Joint Spatio-Temporal Precoding for Practical Non-Stationary Wireless 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

Ngwe Thawdar

US Air Force Research Laboratory

Rome, NY, USA

ngwe.thawdar@us.af.mil

Abstract

The high mobility, density and multi-path evident in modern wireless systems makes the channel highly non-stationary. This causes temporal variation in the channel distribution that leads to the existence of time-varying joint interference across multiple degrees of freedom (DoF, *e.g.*, users, antennas, frequency and symbols), which renders conventional precoding sub-optimal in practice. In this work, we derive a High-Order Generalization of Mercer's Theorem (HOGMT), which decomposes the multi-user non-stationary channel into two (dual) sets of jointly orthogonal subchannels (eigenfunctions), that result in the other set when one set is transmitted through the channel. This duality and joint orthogonality of eigenfuntions ensure transmission over independently flat-fading subchannels. Consequently, transmitting these eigenfunctions with optimally derived coefficients eventually mitigates any interference across its degrees of freedoms and forms the foundation of the proposed joint spatio-temporal precoding. The transferred dual eigenfuntions and coefficients directly reconstruct the data symbols at the receiver upon demodulation, thereby significantly reducing its computational burden, by alleviating the need for any complementary post-coding. Additionally, the eigenfunctions decomposed from the time-frequency delay-Doppler channel kernel are paramount to extracting the second-order channel statistics, and therefore completely characterize the underlying channel. We evaluate this using a realistic non-stationary channel framework built in Matlab and show that our precoding achieves $\geqslant 4$ orders of reduction in BER at SNR$\geqslant$15dB in OFDM systems for higher-order modulations and less complexity compared to the state-of-the-art precoding.

I. INTRODUCTION

Precoding has been widely investigated for stationary channels, where the orthogonality along each DoF is enforced by decomposing them using linear algebraic tools (*e.g.*, singular value decomposition (SVD) or QR decomposition [1]) leading to capacity achieving strategies [2], [3]

typically under the block fading assumption. The statistical non-stationarity that is evident in modern and next Generation propagation environments including V2X, mmWave, and massive-MIMO channels, [4], [5], [6], leads to *catastrophic* error rates even with state-of-the-art pre-coding [7] as such capacity-achieving strategies optimized for stationary channels do not ensure interference-free communication when the channel distribution changes over time. The non-stationarity in such channels engender joint interference across multiple dimensions (space (users/ antennas), time-frequency or delay-Doppler) in communication systems that leverage multiple degrees of freedom (*e.g.*, MU-MIMO, OFDM, OTFS [8]). This time-varying joint interference renders conventional decomposition techniques incapable of achieving flat-fading. 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 channels and minimizing interference.

The multi-user non-stationary channel is represented as a 4-dimensional space (user) time-varying impulse response and is acquired from the CSI obtained from each receiver (user). The core of our precoding is the decomposition of this asymmetric channel, by generalizing Mercer's Theorem [10] to high-dimensional asymmetric processes, into 2-dimensional eigenfunctions that are jointly orthogonal across the DoF. These eigenfunctions serve as independently flat-fading subchannels, and precoding using them is the key to canceling the time-varying joint interference that exists across the DoF. The second order statistics of non-stationary channels vary across time-frequency and delay-Doppler (4-dimensions) and therefore, such channels can be represented as atomic channels with 4-dimensional asymmetric coefficients [11]. Unlike recent literature that only partially characterize the non-stationary channel using a select few local statistics [12], [13], the 2-dimensional eigenfunctions decomposed from 4-dimensional asymmetric coefficients 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 system view of joint spatio-temporal precoding. The spatio-temporal CSI is used to extract a 4-dimensional kernel, which can be decomposed into dual 2-dimensional
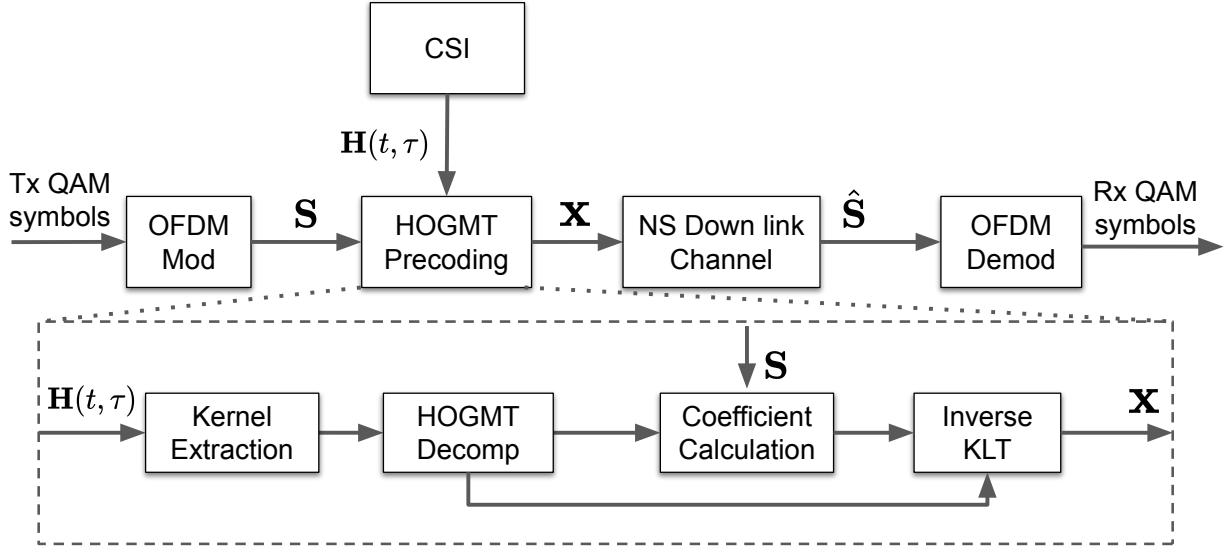
Figure 1: System view of HOGMT-based precoding for non-stationary channel

(space-time) eigenfunctions by derived High Order Generalized Mercer's Theorem (HOGMT). Thus decorrelate the space-time domain at the transmitter and the receiver. The spatio-temporal eigenfunctions corresponding to the receiver are used to derive optimal coefficients that minimize the least square error in the transmitted and received symbols. Then combining their dual spatio-temporal eigenfunctions with these coefficients by inverse KLT. Since the eigenfunctions are independently and jointly orthonomal sub-channels over space and time, precoding using them warrants flat-fading (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 our best of knowledge, precoding for NS channel is first proposed in our previous work [14], which is extended in this work. Additionally, we investigate the computational complexity of our precoding and show that it achieves lower complexity than Dirty Paper Coding (Section VII-A). Therefore, we make the following contributions:

1) Decomposing non-stationary channels into flat-fading subchannels: The 4-dimensional

Applications	Source of NS	Degree of NS	SoTA BER
Vehicle-to-Everything (V2X) **High-Speed Train (HST)** **Unmanned Aerial Vehicle (UAV)**	High Tx-Rx mobility, weather, time-varying scatterers, AoA, AoD and Doppler, UAV attitude.	V2X: SI=10–40 m at 5–15 km/h [15], [16] HST: SI=1.37-2.84m at 2.6GHz and 18MHz BW [17] UAV: SI=15m at C band with 50 MHz BW [18], [19]	V2X: BER>0.01 for BPSK at SNR 20dB [20]
MIMO **Massive MIMO** **Extra-large MIMO (XL-MIMO)** **MU-MIMO**	Time varying multipath, delay spread and spatial visibility regions.	MIMO: 0-3m Correlation Matrix Distance [21] Massive/XL/MU-MIMO: AoA shift from 100°-80° at 2.6 GHz with 50 MHz BW [22]	MIMO: BER>0.07 for QPSK at SNR 20dB [21] with 8×8 MIMO Massive/XL/MU-MIMO: BER>10^{-2} with 320 antenna elements [23]
mmWave, **THz** **Underwater**, **Satellite-Com** **Visible Light Communications** **Reconf Intelligent Surface** **Hypersonic Reentry Com (HRC)**	Time varying blockage, reflection angles and NLoS, atmospheric absorption topology and opportunistic access.	mmWave: Delay spread<20 ns at 28–73 GHz [4], [24] HRC: Coherence time $\approx 5\mu s$ [25]	THz: BER>0.02 for 16-QAM at 26m distance and frequency of 350 GHz [26] Underwater: BER>0.003 QPSK at SNR 20dB [27] HRC: BER>10^{-1} [25]

Table I: Overview of applications in non-stationary channels: Impairments, SoTA and BER performance.

channel kernels is decomposed into dual jointly orthogonal subchannels (eigenfunctions) that are flat-fading in the eigen-domain (Section IV).

2) A Unified Characterization of Wireless Channels: The eigenfunctions decomposed from 4-dimensional coefficients of atomic channels completely characterizes the non-stationary channel, which can generalize to any wireless channel (Section V).

3) Joint Spatio-Temporal Precoding: The eigenfunctions decomposed from 4-dimensional channel kernels are leveraged to cancel the joint spatio-temporal interference in non-stationary channels (Section VI).

4) Post-coding free Precoding: The precoded symbols transmitting through the channel directly reconstruct the modulated sybols at receivers without complementary step, alleviating additional computational burden (Section VI).

