHANNEL DECOMPOSITION AND CHARACTERIZATION

The analysis of non-stationary channels is complicated as its statistics vary across both time-frequency and delay-Doppler domains resulting in 4-D second order statistics [14], which motivates the need for a unified characterization of wireless channels². Wireless channels are completely characterized by their statistics, however they are difficult to extract for non-stationary channels, due to their time dependence. Therefore, we start by expressing the channel H using an atomic channel G and the 4-D channel kernel $\mathcal{H}(t, f; \tau, \nu)$ [14] as in (10),

$$H = \iiint \mathcal{H}(t, f; \tau, \nu) G_{t, f}^{\tau, \nu} dt df d\tau d\nu \quad (10)$$

²Any channel can be generated as a special case of the non-stationary channel. Therefore a characterization of non-stationary channels would generalize to any other wireless channel [10].

where G is a normalized ($\|G\|=1$) linear prototype system whose transfer function $L_G(t, f)$ is smooth and localized about the origin of the TF plane. $G_{t,f}^{\tau,\nu} = S_{t,f+\nu} G S_{t-\tau,f}^+$ means that the atomic channel G shifts the signal components localized at $(t-\tau, f)$ to $(t, f+\nu)$ on the TF plane. $S_{\tau,\nu}$ is TF shift operator defined as $(S_{\tau,\nu}s)(t) = s(t-\tau)e^{j2\pi\nu t}$. Then the channel kernel $\mathcal{H}(t, f; \tau, \nu)$ is given by (11).

$$\begin{aligned} \mathcal{H}(t, f; \tau, \nu) &= \langle H, G_{t,f}^{\tau,\nu} \rangle \quad (11) \\ &= e^{j2\pi f \tau} \iint L_H(t', f') L_G^*(t'-t, f'-f) e^{-j2\pi(\nu t' - \tau f')} dt' df' \end{aligned}$$

The statistics of any wireless channel can always be obtained from the above 4-D channel kernel. Therefore, decomposing this kernel into fundamental basis allows us to derive a unified form to characterize any wireless channel.

A. Channel decomposition

4-D channel kernel decomposition into orthonormal 2-D kernels is unprecedented in the literature, but is essential to mitigate joint interference in the 2-D space and to completely characterize non-stationary channels. While Mercer's theorem [15] provides a method to decompose symmetric 2-D kernels into the same eigenfunctions along different dimensions, it cannot directly decompose 4-D channel kernels due to their high-dimensionality and since the kernel is not necessarily symmetric in the time-frequency delay-Doppler domains. Karhunen–Loeve transform (KLT) [16] provides a method to decompose kernels into component eigenfunctions of the same dimension, however is unable to decompose into orthonormal 2-D space-time eigenfunctions, and hence cannot be used to mitigate interference on the joint space-time dimensions. Therefore, we derive an asymmetric 4-dimensional kernel decomposition method that combines the following two steps as shown in figure 3: A) A generalization of Mercer's theorem that is applicable to both symmetric or asymmetric kernels, and B) An extension of KLT for high-dimensional kernels.

Lemma 1. (Generalized Mercer's theorem (GMT)) *The decomposition of a 2-D process $K \in L^2(X \times Y)$, where X and Y are square-integrable zero-mean processes, is given by,*

$$K(t, t') = \sum_{n=1}^{\infty} \sigma_n \psi_n(t) \phi_n(t') \quad (12)$$

where σ_n is a random variable with $\mathbb{E}\{\sigma_n \sigma_{n'}\} = \lambda_n \delta_{nn'}$, and λ_n is the n^{th} eigenvalue. $\psi_n(t)$ and $\phi_n(t')$ are eigenfunctions.

The proof combines Mercer's Theorem with KLT to generalize it to asymmetric kernels and is provided in Appendix B-A in the external document [13]. From Lemma 1, by letting $\rho(t, t') = \psi_n(t) \phi_n(t')$ in (12) we have (13),

$$K(t, t') = \sum_{n=1}^{\infty} \sigma_n \rho_n(t, t') \quad (13)$$

where the 2-D kernel is decomposed into random variable σ_n with constituent 2-D eigenfunctions, $\rho(t, t')$, this serves as an extension of KLT to 2-D kernels. A similar extension leads to

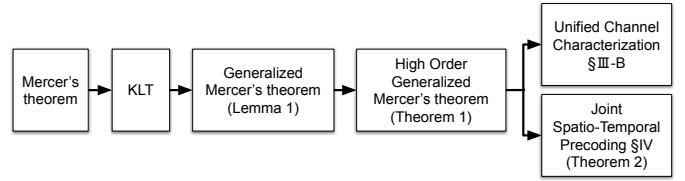


Figure 3: Derivation of High Order Generalized Mercer's Theorem for channel decomposition

the derivation of KLT for N-dimensional kernels which is key to deriving Theorem 1.

