Received 6 March 2020; accepted 6 April 2020. Date of publication 13 April 2020; date of current version 29 May 2020.

Digital Object Identifier 10.1109/OJCOMS.2020.2987543

A Tensor Based Framework for Multi-Domain Communication Systems

ADITHYA VENUGOPAL AND HARRY LEIB

Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada

CORRESPONDING AUTHOR: H. LEIB (e-mail: harry.leib@mcgill.ca).

This work was supported by the Natural Sciences and Engineering Research Council of Canada through the award entitled "Tensor Modulation for Space-Time-Frequency Communication Systems" under Grant RGPIN-2016-03647.

ABSTRACT The demand for mobile data is likely to grow at a pace more than envisaged in the coming years. Further, as applications such as the Internet of Things (IoT) come to fruition, there will be increased diversity in the types of devices demanding Internet connectivity and their requirements. Significant increase in data rate requirements is also expected due to services such as Ultra High Definition (UHD) video streaming and cloud computing. To meet all these demands, physical layer waveform candidates for future generations of communications need to be robust and inherently capable of extending into multiple domains (space, time, frequency, users, transmission media, code etc.) to ensure efficient utilization of resources. Multiple domains can be innately integrated into the design process of modulation schemes by using tensors, which are multi-way arrays. This paper introduces a unified tensor framework, providing a foundation for multi-domain communication systems that can be used to represent, design and analyse schemes that span several domains. Transmitted signals are represented by Nth order time function tensors which are coupled, using a system tensor of order $N+M$, with the received signals which are represented by another tensor of order M through the contracted convolution. We begin with the continuous time representation of the tensor system model and present both the strict multi-domain generalization of the Nyquist criterion for zero interference (inter-tensor and intra-tensor interference) as well as a relaxation. We present an equivalent discrete time system model, and as an example of using the tensor framework we derive tensor based linear equalization methods to combat multi-domain interference. An application to multi-user MIMO-GFDM illustrates the utility of this novel framework for derivation of joint domain signal processing techniques.

INDEX TERMS Linear equalization, MIMO, multi-domain communication systems, tensor modelling, wireless communication systems.

I. INTRODUCTION

WIRELESS communications and the Internet have been two of the most disruptive technologies in recent history and the synergetic relationship between them has led to exponential demand for mobile communication services. As presented in the visual network index (VNI) report released by Cisco [1], the amount of wireless data has exploded and it is predicted to further grow exponentially in the coming years. Significant increase in data rate requirements are expected due to services such as Ultra High Definition (UHD) video streaming and cloud computing. Hence, future generations of wireless communications will need to provide data rates that are orders of magnitude higher than current 4G technologies. Moreover, with the Internet of Things (IoT) poised to become a reality, many diverse devices with an eclectic mix of requirements will soon demand wireless connectivity to the Internet. In order to service such a vast audience while constrained by radio spectrum scarcity, future communication systems will need to be extremely bandwidth efficient. Given these demands, it is clear that a paradigm shift is required in communication systems of the coming generations (5G and beyond) since incremental improvements on current (4G) systems will not suffice [2].

The use of additional domains in the design process of a communication system is an important means to improve its performance via added robustness from diversity or higher data rates from multiplexing. For example, the addition of the space domain through the utilization of multiple input and multiple output (MIMO) techniques was the logical successor of single input single output (SISO) systems. The use

of MIMO technology boasts improved link performance as in the case of space-time coding [3], or higher data rates via spatial multiplexing such as V-BLAST [4]. Multicarrier (MC) systems such as OFDM, GFDM and FBMC are examples of frequency domain utilization and are significant improvements over single carrier (SC) systems. The two-dimensional structure of these systems are well represented through the use of matrices. In recent years, communication systems that exploit several domains have gained popularity. For instance, multi-domain index modulation in the context of vehicle-to-infrastructure (V2I) and high speed train communication systems is discussed in [5]. Here, the domains used for transmission include the indices of transmit antennas, receive antennas, code type, channel impulse response taps and many more (listed in detail in [5]). Inter-domain communications for in-house networks where a single transmission scheme can be used for multiple types of wires are gaining popularity in [6]. The International Telecommunication Union (ITU) G.hn standard identified the classical in-house mediums such as power lines, twisted-pairs and coax, to enable broadband data communication [7]. Future communication systems are expected to follow this trend of multi-domain communication systems and it is crucial that waveform candidates for future generations of wireless communications be natively capable of extending into *multiple* domains (space, time, frequency, users, transmission media and code to name a few) to ensure efficient utilization of resources.