II. BACKGROUND & RELATED WORK

We categorize the related work into three categories:

Non-stationarity of Wireless Channels: Table I provides ample evidence in the literature for the existence of non-stationarity (NS) in modern wireless channels. Non-stationarity is primarily attributed to temporal and sometimes spatial variation in the transceivers and the dynamic nature of the scattering environment, and is measured in terms of the stationarity interval (SI) [15] in time or space. Depending on the features of the channel non-stationarity may arise from the time-varying Doppler in V2X, HST and UAV channels, from the time-varying multipath in MIMO channels and its variants and due to time varying blockage in mmWave, THz and VLC channels. Unfortunately, state of the art (SoTA) methods applied to NS channels are only able to achieve a modest error rate that is inadequate to support high data rate wireless applications like mobile AR/VR/XR, aerial communications and 4K/8K HDR video streaming services. These will require joint spatio-temporal precoding for NS channels that is capable of achieving orders of magnitude improvement in Bit Error Rate (BER) across all types of channels and applications.

Precoding in Non-Stationary Channels: Although precoding non-stationary channels is unprecedented in the literature [7], we list the most related literature for completeness. The challenge in precoding non-stationary channels is the time-dependence of statistics and the channel cannot be modeled as the time-independent matrix. This leads to suboptimal performance using state-of-the-art precoding techniques like Dirty Paper Coding (DPC). Though DPC is theoretically interference-free with perfect CSI, the current implementation by QR decomposition targeting for separate channel matrices, unable to capture the variation over time. Meanwhile, recent literature present attempt to deal with imperfect CSI by modeling the error in the CSI [28], [29], [30], [31], [32], [33], [34], [35], they are limited by the assumption the channel or error statistics are stationary or WSSUS at best.

Spatio-Temporal Precoding: While, precoding has garnered significant research, spatio-temporal interference is typically treated as two separate problems, where spatial precoding at the transmitter aims to cancel inter-user and inter-antenna interference, while equalization at the receiver mitigates inter-carrier and inter-symbol interference. Alternately, [8] proposes to modulate the symbols such that it reduces the cross-symbol interference in the delay-Doppler domain, but requires equalization at the receiver to completely cancel such interference in practical systems. Moreover, this approach cannot completely minimize the joint spatio-temporal interference that occurs in non-stationary channels since their statistics depend on the time-frequency domain in addition to the delay-Dopper domain (explained in Section III-A). While spatio-temporal block

coding techniques are studied in the literature [1] they add redundancy and hence incur a communication overhead to mitigate interference, which we avoid by precoding. These techniques are capable of independently canceling the interference in each domain, however are incapable of mitigating interference that occurs in the joint spatio-temporal domain in non-stationary channels. We design a joint spatio-temporal precoding that leverages the extracted 2-D eigenfunctions from non-stationary channels to mitigate interference that occurs on the joint space-time dimensions, which to the best of our knowledge is unprecedented in the literature.

III. MODELS & PRELIMINARIES

A. Non-stationary wireless channel model

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 shift in the multipath propagation [11] given by (1),

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

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 [11] in (2),

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

$$= \int L_H(t,f)S(f)e^{j2\pi tf} \ df = \int h(t,\tau)s(t-\tau) \ d\tau \tag{3}$$

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 (4). 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 (5).

$$L_H(t,f)= \iint S_H(\tau,\nu)e^{j2\pi(t\nu-f\tau)} \ d\tau \ d\nu \tag{4}$$

$$h(t,\tau)= \int S_H(\tau,\nu)e^{j2\pi t\nu} \ d\nu \tag{5}$$

Figure 2a and 2b show the time-varying impulse response and TF transfer function for a NS channel, respectively.

(a) The time-varying impulse response of antenna 1 to user 1

(b) The time-frequency transfer function of antenna 1 to user 1

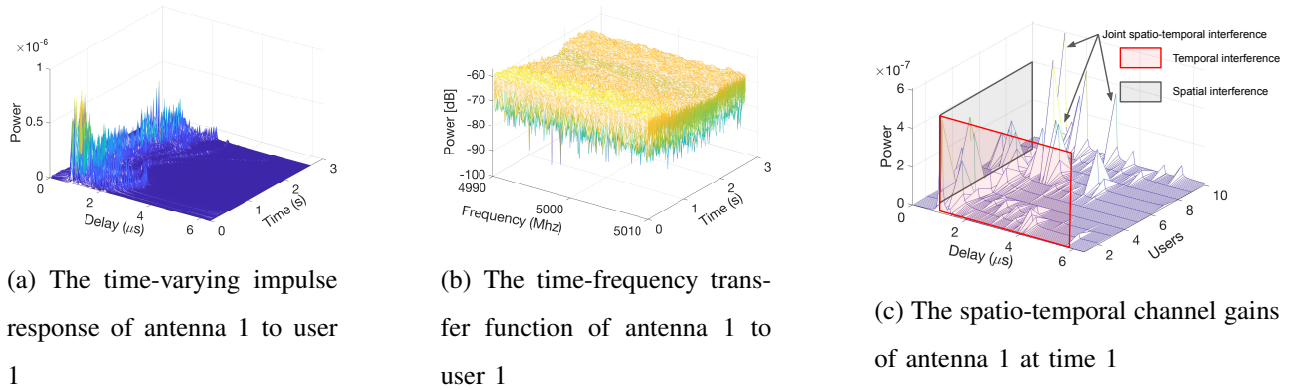(c) The spatio-temporal channel gains of antenna 1 at time 1

Figure 2: Illustration of a non-stationary channel

B. Statistics of non-stationary channels

For stationary channels, the TF transfer function is a stationary process and and the spreading function is a white process (uncorrelated scattering) with

$$\mathbb{E}\{L_H(t,f)L_H^*(t',f')\}=R_H(t-t',f-f') \tag{6}$$

$$\mathbb{E}\{S_H(\tau,\nu)S_H^*(\tau',\nu')\}=C_H(\tau,\nu)\delta(\tau-\tau')\delta(\nu-\nu') \tag{7}$$

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,

$$C_H(\tau,\nu) = \iint R_H(\Delta t, \Delta f)e^{-j2\pi(\nu\Delta t-\tau\Delta f)}\,d\Delta t\,d\Delta f \tag{8}$$

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) $\mathcal{C}_H(t,f;\tau,\nu)$ [11] is defined to extend $C_H(\tau,\nu)$ to the non-stationary channels in (9). 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 (10).

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

$$\mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu)$$
$$= \iint R_L(t, f; \Delta t, \Delta f) e^{-j2\pi(\Delta \nu t - \Delta \tau f)} \ dt \ df \tag{10}$$
$$= \iint R_S(\tau, \nu; \Delta \tau, \Delta \nu) e^{-j2\pi(\Delta t \nu - \Delta f \tau)} \ d\tau \ d\nu$$

where

$$R_L(t, f; \Delta t, \Delta f) = \mathbb{E}\{L_H(t, f + \Delta f) L_H^*(t - \Delta t, f)\}$$

$$R_S(\tau, \nu; \Delta \tau, \Delta \nu) = \mathbb{E}\{S_H(\tau, \nu + \Delta \nu) S_H^*(\tau - \Delta \tau, \nu)\}$$

For stationary channels, CCF reduces to TF correlation function as

$$\mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu) = R_H(\Delta t, \Delta f) \delta(\Delta t) \delta(\Delta f)$$

C. Multi-user non-stationary channel model

For precoding, we express the spatio-temporal downlink channel response by extending the time-varying response $h(t, \tau)$ to incorporate multiple users. Without loss of generality, for convenience of exposition let us consider the case where each user has a single antenna, i.e., MISO case. Denotes $h_{u,u'}(t, \tau)$ [36] as the time-varying impulse response between the u'^{th} transmit antenna and the uth user (For MIMO case, $h_{u,u'}(t, \tau)$ is a matrix). Thus the received signal in (3) is extended to the multi-user case of $h(t, \tau)$ and is given by (11),

$$\mathbf{H}(t, \tau) = \begin{bmatrix} h_{1,1}(t, \tau) & \cdots & h_{1,u'}(t, \tau) \\ \vdots & \ddots & \\ h_{u,1}(t, \tau) & & h_{u,u'}(t, \tau) \end{bmatrix} \tag{11}$$

Therefore, the received signal in (3) is extended as in (11). The first term in (12) is desired signal with fading effects. Spatial and temporal interference correspond to the second and third terms in (12) and the black and red regions in figure 2c, respectively. Additionally, the space-time kernel induces joint spatio-temporal due to the interference from delayed symbols from other users as shown in the figure and the last term of (12). Further, since the spatio-temporal signals in practice

$$r_u(t) = \int \sum_{u'} h_{u,u'}(t, \tau) s_{u'}(t - \tau) d\tau + v_u(t) \tag{11}$$

$$= \underbrace{h_{u,u}(t, 0) s_u(t)}_{\text{Signal with the attenuation coefficient}} + \underbrace{\sum_{u' \neq u} h_{u,u'}(t, 0) s_{u'}(t)}_{\text{Spatial interference}} + \underbrace{\int h_{u,u}(t, \tau) s_u(t - \tau) d\tau}_{\text{Temporal interference}} + \underbrace{\int \sum_{u' \neq u} h_{u,u'}(t, \tau) s_{u'}(t - \tau) d\tau}_{\text{Joint spatio-temporal interference}} + \underbrace{v_u(t)}_{\text{Noise}}$$

$$\tag{12}$$

are 2-dimensional ($r_u(t)$ or $s_u(t)$), canceling all the above interference, necessitates a method to decompose the asymmetric 4-D channel in (11) into 2-D independently fading subchannels.

The CSI processing including estimating the 4-D spatio-temporal channel at the receiver (to obtain CSI), compressing and transmitting, and predicting CSI by outdated CSI are widely investigated in the literature [37], [38], [39], [40], [41] and hence is not the focus of this work. Therefore, in the following sections we investigate the ability to cancel all interference for non-stationary channels under perfect CSI.

Therefore, in the following sections we investigate the ability to cancel all interference (pre-code) non-stationary channels under perfect CSI.