Theorem 1. (High Order GMT (HOGMT)) *The decomposition of $M=Q+P$ dimensional kernel $K \in L^M(X \times Y)$, where $X(\gamma_1, \dots, \gamma_Q)$ and $Y(\zeta_1, \dots, \zeta_P)$ are Q and P dimensional kernels respectively, that are square-integrable zero-mean random processes, is given by (14),*

$$K(\zeta_1, \dots, \zeta_P; \gamma_1, \dots, \gamma_Q) = \sum_{n=1}^{\infty} \sigma_n \psi_n(\zeta_1, \dots, \zeta_P) \phi_n(\gamma_1, \dots, \gamma_Q) \quad (14)$$

where $\mathbb{E}\{\sigma_n \sigma_{n'}\} = \lambda_n \delta_{nn'}$. λ_n is the n^{th} eigenvalue and $\psi_n(\zeta_1, \dots, \zeta_P)$ and $\phi_n(\gamma_1, \dots, \gamma_Q)$ are P and Q dimensional eigenfunctions respectively.

The proof is provided in Appendix B-B in [13]. Theorem 1 is applicable to any M dimensional channel kernel. Examples of such channel kernels may include 1-D time-varying channels, 2-D time-frequency kernels for doubly dispersive channels [17], user, antenna dimensions in MU-MIMO channels and angles of arrivals and departures in mmWave channels. Theorem 1 ensures that the 4-D channel kernel in (11) is decomposed as in (15) into 2-D eigenfunctions that are jointly orthonormal in the time-frequency dimensions as in (16).

$$\mathcal{H}(t, f; \tau, \nu) = \sum_{n=1}^{\infty} \sigma_n \psi_n(t, f) \phi_n(\tau, \nu) \quad (15)$$

$$\begin{aligned} \iint \psi_n(t, f) \psi_{n'}^*(t, f) dt df &= \delta_{nn'} \\ \iint \phi_n(\tau, \nu) \phi_{n'}^*(\tau, \nu) d\tau d\nu &= \delta_{nn'} \end{aligned} \quad (16)$$

Moreover, from (15) and (16) we have that,

$$\iint \mathcal{H}(t, f; \tau, \nu) \phi_n^*(\tau, \nu) d\tau d\nu = \sigma_n \psi_n(t, f) \quad (17)$$

(17) suggests that when the eigenfunction, $\phi_n^*(\tau, \nu)$ is transmitted through the channel a different eigenfunction, $\psi_n(t, f)$ is received with σ_n . Therefore, we refer to ϕ_n and ψ_n as a pair of *dual* eigenfunctions. By definition, these eigenfunctions are constituent sub-channels of the channel kernel that only undergo a scaling when transmitted over channel. Therefore, the dual eigenfunctions are referred to as *flat-fading sub-channels* of H .

B. Unified Characterization Using Channel Statistics

Wireless channels are fully characterized by their (second order) statistics, which we calculate using the extracted eigen-

values and 2-D eigenfunctions. The CCF is calculated as the correlations of $\mathcal{H}(t, f; \tau, \nu)$ [14] and is given by,

$$|\mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu)| \quad (18)$$

$$= \left| \iiint \mathbb{E}\{\mathcal{H}^*(t-\Delta t, f-\Delta f; \tau-\Delta \tau, \nu-\Delta \nu) \mathcal{H}(t, f; \tau, \nu)\} dt df d\tau d\nu \right| \\ = \sum_{n=1}^{\infty} \lambda_n |R_{\psi_n}(\Delta t, \Delta f)| |R_{\phi_n}(\Delta \tau, \Delta \nu)| \quad (19)$$

where (19) is obtained by substituting (17) in (18). $R_{\psi_n}(\Delta t, \Delta f)$ and $R_{\phi_n}(\Delta \tau, \Delta \nu)$ are the correlations of $\psi_n(t, f)$ and $\phi_n(\tau, \nu)$, respectively. The LSF reveals the non-stationarities (in time or frequency) in a wireless channel and is given by the 4-D Fourier transform (\mathbb{F}^4) of the CCF as,