The use of tensors, which are multidimensional arrays [8], allows innate integration of several domains into the design process of modulation schemes. The notions of tensors and tensor decompositions date back to 1927 with the work of Hitchcock [9]. Cattell [10] is credited for introducing the notion of the multi-way model. However, tensors and their decompositions first gained popularity in psychometrics through the works of Tucker [11], and Carroll and Chang [12]. Since then, tensors have been extensively used in chemometrics in the food industry, in fluorescence spectroscopy and flow injection analysis [13]–[15]. In the last years, tensor applications have gained significant interest in various fields such as signal processing [16], [17], data mining [18], graph analysis [19], neuroscience and computer vision [20], [21]. A tensor approach for multidimensional data filtering is presented in [22]. Cumulant-Based Blind Identification of Under determined Mixtures is explored in [23]. A comprehensive overview of multi-linear algebra, tensor products and their decompositions is provided in [8]. Solution of multi-linear equations using tensor inversion is studied in [24], and a higher-order generalization of the Moore-Penrose pseudoinverse is derived in [25]. The notion of the various transpose operations of a tensor is presented in [26].

Matrix decompositions are not unique in general, meaning that a particular matrix may be decomposed in a number of different ways. In order to ensure uniqueness of a matrix decomposition, additional constraints such as positive-definiteness or orthonormality must be imposed. In contrast, such strong constraints are not required for a tensor to offer a unique decomposition due to the use of higher dimensions [27], [28]. This is one of the reasons for the gain in popularity of tensor based approached in wireless communications over recent years. A blind receiver using PARAFAC decompositions for DS-CDMA systems is considered in [29]. Multiple invariance sensor array processing (MI-SAP) is linked to parallel factor (PARAFAC) analysis for both data-domain and subspace formulations in [30]. A blind receiver that uses tensor decompositions for SIMO and MIMO OFDM systems is presented in [31]. A space-time coding model based on a Khatri-Rao product, dubbed KRST, was derived by combining spatial multiplexing and temporal spreading through linear pre-coding and linear post-coding respectively [32]. Tensor based receivers for MIMO communication systems are presented in [33] and [34]. In [35], it is shown that the received signal in over sampled CDMA and OFDM has a multidimensional structure and a constrained Block-PARAFAC model is used for blind equalization where the constraints of the tensor model vary based on the system that is being used. Three dimensional tensors are used to combine space-time coding with spatial multiplexing, dubbed space-time multiplexing (STM) coding, in [36]. Two constrained tensor models dubbed the PARATUCK-(N_1, N) and Tucker-(N_1, N) are introduced in [37], which are then used to derive semi-blind receivers for MIMO OFDM-CDMA systems. A modified alternating least squares (ALS) algorithm for estimating the matrix factors of the Kronecker product is considered in [38], that is used for the design of MIMO wireless communication systems using tensor modelling. Multidimensional Weiner filtering, where the n-mode unfolding of the desired signal is expressed as a weighted combination of orthogonal vectors from the n-mode signal subspace basis is used to determine the theoretical expression of the n-mode Wiener filter, is described in [22].