IV. NON-STATIONARY CHANNEL DECOMPOSITION

4-D channel decomposition into orthonormal 2-D subchannels is unprecedented the literature, but is essential to mitigate joint interference in the 2-D space and to completely characterize non-stationary channels (Any channel can be generated as a special case of the non-stationary channel. Therefore a precoding for non-stationary channels would generalize to any other wireless channel [11]). While SVD is only capable of decomposing LTI channels, Karhunen–Loève transform (KLT) [42] provides a method to decompose random process into component eigenfunctions of the same dimension. However, KLT is unable to decompose the multi-user time-varying 4-D channel in (11), into orthonormal 2-D space-time eigenfunctions (DoF decorrelation), and therefore cannot mitigate interference on the joint space-time dimensions. Mercer's theorem provides a method to decompose symmetric 2-D kernels into the same eigenfunctions along different dimensions, however, it cannot directly decompose 4-D channel kernels due to their high-dimensionality and since the channel kernel in (11) is not necessarily symmetric in the 4 dimensions. Therefore, we derive a generalized version of Mercer's Theorem for asymmetric kernels and extend it to higher-order kernels, which decomposes the asymmetric 4-D channel into 2-D jointly orthogonal subchannels. Consequently, this leads to flat-fading communication and mitigates joint spatio-temporal interference. The above decomposition techniques are compared in Table II, where only SVD has been used in the literature for precoding.

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

$$K(t, t') = \sum\nolimits_{n=1}^{\infty} \sigma_n \psi_n(t) \phi_n(t') \tag{13}$$

Table II: Comparison of HOGMT with other channel decomposition methods

Decomposition methods	Time-varying	Asymetric	DoF decorrelation	High-order
SVD: $H = U\Sigma V^*$		✓	✓	
Karhunen–Loève Transform (KLT): $X(t) = \Sigma_n \sigma_n \phi_n(t)$	✓	✓		
Mercer's Theorem: $K(t,t') = \Sigma_n \lambda_n \phi_n(t)\phi_n(t')$	✓		✓	
Generalized Mercer's Theorem (GMT): $K(t,t') = \Sigma_n \sigma_n \psi_n(t)\phi_n(t')$	✓	✓	✓	
HOGMT: $K(\zeta_1,...,\zeta_P;\gamma_1,...,\gamma_Q)=\sum_n \sigma_n \psi_n(\zeta_1,...,\zeta_P)\phi_n(\gamma_1,...,\gamma_Q)$	✓	✓	✓	✓

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

Proof. Consider a 2-dimensional process $K(t,t') \in L^2(Y \times X)$, where $Y(t)$ and $X(t')$ are square-integrable zero-mean random processes with covariance function K_Y and K_X, respectively. The projection of $K(t,t')$ onto $X(t')$ is obtained as in (14),

$$C(t) = \int K(t,t')X(t') \ dt' \tag{14}$$

Using *Karhunen–Loève Transform* (KLT), $X(t')$ and $C(t)$ are both decomposed as in (15),

$$X(t') = \sum_{i=1}^{\infty} x_i \phi_i(t') \quad \text{and,} \quad C(t) = \sum_{j=1}^{\infty} c_j \psi_j(t) \tag{15}$$

where x_i and c_j are both random variables with $\mathbb{E}\{x_i x_{i'}\}=\lambda_{x_i}\delta_{ii'}$ and $\mathbb{E}\{c_j c_{j'}\}=\lambda_{c_j}\delta_{jj'}$. $\{\lambda_{x_i}\}$, $\{\lambda_{x_j}\}$ $\{\phi_i(t')\}$ and $\{\psi_j(t)\}$ are eigenvalues and eigenfuncions, respectively. Let us denote $n=i=j$ and $\sigma_n=\frac{c_n}{x_n}$, and assume that $K(t,t')$ can be expressed as in (16),

$$K(t,t') = \sum_{n}^{\infty} \sigma_n \psi_n(t)\phi_n(t') \tag{16}$$

We show that (16) is a correct representation of $K(t,t')$ by proving (14) holds under this definition. We observe that by substituting (15) and (16) into the right hand side of (14) we have that,

$$\int K(t,t')X(t') \ dt' = \int \sum_{n}^{\infty} \sigma_n \psi_n(t)\phi_n(t') \sum_{n}^{\infty} x_n \phi_n(t') \ dt'$$

$$= \int \sum_{n}^{\infty} \sigma_n x_n \psi_n(t)|\phi_n(t')|^2 + \sum_{n'\neq n}^{\infty} \sigma_n x_{n'} \psi_n(t)\phi_n(t')\phi_{n'}^*(t')dt'$$

$$= \sum_{n}^{\infty} c_n \psi_n(t) = C(t) \tag{17}$$

which is equal to the left hand side of (14). Therefore, (16) is a correct representation of $K(t, t')$. □

From Lemma 1, by letting $\rho(t, t') = \psi_n(t)\phi_n(t')$ in (13) we have (18),

$$K(t, t') = \sum_{n=1}^{\infty} \sigma_n \rho_n(t, t') \tag{18}$$

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 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 $X(\gamma_1, \cdots, \gamma_Q)$ and $Y(\zeta_1, \cdots, \zeta_P)$ are Q and P dimensional kernels respectively, that are square-integrable zero-mean random processes, is given by (19),*

$$K(\zeta_1,...,\zeta_P; \gamma_1,...,\gamma_Q) = \sum_{n=1}^{\infty} \sigma_n \psi_n(\zeta_1,...,\zeta_P) \phi_n(\gamma_1,...,\gamma_Q) \tag{19}$$

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

Proof. Given a 2-D process $X(\gamma_1, \gamma_2)$, the eigen-decomposition using Lemma 1 is given by,

$$X(\gamma_1, \gamma_2) = \sum_{n}^{\infty} x_n e_n(\gamma_1) s_n(\gamma_2) \tag{20}$$

Letting $\psi_n(\gamma_1, \gamma_2) = e_n(\gamma_1)s_n(\gamma_2)$, and substituting it in (20) we have that,

$$X(\gamma_1, \gamma_2) = \sum_{n}^{\infty} x_n \phi_n(\gamma_1, \gamma_2) \tag{21}$$

where $\phi_n(\gamma_1, \gamma_2)$ are 2-D eigenfunctions with the property (22).

$$\iint \phi_n(\gamma_1, \gamma_2) \phi_{n'}(\gamma_1, \gamma_2) \, d\gamma_1 \, d\gamma_2 = \delta_{nn'} \tag{22}$$

We observe that (21) is the 2-D form of KLT. With iterations of the above steps, we obtain *High-Order KLT* for $X(\gamma_1, \cdots, \gamma_Q)$ and $C(\zeta_1, \cdots, \zeta_P)$ as given by,

$$X(\gamma_1, \cdots, \gamma_Q) = \sum_{n}^{\infty} x_n \phi_n(\gamma_1, \cdots, \gamma_Q) \tag{23}$$

$$C(\zeta_1, \cdots, \zeta_P) = \sum_{n}^{\infty} c_n \psi_n(\zeta_1, \cdots, \zeta_P) \tag{24}$$

where $C(\zeta_1, \cdots, \zeta_P)$ is the projection of $X(\gamma_1, \cdots, \gamma_Q)$ onto $K(\zeta_1, \cdots, \zeta_P; \gamma_1, \cdots, \gamma_Q)$. Then following similar steps as in the proof of Lemma 2, we get (19). □

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 [43], user, antenna dimensions in MU-MIMO channels and angles of arrivals and departures in mmWave channels.

V. CHARACTERIZATION OF NON-STATIONARY CHANNELS

Wireless channel characterization in the literature typically require several local and global (in space-time dimensions) higher order statistics to characterize or model non-stationary channels, due to their time-varying statistics. These statistics cannot completely characterize the non-stationary channel, however are useful in reporting certain properties that are required for the application of interest such as channel modeling, assessing the degree of stationarity etc. Contrarily, we leverage the 2-dimensional eigenfunctions that are decomposed from the most generic representation of any wireless channel as a spatio-temporal channel kernel. These spatio-temporal eigenfunctions can be used to extract any higher order statistics of the channel as demonstrated in Section IV, and hence serves as a complete characterization of the channel. Furthermore, since this characterization can also generalize to stationary channels, it is a unified characterization for any wireless channel. Beyond characterizing the channel, these eigenfunctions are the core of the precoding algorithm.

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 [44], which motivates the need for a unified characterization of wireless channels[1]. 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 $\mathcal{H}(t, f; \tau, \nu)$ [44] as in (25),

$$H = \iiiint \mathcal{H}(t, f; \tau, \nu) G_{t,f}^{\tau,\nu} \, dt \, df \, d\tau \, dv \tag{25}$$

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

---

[1]Any 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 [11].

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 (26).

$$\mathcal{H}(t, f, \tau, \nu) = \left\langle H, G_{t,f}^{\tau,\nu} \right\rangle \tag{26}$$

$$=e^{j2\pi f\tau} \iint L_H\left(t', f'\right) L_G^*\left(t'\text{-}t, f'\text{-}f\right) e^{-j2\pi(\nu t'-\tau f')}\mathrm{d}t'\mathrm{d}f'$$

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. Theorem 1 ensures that the 4-D channel kernel in (26) is decomposed as in (27) into 2-D eigenfunctions that are jointly orthonormal in the time-frequency or delay-Doppler dimensions as in (28).

$$\mathcal{H}(t, f; \tau, \nu) = \sum_{n=1}^{\infty} \sigma_n \psi_n(t, f)\phi_n(\tau, \nu) \tag{27}$$

$$\iint \psi_n(t, f)\psi_{n'}^*(t, f) \ dt \ df=\delta_{nn'}$$
$$\iint \phi_n(\tau, \nu)\phi_{n'}^*(\tau, \nu) \ d\tau \ d\nu=\delta_{nn'} \tag{28}$$

The variation across time-frequency delay-Doppler domains in the 4-D channel kernel is extracted by decomposing into separate 2-D eigenfuntions in time-frequency and delay-Doppler domains, respectively. The decorrelation of dimensionality and the orthonormal properties in (28) allow eigenfuntions and eigenvalues to extract statistics in either time-frequency and delay Doppler 4-D domains or separate 2-D domains, as shown in Corollary 1.