$$\begin{aligned} \mathcal{C}_H(t, f; \tau, \nu) &= \mathbb{F}^4 \{ \mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu) \} \\ &= \iiint \mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu) e^{-j2\pi(t\Delta\nu - f\Delta\tau + \tau\Delta f - \nu\Delta t)} dt df d\tau d\nu \\ &= \sum_{n=1}^{\infty} \lambda_n |\psi_n(\tau, \nu)|^2 |\phi_n(t, f)|^2 \end{aligned} \quad (20)$$

where $|\psi_n(\tau, \nu)|^2$ and $|\phi_n(t, f)|^2$ represent the spectral density of $\psi_n(t, f)$ and $\phi_n(\tau, \nu)$, respectively. Then, the *global (or average) scattering function* $\bar{C}_H(\tau, \nu)$ and (local) TF path gain $\rho_H^2(t, f)$ [14] are calculated in (21) and (22),

$$\bar{C}_H(\tau, \nu) = \mathbb{E}\{|S_H(\tau, \nu)|^2\} = \iint \mathcal{C}_H(t, f; \tau, \nu) dt df \quad (21)$$

$$\rho_H^2(t, f) = \mathbb{E}\{|L_H(t, f)|^2\} = \iint \mathcal{C}_H(t, f; \tau, \nu) d\tau d\nu \quad (22)$$

(21) and (22) are re-expressed in terms of the spectral density of eigenfunctions by using (20) and the properties in (16),

$$\bar{C}_H(\tau, \nu) = \mathbb{E}\{|S_H(\tau, \nu)|^2\} = \sum_{n=1}^{\infty} \lambda_n |\psi_n(\tau, \nu)|^2 \quad (23)$$

$$\rho_H^2(t, f) = \mathbb{E}\{|L_H(t, f)|^2\} = \sum_{n=1}^{\infty} \lambda_n |\phi_n(t, f)|^2 \quad (24)$$

Finally, the *total transmission gain* \mathcal{E}_H^2 is obtained by integrating the LSF out with respect to all four variables,

$$\mathcal{E}_H^2 = \iiint \mathcal{C}_H(t, f; \tau, \nu) dt df d\tau d\nu = \sum_{n=1}^{\infty} \lambda_n \quad (25)$$

Therefore, the statistics of the non-stationary channel is completely characterized by its eigenvalues and eigenfunctions obtained by the decomposition of $\mathcal{H}(t, f; \tau, \nu)$, which are summarized in Table I.

Table I: Unified characterization of non-stationary channel

Statistics	Eigen Characterization
CCF $ \mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu) $	$\sum \lambda_n R_{\psi_n}(\Delta t, \Delta f) R_{\phi_n}(\Delta \tau, \Delta \nu) $
LSF $\mathcal{C}_H(t, f; \tau, \nu)$	$\sum \lambda_n \psi_n(\tau, \nu) ^2 \phi_n(t, f) ^2$
Global scattering function $\bar{C}_H(\tau, \nu)$	$\sum \lambda_n \psi_n(\tau, \nu) ^2$
Local TF path gain $\rho_H^2(t, f)$	$\sum \lambda_n \phi_n(t, f) ^2$
Total transmission gain \mathcal{E}_H^2	$\sum \lambda_n$

IV. JOINT SPATIO-TEMPORAL PRECODING

The kernel $\mathcal{H}(t, f; \tau, \nu)$ in (11) describes the time-frequency delay-Doppler response of the channel from (4) and is essential to extract the statistics of the non-stationary channel H for a single user as they depend on the same 4 dimensions.

For precoding, we express the spatio-temporal channel response by extending the time-varying response $h(t, \tau)$ to incorporate multiple users, i.e., $h_{u,u'}(t, \tau)$ [18], [19], which denotes the time-varying impulse response between the u^{th} transmit antenna and the u^{th} user, where each user has a single antenna. The 4-D spatio-temporal channel estimation is widely investigated [20], [21], [22]. In this paper we assume perfect CSI at transmit. Thus the received signal in (4) is extended as

$$\begin{aligned} r_u(t) &= \int \sum_{u'} h_{u,u'}(t, \tau) s_{u'}(t-\tau) d\tau + v_u(t) \\ &= \int \sum_{u'} k_{u,u'}(t, t') s_{u'}(t') dt' + v_u(t) \end{aligned} \quad (26)$$

where $v_u(t)$ is the noise, $s_u(t)$ is the data signal and $k_{u,u'}(t, t') = h_{u,u'}(t, t-\tau)$ is the channel kernel. Then, the relationship between the transmitted and received signals is obtained by rewriting (26) in its continuous form in (27).

$$r(u, t) = \iint k_H(u, t; u', t') s(u', t') du' dt' + v(u, t) \quad (27)$$

Let $x(u, t)$ be the precoded signal, then the corresponding received signal is $Hx(u, t)$. The aim of precoding in this work is to minimize the interference, i.e., to minimize the least square error, $\|s(u, t) - Hx(u, t)\|^2$.