In this paper, we present a unified tensor framework for multi-domain communication systems. In Section II, we present the basics of tensors and some important definitions that are used in our work. In Section III we introduce the Tensor Framework. Using this framework, we present the foundations for a multi-domain communication system that can be used to represent, design and analyse future wireless or wired communication systems. In Section IV, to demonstrate the efficacy of the tensor framework, we present examples, such as OFDM, GFDM and FBMC to show how this framework can encapsulate such modulation formats. In Section V, we present both the strict multi-domain generalization of the Nyquist criterion for zero interference (inter-tensor and intra-tensor interference) as well as a relaxation. Then we derive tensor based zero forcing (ZF) and minimum mean squared error (MMSE) linear equalization schemes for multi-domain interference and compare their performance for different system parameters. In Section VI, we study three different types of equalizers that treat inter-tensor interference differently in Multi-user MIMO GFDM systems. Section VII presents the conclusions of this work.

In Appendix A, we present the proof for a tensor Cauchy Schwartz inequality and Appendix B contains the proof of Theorem 2.

II. BASICS OF TENSORS WITH APPLICATIONS TO SIGNALS AND SYSTEMS

A tensor is a multi-dimensional array of data [8]. The order of a tensor is the number of domains. A vector is a tensor of order one, a matrix is a tensor of order two and tensors of order greater than two are known as higher order tensors.

Definition 1 (The Contracted Product): The *contracted product* over K dimensions, or modes, of an Nth order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ and an Mth order tensor $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_M}$ where $I_1 = J_1, \ldots, I_K = J_K$ with $K \leq \min(N, M)$ is a $(N + M - 2K)$th order tensor $\mathcal{C} \in \mathbb{C}^{I_{K+1} \times I_{K+2} \times \cdots \times I_N \times J_{K+1} \times J_{K+2} \times \cdots \times J_M}$ defined as [8]

$$\mathcal{C} = \{\mathcal{A}, \mathcal{B}\}_{(1,\ldots,K;1,\ldots,K)} \tag{1}$$

with $\mathcal{C}_{i_{K+1}\ldots i_N j_{K+1}\ldots j_M} = \sum_{i_1 \ldots i_K} \mathcal{A}_{i_1 \ldots i_K i_{K+1} \ldots i_N} \mathcal{B}_{i_1 \ldots i_K j_{K+1} \ldots j_M}$.

In (1), the modes of contraction are the first K modes of $\mathcal{A}$ and $\mathcal{B}$. However, it should be noted that the modes of contraction do not have to be the same in both tensors, since any two modes of same size can be contracted. For example, the first and second modes of tensor $\mathcal{A} \in \mathbb{C}^{3\times4\times5}$ and the second and third modes of tensor $\mathcal{B} \in \mathbb{C}^{2\times3\times4}$ can be contracted to give a tensor $\mathcal{X} = \{\mathcal{A}, \mathcal{B}\}_{(1,2;2,3)}$ with components $\mathcal{X}_{i_3,j_1} = \sum_{i_1=1}^{3} \sum_{i_2=1}^{4} \mathcal{A}_{i_1,i_2,i_3} \mathcal{B}_{j_1,i_1,i_2}$. A contraction that appears commonly throughout this paper is one where the modes of contraction appear at the end of the first tensor and the beginning of the second. Consider a $(P + N)$th order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_P \times J_1 \times \cdots \times J_N}$ and a $(N + Q)$th order tensor $\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_N \times K_1 \times \cdots \times K_Q}$. The contracted product over the last N modes of $\mathcal{A}$ and the first N modes of $\mathcal{B}$ is a $(P + Q)$th order tensor $\mathcal{C}$

$$\mathcal{C} = \{\mathcal{A}, \mathcal{B}\}_{(P+1,\ldots,P+N;1,\ldots,N)} \tag{2}$$

and $\mathcal{C}_{i_1 \ldots i_P k_1 \ldots k_Q} = \sum_{j_1 \ldots j_N} \mathcal{A}_{i_1 \ldots i_P j_1 \ldots j_N} \mathcal{B}_{j_1 \ldots j_N k_1 \ldots k_Q}$. In the rest of this paper, for the sake of brevity, we use the shorthand notation

$$\mathcal{C} = \{\mathcal{A}, \mathcal{B}\}_{(N)} \tag{3}$$

to denote the contraction in (2). For tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_M \times J_1 \times \cdots \times J_N}$, $\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_N \times K_1 \times \cdots \times K_P}$ and $\mathcal{C} \in \mathbb{C}^{K_1 \times \cdots \times K_P \times L_1 \times \cdots \times L_Q}$, we have

$$\big\{\{\mathcal{A}, \mathcal{B}\}_{(N)}, \mathcal{C}\big\}_{(P)} = \big\{\mathcal{A}, \{\mathcal{B}, \mathcal{C}\}_{(P)}\big\}_{(N)}. \tag{4}$$