Corollary 1. *(Unified characterization for non-stationary channel by HOGMT) 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 III.*

Table III: 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 $\overline{C}_H(\tau,\nu)$	$\sum\lambda_n\|\psi_n(\tau,\nu)\|^2$
Local TF path gain $\rho_H^2(t, f)$	$\sum\lambda_n\|\phi_n(t, f)\|^2$
Total transmission gain $\mathcal{E}_H^2$	$\sum\lambda_n$

Proof. Wireless channels are fully characterized by their (second order) statistics, which we calculate using the extracted eigenvalues and 2-D eigenfunctions. The CCF is calculated as the correlations of $\mathcal{H}(t, f; \tau, \nu)$ and is given by,

$$|\mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu)| \tag{29}$$

$$= \left| \iiiint \mathbb{E}\{\mathcal{H}^*(t\text{-}\Delta t, f\text{-}\Delta f; \tau\text{-}\Delta \tau, \nu\text{-}\Delta \nu)\mathcal{H}(t, f; \tau, \nu)\} dt df d\tau d\nu \right|$$

$$= \sum_{n=1}^{\infty} \lambda_n |R_{\psi_n}(\Delta t, \Delta f)||R_{\phi_n}(\Delta \tau, \Delta \nu)| \tag{30}$$

where (30) is obtained by substituting (27) in (29). $R_{\psi_n}(\Delta t, \Delta f)$ 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,

$$\mathcal{C}_H(t, f; \tau, \nu) = \mathbb{F}^4\{\mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu)\}$$

$$= \iiiint \mathcal{R}(\Delta t, \Delta f; \Delta \tau, \Delta \nu) e^{-j2\pi(t\Delta \nu - f\Delta \tau + \tau \Delta f - \nu \Delta t)} dt df d\tau d\nu$$

$$= \sum_{n=1}^{\infty} \lambda_n |\psi_n(\tau, \nu)|^2 |\phi_n(t, f)|^2 \tag{31}$$

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* $\overline{C}_H(\tau, \nu)$ and (local) TF path gain $\rho_H^2(t, f)$ [44] are calculated in (32) and (33),

$$\overline{C}_H(\tau, \nu) = \mathbb{E}\{|S_H(\tau, \nu)|^2\} = \iint \mathcal{C}_H(t, f; \tau, \nu) \, dt \, df \tag{32}$$

$$\rho_H^2(t, f) = \mathbb{E}\{|L_H(t, f)|^2\} = \iint \mathcal{C}_H(t, f; \tau, \nu) \, d\tau \, dv \tag{33}$$

(32) and (33) are re-expressed in terms of the spectral density of eigenfunctions by using (31) and the properties in (28),

$$\overline{C}_H(\tau, \nu) = \mathbb{E}\{|S_H(\tau, \nu)|^2\} = \sum_{n=1}^{\infty} \lambda_n |\psi_n(\tau, \nu)|^2 \tag{34}$$

$$\rho_H^2(t, f) = \mathbb{E}\{|L_H(t, f)|^2\} = \sum_{n=1}^{\infty} \lambda_n |\phi_n(t, f)|^2 \tag{35}$$

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

$$\mathcal{E}_H^2 = \iiiint \mathcal{C}_H(t, f; \tau, \nu) \, dt \, df \, d\tau \, dv = \sum_{n=1}^{\infty} \lambda_n \tag{36}$$

$\square$

Consequently, the statistics of any wireless channel can be expressed by the eigenfunctions and eigenvalues obtained by the decomposition in (27). Therefore, we refer to Corollary 1 as a *unified characterization* of wireless channels. Further, Corollary 1 also suggests that the non-stationary channels are completely explained/characterized by the components decomposed by (27), thereby serving as a validation of the correctness of HOGMT.

VI. JOINT SPATIO-TEMPORAL PRECODING

A. Dual space-time variation of non-stationary channels

The received signal in (11) can be expressed by the channel kernel as in (37) [11]

$$r_u(t) = \int \sum_{u'} k_{u,u'}(t,t') s_{u'}(t') dt' + v_u(t) \tag{37}$$

where $v_u(t)$ is the noise, $s_u(t)$ is the data signal and $k_{u,u'}(t,t')=h_{u,u'}(t,t-t')$ is the channel kernel. Then, the relationship between the transmitted and received signals is obtained by rewriting (37) in its continuous form in (38).

$$r(u,t) = \iint k_H(u,t;u',t') s(u',t') \ du' \ dt' + v(u,t) \tag{38}$$

It is clear that this joint space-time interference (from both u' and t') varies along the space and time dimensions (i.e., across both u and t). This is referred to as the *dual space-time variation property* and it indicates that precoding using arbitrary joint space-time orthogonal basis is not sufficient to ensure interference-free communication, unless these basis remain orthogonal after propagating through the channel (as shown in Lemma 2).

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 all existing interference of the channel, 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,*

$$\|s(u,t) - Hx(u,t)\|^2 = 0 \tag{39}$$

where $\varphi_n(u,t)$ is the 2-D eigenfunction of $x(u,t)$, obtained by KLT decomposition as in (40)

$$x(u,t) = \sum_{n=1}^{\infty} x_n \varphi_n(u,t) \tag{40}$$

where x_n is a random variable with $E\{x_n x_{n'}\}=\lambda_n \delta_{nn'}$.

Proof. $H\varphi_n(u,t)$ is the projection of $k_H(u,t;u',t')$ onto $\varphi_n(u',t')$ denoted by $c_n(u,t)$ and is given by,

$$c_n(u,t) = \iint k_H(u,t;u',t')\varphi_n(u',t') \ du' \ dt' \tag{41}$$

Using the above, (39) is expressed as,

$$||s(u,t) - Hx(u,t)||^2 = ||s(u,t) - \sum_n^\infty x_n c_n(u,t)||^2 \tag{42}$$

Let $\epsilon(x) = ||s(u,t) - \sum_n^\infty x_n\varphi_n(u,t)||^2$. Then its expansion is given by,

$$\epsilon(x) = \langle s(u,t), s(u,t)\rangle - 2\sum_n^\infty x_n\langle c_n(u,t), s(u,t)\rangle \tag{43}$$

$$+ \sum_n^\infty x_n^2\langle c_n(u,t), c_n(u,t)\rangle + \sum_n^\infty \sum_{n'\neq n}^\infty x_n x_{n'}\langle c_n(u,t), c_{n'}(u,t)\rangle$$

Then the solution to achieve minimal $\epsilon(x)$ is obtained by solving for $\frac{\partial\epsilon(x)}{\partial x_n} = 0$ as in (44).

$$x_n^{opt} = \frac{\langle s(u,t), c_n(u,t)\rangle - \frac{1}{2}\sum_{n'\neq n}^\infty x_{n'}\langle c_{n'}(u,t), c_n(u,t)\rangle}{\langle c_n(u,t), c_n(u,t)\rangle} \tag{44}$$

where $\langle a(u,t), b(u,t)\rangle = \iint a(u,t)b^*(u,t) \ du \ dt$ denotes the inner product. Let $\langle c_{n'}(u,t), c_n(u,t)\rangle = 0$, i.e., the projections $\{c_n(u,t)\}_n$ are orthogonal basis. Then we have a closed form expression for x^{opt} as in (45).

$$x_n^{opt} = \frac{\langle s(u,t), c_n(u,t)\rangle}{\langle c_n(u,t), c_n(u,t)\rangle} \tag{45}$$

Substitute (45) in (43), it is straightforward to show that $\epsilon(x) = 0$. □

B. Dual jointly orthogonal space-time decomposition

Lemma 2 formalizes the requirements for the joint space-time orthogonal basis $\{\varphi_n(u,t)\}$ to achieve a precoding scheme that ensures interference-free reception. From Theorem 1, the 4-D channel kernel is decomposed as,

$$k_H(u,t;u',t') = \sum_{n=1}^\infty \sigma_n\psi_n(u,t)\phi_n(u',t') \tag{46}$$

with properties as in (47),

$$\iint \psi_n(u,t)\psi_{n'}^*(u,t) \ du \ dt = \delta_{nn'}$$
$$\iint \phi_n(u',t')\phi_{n'}^*(u',t') \ du' \ dt' = \delta_{nn'} \tag{47}$$

(46) and (47) suggest that the 4-D kernel is decomposed into jointly orthogonal subchannels,

$\{\psi_n(u,t)\}$ and $\{\phi_n(u',t')\}$. Moreover, combining (46) and (47) leads to (48), which shows the *duality* of the subchannels.

$$\iint k_H(u,t;u',t')\phi_n^*(u',t') \ du' \ dt' = \sigma_n\psi_n(u,t). \tag{48}$$

This duality suggests that when $\{\phi_n\}$ is transmitted through the 4-D channel, it transforms it to $\{\psi_n\}$ with random variables $\{\sigma_n\}$. Therefore, we refer to ϕ_n and ψ_n as a pair of *dual* eigenfunctions. Meanwhile, the transformation of ϕ_n to ψ_n is scaled by σ_n meaning that decomposed dual joint space-time orthogonal subchannels are flat-fading.