Lemma 2. *Given a non-stationary channel H with kernel $k_H(u, t; u', t')$, if each projection in $\{H\varphi_n(u, t)\}$ are orthogonal to each other, there exists a precoded signal scheme $x(u, t)$ that ensures interference-free communication at the receiver,*

$$\|s(u, t) - Hx(u, t)\|^2 = 0 \quad (28)$$

where $\varphi_n(u, t)$ is the eigenfunction of $x(u, t)$, obtained by KLT decomposition as $x(u, t) = \sum_{n=1}^{\infty} x_n \varphi_n(u, t)$.

The proof is provided in Appendix C-A in [13]. Therefore, precoding using $\{\varphi_n\} = \{\phi_n\}$ or $\{\varphi_n\} = \{\psi_n\}$ obtained by decomposing the channel kernel using Theorem (1), (i.e., constructing $x(u, t)$ using $\{\phi_n\}$ or $\{\psi_n\}$ with coefficients x_n using inverse KLT, eventually leads to interference-free communication, as it satisfies (28) by ensuring that $\{H\varphi_n(u, t)\}$ are orthogonal using the properties in (16).

Theorem 2. *(HOGMT-based precoding) Given a non-stationary channel H with kernel $k_H(u, t; u', t')$, the precoded signal $x(u, t)$ that ensures interference-free communication at the receiver is constructed by inverse KLT as,*

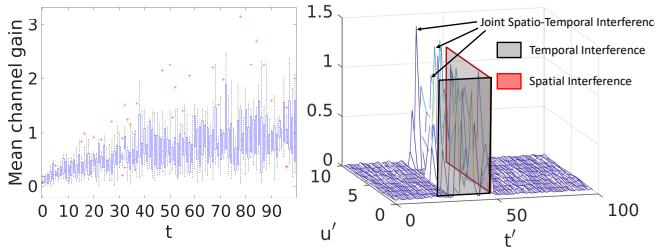
$$x(u, t) = \sum_{n=1}^{\infty} x_n \phi_n^*(u, t), \text{ where, } x_n = \frac{\langle s(u, t), \psi_n(u, t) \rangle}{\sigma_n} \quad (29)$$

where $\{\sigma_n\}$, $\{\psi_n\}$ and $\{\phi_n\}$ are obtained by decomposing the kernel $k_H(u, t; u', t')$ using Theorem 1 as in (30),

$$k_H(u, t; u', t') = \sum_{n=1}^{\infty} \sigma_n \psi_n(u, t) \phi_n(u', t') \quad (30)$$

with properties as in (31),

$$\begin{aligned} \iint \psi_n(u, t) \psi_{n'}^*(u, t) du dt &= \delta_{nn'} \\ \iint \phi_n(u', t') \phi_{n'}^*(u', t') du' dt' &= \delta_{nn'} \end{aligned} \quad (31)$$



(a) Distribution of mean channel gains at each time instance. (b) Channel kernel $k_H(u, t; u', t')$ for $u=1$ and $t=50$

Figure 4: Non-stationary channel statistics and kernel

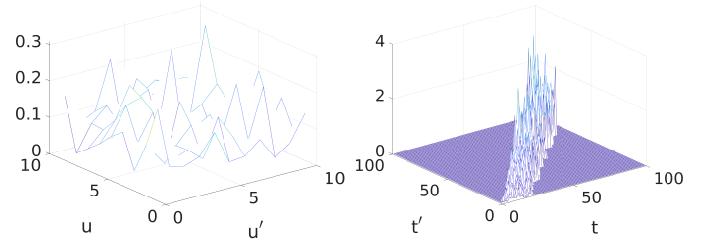
The proof is provided in Appendix C-B in [13]. Although precoding involves a linear combination of $\phi_n^*(u, t)$ with x_n it is a non-linear function ($\mathcal{W}(\cdot)$) with respect to the data signal $s(u, t)$, i.e., $x(u, t) = \mathcal{W}(\{\phi_n(u, t)\}; \{\psi_n(u, t)\}, \{\sigma_n\}, s(u, t))$. (30) and (31) suggest that the 4-D kernel is decomposed into jointly orthogonal sub-channels $\{\psi_n(u, t)\}$ and $\{\phi_n(u', t')\}$. Therefore, the precoding in Theorem 2 can be explained as transmitting the eigenfunctions $\{\phi_n^*(u, t)\}$ after multiplying with specific coefficients $\{x_n\}$. Consequently, when transmitted through the channel H , it transforms $\{\phi_n(u, t)\}$ to its dual eigenfunctions $\{\psi_n(u, t)\}$ with $\{\sigma_n\}$ as in (32),