Definition 2 (Outer Product): The outer product of two tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \cdots \times J_M}$ is denoted by $\mathcal{A} \circ \mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$ with components

$$(\mathcal{A} \circ \mathcal{B})_{i_1,\ldots,i_N,j_1\ldots,j_M} = \mathcal{A}_{i_1,\ldots,i_N} \mathcal{B}_{j_1,\ldots,j_M} \tag{5}$$

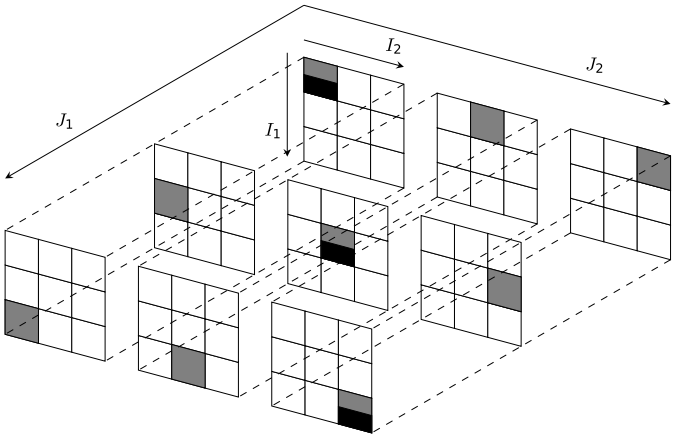

FIGURE 1. Pseudo-diagonal (gray) and diagonal (black) elements of a tensor of size 3 × 3 × 3 × 3.

The tensor outer product is distributive and associative. It is not in general commutative. For tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, $\mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and $\mathcal{C} \in \mathbb{C}^{J_1 \times \cdots \times J_M}$ we have

$$(\mathcal{A} + \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ \mathcal{C} + \mathcal{B} \circ \mathcal{C} \tag{6}$$

Similarly, For tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$,$\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_M}$ and $\mathcal{C} \in \mathbb{C}^{K_1 \times \cdots \times K_P}$ we have

$$(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C}). \tag{7}$$

Definition 3 (Diagonal and Pseudo-Diagonal Tensors): A tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \cdots \times I_N}$ is *diagonal* if

$$\mathcal{A}_{i_1,\ldots,i_N} = \begin{cases} k_{i_1,\ldots,i_N} & \text{if } i_1 = i_2 = \ldots i_N \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $k_{i_1 \ldots i_N}$ is an arbitrary scalar. A pseudo-diagonal tensor is a tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times I_1 \times \cdots \times I_N}$ with components

$$\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_N} = \begin{cases} k_{i_1,\ldots,i_N,j_1,\ldots,j_N} & \text{if } i_1 = j_1 \ldots i_N = j_N \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Authors in [24] and [25] define tensors of the form (9) as diagonal tensors. However, given the stricter, more prevalent, definition of a diagonal tensor [8] the notion of pseudo-diagonality is used in this paper. The non-zero entries of a pseudo-diagonal tensor are known as its pseudo-diagonal entries. A diagonal tensor is a pseudo-diagonal tensor with zeros in some specific pseudo-diagonal entries. Fig. 1 shows a fourth order tensor of size $J_1 \times J_2 \times I_1 \times I_2$ with $I_1 = I_2 = J_1 = J_2 = 3$ with the pseudo-diagonal elements highlighted in gray and diagonal elements highlighted in black. We can see that all the diagonal elements are also pseudo-diagonal elements. We define an *identity tensor* of order $2N$ as a pseudo-diagonal tensor $\mathcal{I}_N \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N \times I_1 \times I_2 \times \cdots \times I_N}$ with entries