C. HOGMT-based precoding

Lemma 2 suggest precoding using $\{\varphi_n\}=\{\phi_n\}$ i.e., constructing $x(u,t)$ using $\{\phi_n\}$ with optimally derived coefficients x_n using inverse KLT, eventually leads to interference-free communication, as the projections of channel kernels onto $\{\phi_n\}$ is $\{\sigma_n\psi_n\}$, which satisfies the orthogonal projection requirements in lemma 2 due to the orthogonal properties in (47).

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,*

$$x(u,t)=\sum_{n=1}^{\infty} x_n\phi_n^*(u,t), where, x_n=\frac{\langle s(u,t),\psi_n(u,t)\rangle}{\sigma_n} \tag{49}$$

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 (46).

Proof. The 4-D kernel $k_H(u,t;u',t')$ is decomposed into two separate sets of eigenfunction $\{\phi_n(u',t')\}$ and $\{\psi_n(u,t)\}$ using Theorem 1 as in (46). By transmitting the conjugate of the eigenfunctions, $\phi_n(u,t)$ through the channel H, we have that,

$$H\phi_n^*(u,t)=\iint k_H(u,t;u',t')\phi_n^*(u',t') \ du' \ dt' = \sigma_n\psi_n(u,t) \tag{50}$$

where $\psi_n(u,t)$ is also a 2-D eigenfunction with the orthogonal property as in (47). From Lemma 2, if the projection $c_n(u,t)$ in (41) is $c_n(u,t) = \sigma_n\psi(u,t)$, which satisfies the orthogonality $\langle \sigma_{n'}\psi_{n'}(u,t), \sigma_n\psi_n(u,t)\rangle = 0$, we achieve the optimal solution as in (45). Therefore, let $x(u,t)$ be the linear combination of $\{\phi_n^*(u,t)\}$ with coefficients $\{x_n\}$ as in (51),

$$x(u,t) = \sum_{n}^{\infty} x_n\phi_n^*(u,t) \tag{51}$$

Then (42) is rewritten as in (52),

$$||s(u,t) - Hx(u,t)||^2 = ||s(u,t) - \sum_n^{\infty} x_n \sigma_n \psi(u,t)||^2 \tag{52}$$

Therefore, optimal x_n in (45) is obtained as in (53),

$$x_n^{opt} = \frac{\sigma_n \langle s(u,t), \psi_n(u,t) \rangle}{\lambda_n \langle \psi_n(u,t), \psi_n(u,t) \rangle} = \frac{\langle s(u,t), \psi_n(u,t) \rangle}{\sigma_n} \tag{53}$$

Substituting (53) in (51), the transmit signal is given by (54),

$$x(u,t) = \sum_n^{\infty} \frac{\langle s(u,t), \psi_n(u,t) \rangle}{\sigma_n} \phi_n^*(u,t). \tag{54}$$

$\square$

As $\{\psi_n(u,t)\}$ are joint space-time orthogonal basis, data signal $s(u,t)$ can be expressed by

$$s(u,t) = \sum_{n=1}^{\infty} s_n \psi_n(u,t), \text{where}, s_n = \langle s(u,t), \psi_n(u,t) \rangle \tag{55}$$

Therefore, the precoding in Theorem 2 can be explained as transmitting the eigenfunctions $\{\phi_n^*(u,t)\}$ after multiplying with derived coefficients $\{x_n\}$, which will transfer to $\{s_n\}^2$. Then the data signal $s(u,t)$ is directly reconstructed at the receiver by the dual eigenfunctions $\{\psi_n(u,t)\}$ with transferred coefficients $\{s_n\}$ to the extent of noise $v_u(t)$ as the net effect of precoding and propagation in the channel ensures that from (37), $r(u,t) = Hx(u,t) + v_u(t) \rightarrow s(u,t) + v_u(t) = \hat{s}(u,t)$ 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 modulated symbol is reconstructed directly at the receiver with an estimation error that of $v_u(t)$ (equation (13) in section III), 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 step 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 [45].
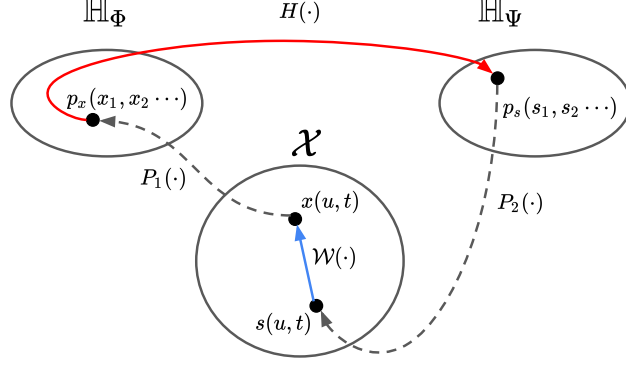
Figure 3: Geometric interpretation of HOGMT-precoding

D. Geometric interpretation of precoding

Figure 3 shows a geometric interpretation of HOGMT precoding without considering the noise item. Given two Hilbert Space $\mathbb{H}_\Phi$ and $\mathbb{H}_\Psi$, where basis are eigenfunctions $\{\phi_n\}$ and $\{\psi_n\}$, respectively, the precoded signal $x(u,t)=\mathcal{W}(k_H; s(u,t)) \in \mathcal{X}$ can be seen as a point $p_x(x_1, x_2, \cdots) = P_1(x(u,t)) \in \mathbb{H}_\Phi$, where x_n obtained in (49) is n^{th} coordinate. Then the 4-D channel have the transform $H(\cdot) : p_x(x_1, x_2, \cdots) \in \mathbb{H}_\Phi \rightarrow p_s(s_1, s_2, \cdots) \in \mathbb{H}_\Psi$, where the point $p_s(s_1, s_2, \cdots)$ represented in reality (project to space-time space $\mathcal{X}$) is directly the data signal $s(u,t) = P_2(p_s(s_1, s_2, \cdots)) \in \mathcal{X}$. The dual spatio-temporal variation of the 4-D channel not only transfer the coordinate $(x_1, x_2, \cdots)$ to $(s_1, s_2, \cdots)$, but also transfer Hilbert Space $\mathbb{H}_\Phi$ to $\mathbb{H}_\Psi$. As HOGMT extract the duality (explain $H(\cdot)$) and dual orthogonality (explain $P_1(\cdot)$ and $P_2(\cdot)$) in this dual variation, we can use inverse method to construct the precoded signal, i.e., for the target closed loop $\mathcal{W}(\cdot) + P_1(\cdot) + H(\cdot) + P_2(\cdot) = 0$, we have $\mathcal{W}(\cdot) = -P_2(\cdot) - H(\cdot) - P_1(\cdot)$, meaning, $\mathcal{W}(\cdot)$ can be obtained by the inverse process $s(u,t) \rightarrow p_s(s_1, s_2, \cdots) \rightarrow p_x(x_1, x_2, \cdots) \rightarrow x(u,t)$, which is equivalent to (49).

VII. IMPLEMENTATION OF HOGMT PRECODING

HOGMT decomposition is the most important and computational part for HOGMT precoding. Ideally, nonlinear approximation by eigenfunctions is optimal in terms of mean square errors [45]. However, extraction of eigenfunctions is very undesirable [46]. There is no direct method to implement nonlinear HOGMT because of limitation of tools. Here we consider the linear

---

[2]Although precoding involves a linear combination of $\phi_n^*(u,t)$ with x_n. Generally, HOGMT-precoding is a non-linear function ($\mathcal{W}(\cdot)$) with respect to the data signal $s(u,t)$, i.e., $x(u,t)=\mathcal{W}(k_H; s(u,t))$.

alternatives.

A. Equivalent tensor form

Denote the tensor form of $H(t, \tau)$ in (11) as $\mathbf{H} \in \mathbb{C}^{L_u \times L_t \times L_{u'} \times L_\tau}$, where L_u and $L_{u'}$ are the number of users and transmit antennas, and assume $L_u \geq L_{u'}$. L_t and L_τ are the number of data symbols and delay taps. Then the 4-D kernel tensor $\mathbf{K} \in \mathbb{C}^{L_u \times L_t \times L_{u'} \times L_{t'}}$ is obtained by shifting coordination of $\mathbf{H}$, where $L_{t'} = L_t$. HOGMT in (46) decompose the 4-D process into 2-D eigenfunctions, meaning it decompose the 4-D tensor into 2-D jointly orthogonal blocks (eigenmatrices) as in (56)

$$\mathbf{K} = \sum_n \sigma_n \Psi_n \otimes \Phi_n \tag{56}$$

where $\otimes$ is Kronecker product. $\Psi_n \in \mathbb{C}^{L_u \times L_t}$ and $\Phi_n \in \mathbb{C}^{L_{u'} \times L_{t'}}$ are eigenmatrices with properties in (57),

$$\langle \Psi_n, \Psi_{n'}^* \rangle_F = \delta_{nn'} \quad \text{and,} \quad \langle \Phi_n, \Phi_{n'}^* \rangle_F = \delta_{nn'} \tag{57}$$

which is Frobenius product form of (28). Then the duality in (48) is transfered to (58),

$$\langle \mathbf{K}, \Phi_n^* \rangle_F = \sigma_n \Psi_n \tag{58}$$

The transmit space-time signal block is thus the combination of eigenmatrices with optimally derived coefficients from Theorem 2. Higher-order SVD (HOSVD) is one choice to decompose tensor into eigenvectors at each dimension. Then Ψ_n is Kroneker product of eigenvectors at u and t domain, whereas Φ_n is Kroneker product of eigenvecors at u' and t' domain. However, the HOSVD is extremely complex for 4-D tensor, especially for non-truncated kernels [47]. Considering the proposed method just require the decorrelation of space-time domain at transmit and receiver instead of each dimension, we further proposed an implementable (and low-dimensional and low-complexity) alternative as in Lemma 3.