$$\iint k_H(u, t; u', t') \phi_n^*(u, t) du dt = \sigma_n \psi_n(u', t'). \quad (32)$$

which means that, $H\phi_n^*(u, t) = \sigma_n \psi_n(u, t)$, as proved in Appendix C-B in [13]. Then the data signal $s(u, t)$ is directly reconstructed at the receiver (to the extent of noise $v_u(t)$) as the net effect of precoding and propagation in the channel ensures that from (26), $r(u, t) = Hx(u, t) + v_u(t) \rightarrow s(u, t) + v_u(t) = \hat{s}(u, t)$ using Lemma 2, where $\hat{s}(u, t)$ is the estimated signal. Therefore, the spatio-temporal decomposition of the channel in Theorem 1 allows us to precode the signal such that all interference in the spacial domain, time domain and joint space-time domain are cancelled when transmitted through the channel, leading to a joint spatio-temporal precoding scheme. Further, this precoding ensures that the data signal is reconstructed directly at the receiver with an estimation error that of $v_u(t)$, thereby completely pre-compensating the spatio-temporal fading/ interference in non-stationary channels to the level of AWGN noise. Therefore, this precoding does not require complementary decoding at receiver, which vastly reducing its hardware and computational complexity compared to state-of-the-art precoding methods like Dirty Paper Coding (DPC) or linear precoding (that require a complementary decoder [23]).

Theorem 2 does not make any assumptions on the type, dimensions or size of the channel kernel. Corollary 1 demonstrates the application of Theorem 2 to an example of a deterministic multi-user channel where only spatial interference from other users' exist. The received signal is given by $r_u = \sum_{u'} h_{u, u'} s_{u'} + v_n$ and its continuous form is given by,

$$r(u) = \int k_H(u, u') s(u') du' + v(u) \quad (33)$$



(a) 4-D kernel $k_H(u, t; u', t')$ at $t=1$ and $t'=1$ which shows the $u=1$ and $u'=1$ which shows the spatial interference (b) 4-D kernel $k_H(u, t; u', t')$ at $u=1$ and $u'=1$ which shows the temporal interference

Figure 5: Separately spatial and delay effects of 4-D kernel

Corollary 1. *Given a deterministic multi-user channel kernel $k(u, u')$, the precoded signal $x(u)$ that warrants spatial interference-free reception is given by (34).*

$$x(u) = \sum_{n=1}^{\infty} \frac{\langle s(u), \psi_n(u) \rangle}{\sigma_n} \phi_n^*(u) \quad (34)$$

where $\{\sigma_n\}$, $\{\psi_n\}$ and $\{\phi_n\}$ follow from Theorem 1 for the 2-D case, i.e., $k_H(u; u') = \sum_{n=1}^{\infty} \sigma_n \psi_n(u) \phi_n(u')$.

Proof of Corollary 1 is provided in Appendix C-C of [13]. The precoding in Corollary 1 holds for any other 2-D channel kernels like the stationary/non-stationary single-user channel kernel given by $k_H(t, t')$ by replacing $k_H(u, u')$ with $k_H(t, t')$.

V. RESULTS

We analyze the accuracy of the proposed unified channel characterization and joint spatio-temporal precoding using a non-stationary channel simulation framework in Matlab. The simulation environment considers 10 mobile receivers (users) and 100 time instances of a non-stationary 4-D kernel $k_H(u, t; u', t')$, where the number of delayed symbols (delay taps) causing interference are uniformly distributed between [10, 20] symbols for each user at each time instance.

The pre-processing of the 4-D channel kernel and 2-D data symbols involves mapping them to a low-dimensional space using an invertible mapping $f: u \times t \rightarrow m$. Although Theorem 1 decomposes the channel kernel into infinite eigenfunctions, we show that it is sufficient to decompose the channel kernel into a finite number of eigenfunctions and select only those whose eigenvalues are greater than a threshold value ϵ^2 for precoding, i.e., $\{(\phi_n(\cdot), \psi_n(\cdot)) : \sigma_n > \epsilon\}$. These eigenfunctions are used to calculate the coefficients for joint spatio-temporal precoding, which subsequently construct the precoded signal after inverse KLT and combining the real and imaginary parts.