$$\mathcal{I}_{N_{i_1,\ldots,i_N,i'_1,\ldots,i'_N}} = \delta_{i_1,i'_1} \cdots \delta_{i_N,i'_N} \tag{10}$$

where $\delta_{x,y}$ is the Kronecker delta defined as

$$\delta_{x,y} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

The sub-script N is used to denote the order of the identity tensor. For example, an identity tensor $\mathcal{I}_N$ is of order $2N$ while $\mathcal{I}_M$ is of order $2M$. For a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \cdots \times I_N \times J_1 \times I_2 \cdots \times J_M}$ we have

$$\{\mathcal{X}, \mathcal{I}_M\}_{(M)} = \{\mathcal{I}_N, \mathcal{X}\}_{(N)} = \mathcal{X}. \tag{12}$$

Definition 4 (Inner Product and Frobenium Norm of a Tensor): The *inner product* of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is defined as

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B} \rangle &= \{\mathcal{A}, \mathcal{B}\}_{(1,\ldots,N;1,\ldots,N)} \\ &= \sum_{i_1} \cdots \sum_{i_N} \mathcal{A}_{i_1,\ldots,i_N} \mathcal{B}_{i_1,\ldots,i_N}\end{aligned} \tag{13}$$

The *Frobenium norm* of a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is defined as

$$\|\mathcal{X}\|_F = \left(\sum_{i_1} \cdots \sum_{i_N} |\mathcal{X}_{i_1,\ldots,i_N}|^2 \right)^{\frac{1}{2}}. \tag{14}$$

Definition 5 (Transpose and Hermitian of a Tensor): A matrix has two indices and the transpose of a matrix is a permutation of these two indices. Since there are several dimensions in a tensor, there are many permutations of its indices and hence there are several ways to write the transpose of a tensor. Authors in [26] define the transpose of a tensor using permutations. For the purpose of this work, we define the transpose of a tensor in the following way. Consider a tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times \cdots \times I_N J_1 \times \cdots \times J_M}$ with a transposition such that the final M modes are swapped with the first N modes such that $\mathcal{Y}^T_{j_1,\ldots,j_M,i_1,\ldots i_N} = \mathcal{Y}_{i_1,\ldots,i_N,j_1,\ldots,j_M}$. For tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_M \times K_1 \times \cdots \times K_P}$ we have

$$\left(\{\mathcal{A}, \mathcal{B}\}_{(M)}\right)^H = \left\{\mathcal{B}^H, \mathcal{A}^H\right\}_{(M)} \tag{15}$$

where superscript H denotes the Hermitian operation.

For tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times \cdots \times J_M \times I_1 \times \cdots \times I_N}$ we have

$$\{\mathcal{B}, \mathcal{A}\}_{(N)} = \left\{\mathcal{A}, \mathcal{B}^T\right\}_{(N)}. \tag{16}$$

Definition 6 (Tensor to Matrix Transformation): For a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$, we define a transformation f_N that transforms $\mathcal{A}$ to a matrix $\mathbf{A} \in \mathbb{C}^{I_1 \ldots I_N \times J_1 \ldots J_M}$ such that $f(\mathcal{A}) = \mathbf{A}$. Component-wise we have

$$\mathcal{A}_{i_1,i_2,\ldots,i_N,j_1,\ldots,j_M} \xrightarrow{f_N} \mathbf{A}_{[i_1 + \sum_{k=2}^{N} (i_k - 1) \prod_{l=1}^{k-1} I_l], [j_1 + \sum_{k=2}^{M} (j_k - 1) \prod_{l=1}^{k-1} J_l]} \tag{17}$$