Lemma 3. *(Dimensionality reduction for HOGMT) Given a 4-D tensor* $\mathbf{K} \in \mathbb{C}^{L_u \times L_t \times L_{u'} \times L_{t'}}$ *and an invertible mapping* $f : u \times t \rightarrow m$, *let* $\mathbf{K}' = f(\mathbf{K}) \in \mathbb{C}^{L_m \times L_{m'}}$. *We have*

$$\mathbf{K}' = \mathbf{U \Sigma V}^* = \sum_n z_n \mathbf{u}_n \otimes \mathbf{v}_n^* \tag{59}$$

where z_n *is the singular value.* $\mathbf{u}_n \in \mathbb{C}^{L_m \times 1}$ *and* $\mathbf{v}_n \in \mathbb{C}^{L_{m'} \times 1}$ *are eigenvectors. For (56), there*
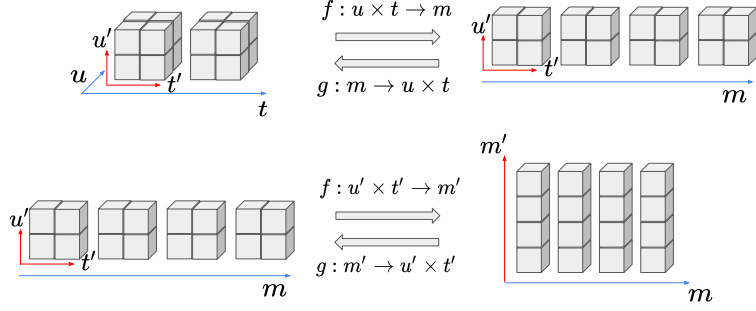
Figure 4: Illustration of dimensionality reduction

exists the equivalent

$$\sigma_n = z_n, \quad \Psi_n = g(\mathbf{u}_n) \quad and \quad \Phi_n = g(\mathbf{v}_n^*) \tag{60}$$

where $g{:}m \to u{\times}t$ is the inverse mapping of f.

Proof.

$$\mathbf{K} = g(\mathbf{K}') = g(\mathbf{U\Sigma V}^*) = g\left(\sum_n z_n \mathbf{u}_n \otimes \mathbf{v}^*{}_n\right)$$

$$= \sum_n z_n g(\mathbf{u}_n) \otimes g(\mathbf{v}^*{}_n) \tag{61}$$

Substituting (61) in (56), we directly have the equivalent (60). □

Figure 4 shows the transition of a $2 \times 2 \times 2 \times 2$ tensor $\mathbf{K}$ to a 4×4 matrix $\mathbf{K}'$. The linear mapping f and g are straightforward.

B. Computational complexity

DPC incurs a much higher runtime complexity (Factorial complexity [48]) compared to HOGMT-precoding (polynomial complexity). Assuming $L_u \geq L_{u'}$, the complexity of HOGMT and DPC are given by Table IV,

Table IV: Complexity Comparison

Strategy	Computational Complexity
HOGMT with Lemma 3	$\mathcal{O}(L_u L_{u'}^2 L_t^3)$
HOGMT with HOSVD	$\mathcal{O}((\frac{L_u + L_{u'} + 2L_t}{4})^5 + L_u L_{u'} L_t^2)$
DPC	$\mathcal{O}(L_t((L_u L_{u'})^{3.5} + L_u L_{u'}^2)L_{u'}!)$

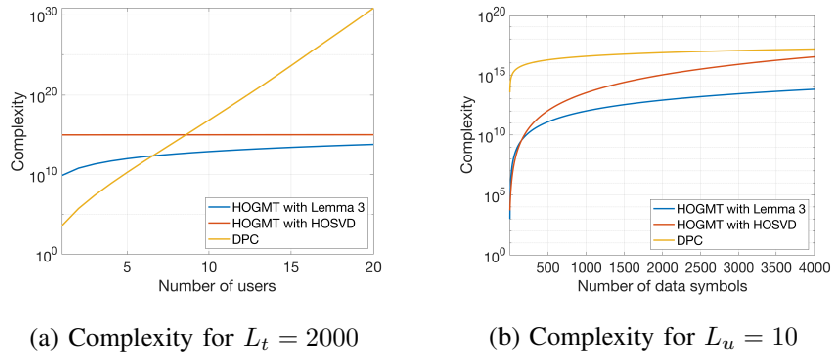(a) Complexity for $L_t = 2000$ (b) Complexity for $L_u = 10$

Figure 5: Complexity comparison for HOGMT with two implementation methods and DPC

Figure 5a compare the complexity of DPC and two implementations of HOGMT with 2000 data symbols with respect to the number of users. When users are more than 8, DPC is much more complex than HOGMT. HOGMT with Lemma 3 has less complexity than HOGMT with HOSVD, though gap narrows as L_u approaches L_t. However, the gap widens in Figure 5b, as L_t is further larger than L_u.

VIII. EVALUATION RESULTS

A. *Practical non-stationary channel simulation framework*

We analyze the accuracy of the proposed joint spatio-temporal precoding using 3GPP 38.901 UMa NLOS senario built on QuaDriga in Matlab. The channel parameters and the layout of the base station (BS) and the user equipment (UE) are shown in Table V. The QuaDriga channel toolbox has been shown to accurately reflect realistic modern channels (*e.g.*, V2X, HST) using practical measurements in [49]. Moreover, this simulated testbed gives the freedom to address a variety of adverse and different scenarios of non-stationary channels that may not be observed without extensive measurement campaigns.

Table V: Non-stationary channel parameters

Senario: 3GPP 38.901 UMa NLOS [50]; Bandwidth: 20 Mhz; Center frequency: 5 Ghz; Subcarriers : 64
BS layout: Array type:3GPP 3-D [51]; Antenna height: 10 m; Antenna number: 10
UE layout: Array type:vehicular [52]; Antenna height: 1.5 m; UE number: 10; Speed: 120 ± 18 km/h

(a) Power delay profile in dB

(b) ACF for the first 1 ms
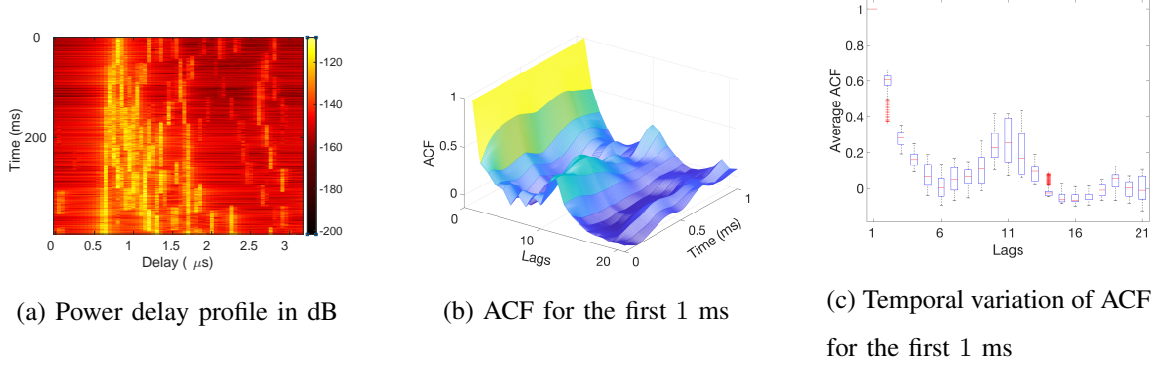
(c) Temporal variation of ACF for the first 1 ms

Figure 6: Power delay profile and statistics distribution for $h_{u,u'}(t,\tau)$ with $u = 1$ and $u' = 1$

Figure 6a shows the power delay profile from the antenna $u' = 1$ to the user $u = 1$ in terms of the 4-D channel i.e., $h_{u,u'}(t,\tau)$, where there exist a drift due to the mobility of the user, leading to the time-varying distribution in overview. We further measure autocorrelation function (ACF) of time-varying impulse response $h_{1,1}(t,\tau)$ in Figure 6b for the first 1 ms. ACF changing over time corroborates its non-stationarity, even within 1 ms. Figure 6c shows the distribution of statistics of ACF (mean and variance), meaning the non-stationarity degree to some extent.

Figure 7a and figure 7b show the spatial channel gains for two fixed time instance (time instances $t=1$ ms and $t=2$ ms both with $\tau = 0$), respectively. The channel gains from other users leads to the spatial (inter-user) interference. We observe that the spatial interference also changes over time. Figure 6 and figure 7 separately show the cause of time-varying temporal interference and spatial interference for $h_{u,u'}(t,\tau)$.