Figure 4a shows distribution of the statistics of the channel gain (mean and variance) for each time instance of the non-stationary channel and further corroborates its non-stationarity.

Figure 4b shows the channel response for user $u=1$ at $t=50$ which indicates the manifestation of interference. For $u=1$, the spatial interference from other users (inter-user interference) occurs on the red plane (at $t'=50$), while the temporal interference at $t=50$ occurs due to previous delayed symbols (inter-

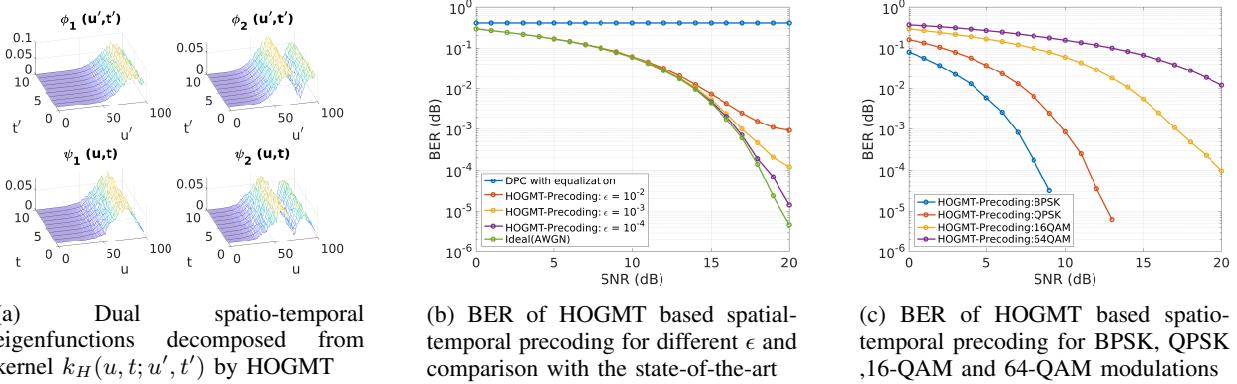


Figure 6: HOGMT based spatio-temporal precoding

symbol interference) on the grey plane (at $u'=1$). However, the received symbols at user $u=1$ is also affected by other delayed symbols from other users (i.e., $t' < 50$ for $u' \neq 1$), which leads to joint spatio-temporal interference, and necessitates joint precoding over space-time dimensions. The variation of the channel kernels over time and for different users is demonstrated and further explained in Appendix D in [13] for completion. This is the cause of joint space-time interference which necessitates joint precoding in the 2-dimensional space using eigenfunctions that are jointly orthogonal.

Figure 5a shows the spatial (inter-user) interference caused by other users for a fixed time instance (time instances $t=1$ and $t'=1$) in terms of the 4-D channel kernel, i.e., $k_H(u, 1, u', 1)$. Figure 5b shows the temporal (inter-symbol) interference for a user $u=1$ caused by its own (i.e., $u'=1$) delayed symbols (e.g., due to multipath), i.e., $k_H(1, t; 1, t')$. We also observe that the temporal interference for each user occurs from its own 20 immediately delayed symbols.

Figure 6a shows two pairs of dual spatio-temporal eigenfunctions ($\phi_n(u', t')$, $\psi_n(u, t)$) (absolute values) obtained by decomposing $k_H(u, t; u', t')$ in (30). We see that this decomposition is indeed asymmetric as each $\phi_n(u', t')$ and $\psi_n(u', t')$ are not equivalent (but shifted), and that each $\phi_1(u', t')$ and $\psi_1(u, t)$ are jointly orthogonal with $\phi_2(u', t')$ and $\psi_2(u, t)$ as in (16), respectively. Therefore, when $\phi_1(u', t')$ (or $\phi_2(u', t')$) is transmitted through the channel, the dual eigenfunctions, $\psi_1(u, t)$ (or $\psi_2(u, t)$) is received with σ_1 and σ_2 , respectively. Therefore, the non-stationary 4-D channel is decomposed to dual flat-fading sub-channels.