Stationary interval: The *Correlation Matrix Distance (CMD)* is a measure for the degree of stationarity of narrowband MU/MIMO channels [15] and is defined in (62).

$$d_{corr}(t, \Delta t) = 1 - \frac{\langle \mathbf{R}(t), \mathbf{R}(t + \Delta t)\rangle_F}{||\mathbf{R}(t)||_F ||\mathbf{R}(t + \Delta t)||_F} \tag{62}$$

where $|| \cdot ||_F$ is the Frobenius norm and $\mathbf{R}$ is the correlation matrix. In our work, the CMD is calculated at both the transmitter (Tx) and Receiver (Rx) sides, and the corresponding correlation matrices for the narrowband channel $\tilde{\mathbf{H}}(t) = \int \mathbf{H}(t,\tau)d\tau$ over period T are given by (63).

$$\mathbf{R_{Tx}}(t) = \frac{1}{T} \int_t^{t+T} \tilde{\mathbf{H}}(t)^T \tilde{\mathbf{H}}(t)^* dt \tag{63}$$

$$\mathbf{R_{Rx}}(t) = \frac{1}{T} \int_t^{t+T} \tilde{\mathbf{H}}(t) \tilde{\mathbf{H}}(t)^H dt$$

Consequently, the time-varying stationary interval is defined as the largest duration over which

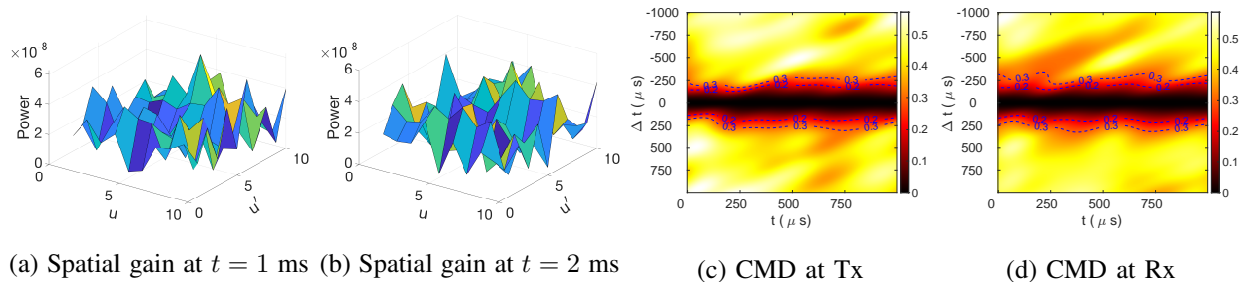(a) Spatial gain at $t = 1$ ms (b) Spatial gain at $t = 2$ ms   (c) CMD at Tx   (d) CMD at Rx

Figure 7: Time-varying spatial channel gains and stationary intervals

CMD remains below a predefined threshold d_0, i.e., $\mathcal{T}(t) = |\Delta t_{max}(t) - \Delta t_{min}(t)|$ where,
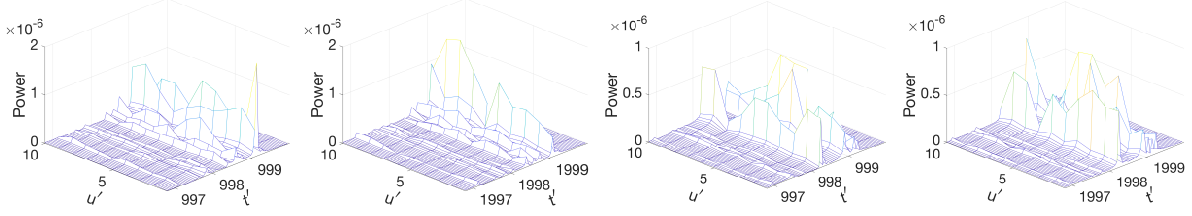
$$\Delta t_{max}(t) = \arg\max_{\Delta t \leq 0} d_{corr}(t, \Delta t) \geq d_0$$

$$\Delta t_{min}(t) = \arg\min_{\Delta t \geq 0} d_{corr}(t, \Delta t) \geq d_0$$

Figure 7c and figure 7d show the CMD at the Tx and Rx, respectively. We observe that, the stationary interval for $d_0 = 0.2$ and $d_0 = 0.3$ (in table I, this threshold for V2X is 0.2, and for HST is 0.7 - 0.9.) are about 400 μs and 500 μs, respectively. Here, the stationarity interval is presented in time instead of distance, as the varying mobility profiles of the multiple users lead to different distances over which stationarity holds. We observe a lower SI compared to that reported in table I, due to the following reasons: a) we consider the multi-user channel, where the varying mobility profiles of spatial elements causes rapid changes in the spatial correlations, b) the rich and dynamic propagation environment in the 3GPP UMa NLOS channel model results in adverse scattering in the simulation environment compared to those observed in the reported experiments in table I. However, as the precoding is able to achieve low BER for smaller SI (higher degree of NS), it would also ensure low BER for larger SI observed in the reported measurement campaigns.

B. Performance of the proposed precoding

The proposed joint spatio-temporal precoding involves extracting the 4-D channel kernel $k_H(u, t; u', t') = h_{u,u'}(t, t - t')$. Figure 8 shows the 4-D channel kernel for $u = 1$ and $u = 2$ at $t = 1000$ μs and $t = 2000$ μs , respectively, where at each instance, the response for user $u = 1$ and $u = 2$ are not only affected by their own delay and other user's spatial interference, but also affected by other users' delayed symbols, which leads to space-time varying joint space-

(a) $u = 1$ and $t = 1000$ μs (b) $u = 1$ and $t = 2000$ μs (c) $u = 2$ and $t = 1000$ μs (d) $u = 1$ and $t = 2000$ μs

Figure 8: 4-D kernel $k_H(u, t; u', t')$

(a) $\phi_1(u', t')$ (b) $\psi_1(u, t)$ (c) $\phi_2(u', t')$ (d) $\psi_2(u, t)$
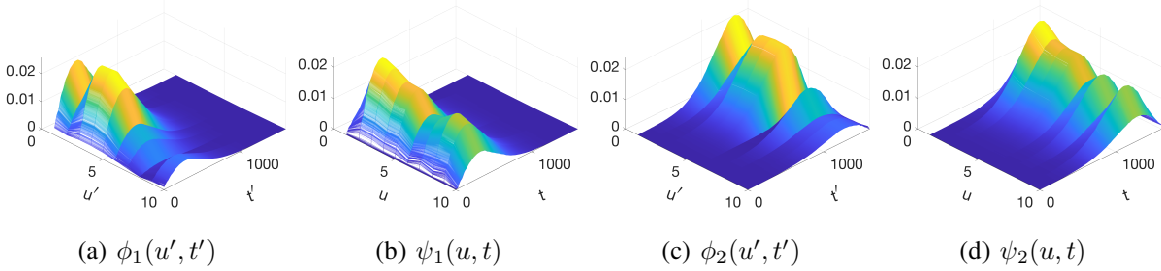
Figure 9: Dual spatio-temporal eigenfunctions decomposed from kernel $k_H(u, t; u', t')$

time interference. This dual spatio-temporal variation necessitates joint spatio-temporal precoding using dual 2-dimensional eigenfunctions, which are dual joint space-time orthogonal.

Figure 9 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 (46). We see that this decomposition is indeed asymmetric as each $\phi_n(u', t')$ and $\psi_n(u', t')$ are not equivalent, 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 (47), 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. Although Theorem 1 decomposes the channel kernel into infinite eigenfunctions, it is sufficient to approximate the channel kernel with a finite number of eigenfunctions with most eigenvalues in terms of mean square errors [45]. These eigenfunctions are used to calculate the coefficients for joint spatio temporal precoding, which subsequently construct the precoded signal after inverse KLT.

In Theorem 2, the energy allocated to n^{th} eigenfunction ϕ_n is $e_n = x_n^2$. As the data signal is directly reconstructed by eigenfunctions $\{\psi_n\}$ with $\{s_n\}$, the reconstruction with n^{th} eigenfunc-

(a) Normalized cumulative function of eigenvalues λ, cost energy e_n and cancelled interference energy e'_n, where both e_n and e'_n are in dB

(b) BER of HOGMT based spatial-temporal precoding using different number of eigenfunctions and comparison with the state-of-the-art

(c) BER of HOGMT based spatio-temporal precoding for BPSK, QPSK, 16-QAM and 64-QAM modulations with 99% eigenfunctions
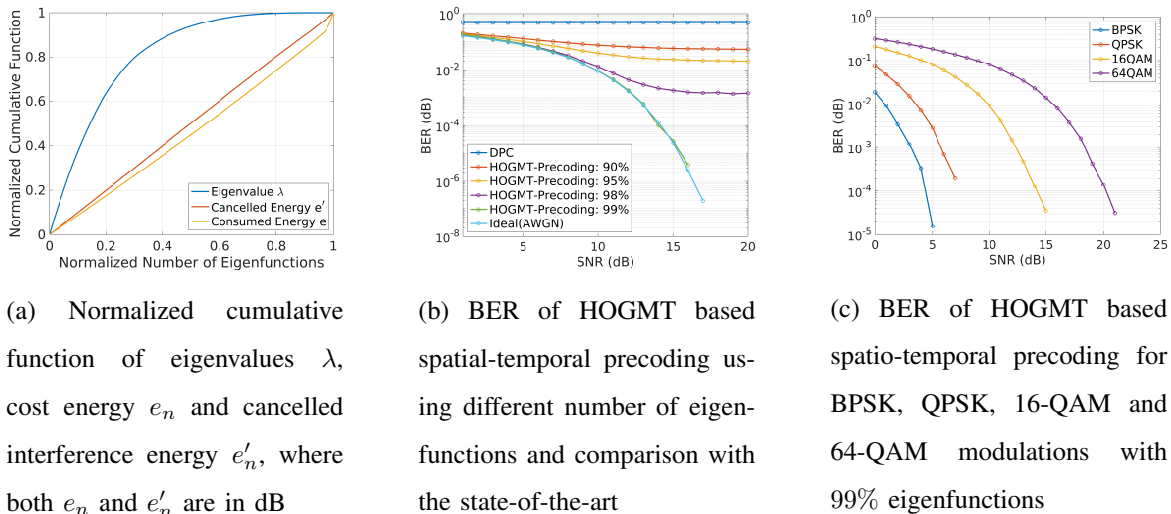
Figure 10: HOGMT based spatio-temporal precoding

tion ψ_n is equivalent to cancelling an interference $-\psi_n$ with energy $e'_n = s_n^2$. The eigenvalue λ_n in (36) is the transmission gain for n^{th} eigenfuntion. From (49), we have $e_n = e'_n/\lambda_n$, meaning more energy is required to cancel interference with more energy and less transmission gains.