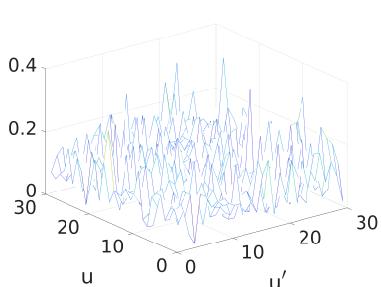
Figures 6b shows the BER at the receiver, using joint spatio-temporal precoding (HOGMT-precoding) at the transmitter with 16-QAM modulated symbols for non-stationary channels. Since this precoding is able to cancel all interference that occurs in space, time and across space-time dimensions which are shown in figure 4b, it achieves significantly lower BER over existing precoding methods that employ DPC at the transmitter to cancel spatial interference and Zero Forcing equalization at the receiver to mitigate temporal interference. Further, we show that with sufficient eigenfunctions ($\epsilon = 10^{-4}$ in this case), proposed method can achieve near ideal BER, only 0.5dB more SNR to achieve the same BER as the

ideal case, where the ideal case assumes all interference is cancelled and only AWGN noise remains at the receiver. This gap exists since practical implementation employs a finite number of eigenfunctions as opposed to an infinite number in (30). Figure 6c compares the BER of HOGMT based spatio-temporal precoding for various modulations (BPSK, QPSK, 16-QAM and 64-QAM) for the same non-stationary channel with $\epsilon = 10^{-3}$. As expected we observe that the lower the order of the modulation, the lower the BER but at the cost of lower data rate. However, we observe that even with high-order modulations (e.g., 64-QAM) the proposed precoding achieves low BER ($\approx 10^{-2}$ at SNR=20dB), allowing high data-rates even over challenging non-stationary channels. The choice of the order of the modulation is therefore, based on the desired BER and data rate for different non-stationary scenarios.

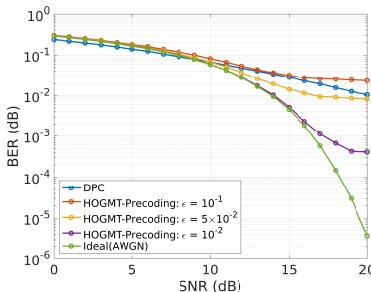
Figure 7 shows an example of precoding for an example of a multi-user deterministic channel defined by the 2-D kernel, $K_H(u, u')$ with 30 mobile users as defined in (33), where only spatial interference from other users exists. The spatial interference from other users is portrayed in figure (7a) in terms of $K_H(u, u')$. The achieved BER at the receiver by precoding using Corollary 1 is shown in figure 7b and is compared with the state-of-the-art DPC for spatial precoding [3]. The performance of HOGMT based spatial precoding for various modulations with $\epsilon = 10^{-3}$ is compared in figure 7c. We observe that with low number of eigenfunctions ($\epsilon = 10^{-1}$), the proposed precoding results in higher BER than DPC because precoding using these limited eigenfunctions is not sufficient to cancel all the spatial interference. However, with more eigenfunctions ($\epsilon = 10^{-2}$), it achieves significantly lower BER compared to DPC for SNR > 9 dB and consequently only requires 3dB more SNR to achieve the same BER as the ideal case. In contrast, while DPC is optimal in the sum rate at the transmitter is not optimal in terms of the BER at the receiver and its BER performance depends on the complementary decoding performed at the receiver.

VI. CONCLUSION

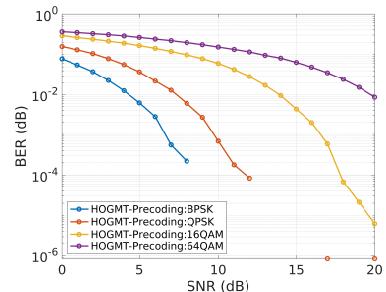
In this work, we derived a high-order generalized version of Mercer's Theorem to decompose the general non-stationary channel kernel into 2-dimensional jointly orthogonal



(a) 2-D kernel $k_H(u, u')$ where only spatial interference exists



(b) BER of HOGMT based spatial precoding for different ϵ and comparison with the state-of-the-art



(c) BER of HOGMT based spatial precoding for BPSK, QPSK, 16-QAM and 64-QAM modulations

Figure 7: HOGMT based spatial precoding

flat fading sub-channels (eigenfunctions). Through theoretical analysis and simulations, we draw three firm conclusions for non-stationary channels: 1) The 2-dimensional eigenfunctions are sufficient to completely derive the second-order statistics of the non-stationary channel and consequently leads to an unified characterization of any wireless channel, 2) precoding by combining these eigenfunctions with optimally derived coefficients mitigates the spatio-temporal interference, and 3) the precoded symbols when propagated over the non-stationary channel directly reconstruct the data symbols at the receiver when combined with calculated coefficients, consequently alleviating the need for complex complementary decoding at the receiver. Therefore, the encouraging results from this work will form the core of robust and unified characterization and highly reliable communication over nonstationary channels, supporting emerging application.

VII. ACKNOWLEDGEMENT

This work is funded by the Air Force Research Laboratory Visiting Faculty Research Program (SA10032021050367), Rome, New York, USA.