Figure 10a shows the normalized cumulative function of eigenvalues λ, cost energy e_n and cancelled interference energy e'_n with respect to the normalized number of eigenfunctions (with descending order of eigenvalues). We observe that, the interference can be fully cancelled by using all eigenfuntions. However, for the last few eigenfuntions with least eigenvalues, it needs more energy, especially using more than 98% eigenfuntions. Meanwhile, the interference cancellation, i.e., e'_n is basically linear with respect to the number of eigenfuntions used. That's because the linear implementation only decomposing eigenfuntions for approximating the kernel. Eigenfunctions with more eigenvalue extract more information of the channel channel, however, are just orthogonal basis with basically equal contributions for reconstructing the data signal. Theoretically, there exists an optimal nonlinear implementation with respect to maximum energy efficiency, which is beyond the scope of this work.

Figures 10b 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 space-time varying interference that occurs in space, time and across space-time dimensions which are shown in figure 8, it achieves significantly lower BER over DPC, which is existing interference-free precoding, however, is applicable for the

time-invariant channel matrix thus shows catastrophic performance for non-stationary channels. Further, we show that with more eigenfuntions, proposed methods achieve lower BER. With more than 99% eigenfunctions, proposed method can achieve near ideal BER, where the ideal case assumes all interference is cancelled and only AWGN noise remains at the receiver. Figure 10c 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 using 99% eigenfunctions. 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 ($\approx 10^{-4}$ 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.

IX. Conclusion

In this work, we derived a high-order generalized version of Mercer's Theorem to decompose the high-order asymmetric kernels into dual 2-dimensional jointly orthogonal eigenfuntions. Through theoretical analysis and simulations, we draw three firm conclusions for non-stationary channels: 1) the 2-dimensional eigenfunctions decomposed from 4-dimensional coefficients of atomic channels across time-frequency and delay-Doppler domain 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) The duality and joint orthogonality of 2-dimensional eigenfunctions decomposed from 4-dimensional non-stationary channels manifest independently flat-fading, 3) precoding by these eigenfunctions with optimally derived coefficients mitigates the spatio-temporal interference , 4) the precoded symbols when propagated over the non-stationary channel directly reconstruct the modulated symbols at the receiver when combined with the calculated coefficients, consequently alleviating the need for complex complementary step at the receiver and 5) proposed precoding has less complexity than DPC. Therefore, the encouraging results from this work will form the core of robust and unifed characterization and highly reliable communication over non-stationary channels, supporting emerging application.

X. ACKNOWLEDGEMENT

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. Kons, M. Tsatsanis, A. Monk, C. Ibars, 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] 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

[11] F. Hlawatsch and G. Matz, *Wireless Communications Over Rapidly Time-Varying Channels*, 1st ed. USA: Academic Press, Inc., 2011.

[12] 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.

[13] J. Bian, C.-X. Wang, X. Gao, X. You, and M. Zhang, "A General 3D Non-Stationary Wireless Channel Model for 5G and Beyond," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3211–3224, 2021.

[14] Z. Zou, M. Careem, A. Dutta, and N. Thawdar, "Unified characterization and precoding for non-stationary channels," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 5140–5146.

[15] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Non-Stationary Narrowband MIMO Inter-Vehicle Channel Characterization in the 5-GHz Band," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 2007–2015, 2010.

[16] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Car-to-car channel models based on wideband MIMO measurements at 5.3 GHz," in *2009 3rd European Conference on Antennas and Propagation*, 2009, pp. 635–639.

[17] T. Zhou, C. Tao, and K. Liu, "Analysis of nonstationary characteristics for high-speed railway scenarios," *Wireless Communications and Mobile Computing*, vol. 2018, pp. 1–7, 06 2018.

[18] D. W. Matolak, "Air-ground channels & models: Comprehensive review and considerations for unmanned aircraft systems," in *2012 IEEE aerospace conference*. IEEE, 2012, pp. 1–17.

[19] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A Survey of Channel Modeling for UAV Communications," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 4, pp. 2804–2821, 2018.

[20] J. J. Jaime-Rodríguez, C. A. Gómez-Vega, C. A. Gutiérrez, J. M. Luna-Rivera, D. U. Campos-Delgado, and R. Velázquez, "A Non-WSSUS Channel Simulator for V2X Communication Systems," *Electronics*, vol. 9, no. 8, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/8/1190

[21] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels," in *2005 IEEE 61st Vehicular Technology Conference*, vol. 1. IEEE, 2005, pp. 136–140.

[22] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *2012 6th European Conference on Antennas and Propagation (EUCAP)*. IEEE, 2012, pp. 433–437.

[23] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The world's first real-time testbed for massive MIMO: Design, implementation, and validation," *IEEE Access*, vol. 5, pp. 9073–9088, 2017.

[24] G. Lovnes, J. Reis, and R. Raekken, "Channel sounding measurements at 59 GHz in city streets," in *5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Wireless Networks - Catching the Mobile Future.*, vol. 2, 1994, pp. 496–500 vol.2.

[25] L. Shi, B. Yao, L. Zhao, C. Wang, H. Wei, and Y. Liu, "Effective transmission method with adaptive nonstationary channel equalization for hypersonic reentry communications," *IEEE Transactions on Plasma Science*, vol. 47, no. 5, pp. 2590–2601, 2019.

[26] S. Wang, Z. Lu, W. Li, S. Jia, L. Zhang, M. Qiao, X. Pang, N. Idrees, M. Saqlain, X. Gao *et al.*, "26.8-m THz wireless transmission of probabilistic shaping 16-QAM-OFDM signals," *APL photonics*, vol. 5, no. 5, p. 056105, 2020.

[27] B. Sharif, J. Neasham, O. Hinton, and A. Adams, "Closed loop doppler tracking and compensation for non-stationary underwater platforms," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158)*, vol. 1. IEEE, 2000, pp. 371–375.

[28] Y. Hatakawa, T. Matsumoto, and S. Konishi, "Development and experiment of linear and non-linear precoding on a real-time multiuser-mimo testbed with limited CSI feedback," in *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)*, 2012, pp. 1606–1611.

[29] F. Hasegawa, H. Nishimoto, N. Song, M. Enescu, A. Taira, A. Okazaki, and A. Okamura, "Non-Linear Precoding for 5G NR," in *2018 IEEE Conference on Standards for Communications and Networking (CSCN)*, 2018, pp. 1–7.

[30] X. Guo, D. Yang, Z. Luo, H. Wang, and J. Kuang, "Robust THP Design for Energy Efficiency of Multibeam Satellite Systems with Imperfect CSI," *IEEE Communications Letters*, vol. 24, no. 2, pp. 428–432, 2020.

[31] F. A. Dietrich, P. Breun, and W. Utschick, "Robust Tomlinson–Harashima Precoding for the Wireless Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 631–644, 2007.

[32] D. Castanheira, A. a. Silva, and A. Gameiro, "Linear and Nonlinear Precoding Schemes for Centralized Multicell MIMO-OFDM Systems," *Wirel. Pers. Commun.*, vol. 72, no. 1, p. 759–777, Sep. 2013. [Online]. Available: https://doi.org/10.1007/s11277-013-1041-z

[33] R. Wang, M. Tao, and Z. Xiang, "Nonlinear Precoding Design for MIMO Amplify-and-Forward Two-Way Relay Systems," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 9, pp. 3984–3995, 2012.

[34] M. Mazrouei-Sebdani, W. A. Krzymień, and J. Melzer, "Massive MIMO With Nonlinear Precoding: Large-System Analysis," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2815–2820, 2016.

[35] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized Precoding for Massive MU-MIMO," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4670–4684, 2017.

[36] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-esposti, 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.

[37] 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.

[38] 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.

[39] 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.

[40] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Overview of Deep Learning-based CSI Feedback in Massive MIMO Systems," 2022. [Online]. Available: https://arxiv.org/abs/2206.14383

[41] M. A. Abdul Careem and A. Dutta, "Real-time prediction of non-stationary wireless channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 7836–7850, 2020.

[42] L. Wang, *Karhunen-Loeve expansions and their applications.* London School of Economics and Political Science (United Kingdom), 2008.

[43] 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.

[44] G. Matz, "On non-WSSUS wireless fading channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2465–2478, 2005.

[45] A. Cohen and J.-P. D'Ales, "Nonlinear approximation of random functions," *SIAM Journal on Applied Mathematics*, vol. 57, no. 2, pp. 518–540, 1997.

[46] K. Liu, T. Kadous, and A. Sayeed, "Orthogonal time-frequency signaling over doubly dispersive channels," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2583–2603, 2004.

[47] R. Badeau and R. Boyer, "Fast multilinear singular value decomposition for structured tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1008–1021, 2008.

[48] Y. Mao and B. Clerckx, "Beyond Dirty Paper Coding for Multi-Antenna Broadcast Channel With Partial CSIT: A Rate-Splitting Approach," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6775–6791, 2020.

[49] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein, "Quadriga-quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Institute, Tech. Rep., 2021.

[50] 3GPP TR 38.901 v16.1.0, "5G; Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Tech. Rep., 11 2020.

[51] 3GPP TR 37.885 v15.1.0, "Technical Specification Group Radio Access Network; Study on evaluation methodology of new Vehicle-to-Everything V2X use cases for LTE and NR," 3rd Generation Partnership Project (3GPP), Tech. Rep., 09 2018.

[52] 3GPP TR 36.873 v12.5.0, "Technical Specification Group Radio Access Network; Study on 3D channel model for LTE," 3rd Generation Partnership Project (3GPP), Tech. Rep., 06 2017.