REFERENCES

- [1] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010.
- [2] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, “Massive mimo linear precoding: A survey,” *IEEE systems journal*, vol. 12, no. 4, pp. 3920–3931, 2017.
- [3] M. Costa, “Writing on dirty paper (corresp.),” *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [4] C.-X. Wang, J. Bian, J. Sun, W. Zhang, and M. Zhang, “A survey of 5g channel measurements and models,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3142–3168, 2018.
- [5] Z. Huang and X. Cheng, “A general 3d space-time-frequency non-stationary model for 6g channels,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 535–548, 2020.
- [6] C. F. Mecklenbrauker, A. F. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernadó, T. Zemen, O. Klemp, and N. Czink, “Vehicular channel characterization and its implications for wireless system design and performance,” *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1189–1212, 2011.
- [7] A. Ali, E. D. Carvalho, and R. W. Heath, “Linear receivers in non-stationary massive mimo channels with visibility regions,” *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 885–888, 2019.
- [8] R. Hadani, S. Rakib, S. Koms, M. Tsatsanis, A. Monk, C. Ibarras, J. Delfeld, Y. Hebron, A. J. Goldsmith, A. F. Molisch, and A. R. Calderbank, “Orthogonal time frequency space modulation,” *CoRR*, vol. abs/1808.00519, 2018. [Online]. Available: <http://arxiv.org/abs/1808.00519>
- [9] G. Matz and F. Hlawatsch, “Time-varying communication channels: Fundamentals, recent developments, and open problems,” in *2006 14th European Signal Processing Conference*, 2006, pp. 1–5.
- [10] F. Hlawatsch and G. Matz, *Wireless Communications Over Rapidly Time-Varying Channels*, 1st ed. USA: Academic Press, Inc., 2011.
- [11] M. Pätzold and C. A. Gutierrez, “Modelling of non-wssus channels with time-variant doppler and delay characteristics,” in *2018 IEEE Seventh International Conference on Communications and Electronics (ICCE)*, 2018, pp. 1–6.
- [12] J. Bian, C.-X. Wang, X. Gao, X. You, and M. Zhang, “A general 3d non-stationary wireless channel model for 5g and beyond,” 2021.
- [13] Z. Zou, M. Careem, A. Dutta, and N. Thawdar, “Proofs and supplementary material: Unified characterization and precoding for non-stationary channels,” 2022. [Online]. Available: <https://arxiv.org/abs/2202.01827>
- [14] G. Matz, “On non-wssus wireless fading channels,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2465–2478, 2005.
- [15] J. Mercer, “Functions of positive and negative type, and their connection with the theory of integral equations,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 209, pp. 415–446, 1909. [Online]. Available: <http://www.jstor.org/stable/91043>
- [16] L. Wang, *Karhunen-Loeve expansions and their applications*. London School of Economics and Political Science (United Kingdom), 2008.
- [17] G. Matz, “Doubly underspread non-wssus channels: analysis and estimation of channel statistics,” in *2003 4th IEEE Workshop on Signal Processing Advances in Wireless Communications - SPAWC 2003 (IEEE Cat. No.03EX689)*, 2003, pp. 190–194.
- [18] A. Ghazal, Y. Yuan, C.-X. Wang, Y. Zhang, Q. Yao, H. Zhou, and W. Duan, “A non-stationary imt-advanced mimo channel model for high-mobility wireless communication systems,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2057–2068, 2017.
- [19] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-espsti, H. Hofstetter, P. Kyosti, D. Laurenson, G. Matz, A. F. Molisch, C. Oestges, and H. Ozcelik, *Survey of channel and radio propagation models for wireless MIMO systems*. EURASIP Journal on Wireless Communications and Net-working, 2007.
- [20] S. Srivastava, M. S. Kumar, A. Mishra, S. Chopra, A. K. Jagannatham, and L. Hanzo, “Sparse doubly-selective channel estimation techniques for ostbc mimo-ofdm systems: A hierarchical bayesian kalman filter based approach,” *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4844–4858, 2020.
- [21] X. Ma, L. Yang, and G. Giannakis, “Optimal training for mimo frequency-selective fading channels,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 453–466, 2005.
- [22] M. Milojevic, G. Del Galdo, and M. Haardt, “Tensor-based framework for the prediction of frequency-selective time-variant mimo channels,” in *2008 International ITG Workshop on Smart Antennas*, 2008, pp. 147–152.
- [23] M. H. Vu, “Exploiting transmit channel side information in mimo wireless systems,” Ph.D. dissertation, Stanford University, 2006.