The subscript in f_N denotes a partition of the modes of the tensor being transformed. The product of the first N modes of the tensor becomes the number of rows of the matrix and the product of the remaining modes of the tensor becomes the number of columns of the matrix. For example, consider a tensor $\mathcal{B} \in \mathbb{C}^{2 \times 3 \times 4 \times 5 \times 6}$ and a transformation f_3 such that $f_3(\mathcal{B}) = \mathbf{B}$. The size of $\mathbf{B}$ is $(2 \cdot 3 \cdot 4) \times (5 \cdot 6)$ and $\mathcal{B}_{i,j,k,l,m} \xrightarrow{f_3} \mathbf{B}_{(i+2(j-1)+6(k-1)),(l+5(m-1))}$. These transformations are called column or row major formats in many computer languages and represent a particular type of matrix unfolding of a tensor. It is shown in [24], for the case of fourth-ordered tensors of the form $\mathcal{X} \in \mathbb{C}^{I \times J \times I \times J}$, that the above transformation function is a bijection with a bijective inverse mapping f_N^{-1} to convert the matrix $\mathbf{A}$ back into the original tensor $\mathcal{A}$. Authors in [39] extend this result to the case of tensors of any order.

Definition 7 (Tensor Inverse): The group of invertible $N \times N$ matrices with matrix multiplication is called the general linear group denoted by $\mathbb{M}_{N,N}(\mathbb{C})$ [24]. Denote

$$\begin{aligned}&\mathbb{T}_{I_1,I_2,\ldots,I_N,I_1,I_2,\ldots,I_N}(\mathbb{C}) \\ &= \left\{\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times I_1 \times \cdots \times I_N} : det(f_N(\mathcal{A})) \neq 0\right\}\end{aligned} \tag{18}$$

Authors in [39] and [24] (for the special case of fourth order tensors) have shown that the set $\mathbb{T}_{I_1,I_2,\ldots,I_N,I_1,I_2,\ldots,I_N}(\mathbb{C})$ forms a group equipped with the contraction $\{\}_{(N)}$ as defined in (3) and the transformation f_N is an isomorphism between $\mathbb{T}_{I_1,I_2,\ldots,I_N,I_1,I_2,\ldots,I_N}(\mathbb{C})$ and $\mathbb{M}_{(I_1 I_2 \cdots I_N),(I_1 I_2 \cdots I_N)}(\mathbb{C})$. This indicates that for any tensor $\mathcal{A} \in \mathbb{T}_{I_1,I_2,\ldots,I_N,I_1,I_2,\ldots,I_N}(\mathbb{C})$ there exists a tensor $\mathcal{B} \in \mathbb{T}_{I_1,I_2,\ldots,I_N,I_1,I_2,\ldots,I_N}(\mathbb{C})$ such that [39]:

$$\{\mathcal{A}, \mathcal{B}\}_{(N)} = \{\mathcal{B}, \mathcal{A}\}_{(N)} = \mathcal{I}_N. \tag{19}$$

where $\mathcal{I}_N$ is the identity tensor. $\mathcal{B}$ is called the inverse of $\mathcal{A}$ and is denoted by $\mathcal{A}^{-1}$. The *Moore-Penrose inverse* of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \cdots \times I_N \times J_1 \times J_2 \times \cdots \times J_N}$,which is a generalization of the matrix Moore-Penrose inverse, is a tensor $\mathcal{A}^+ \in \mathbb{C}^{J_1 \times J_2 \cdots \times J_N \times I_1 \times I_2 \times \cdots \times I_N}$ that satisfies [25]:

$$\begin{aligned}\left\{\left\{\mathcal{A}, \mathcal{A}^+\right\}_{(N)}, \mathcal{A}\right\}_{(N)} &= \mathcal{A} \\ \left\{\left\{\mathcal{A}^+, \mathcal{A}\right\}_{(N)}, \mathcal{A}^+\right\}_{(N)} &= \mathcal{A}^+ \\ \left\{\mathcal{A}, \mathcal{A}^+\right\}^H_{(N)} &= \left\{\mathcal{A}, \mathcal{A}^+\right\}_{(N)} \\ \left\{\mathcal{A}^+, \mathcal{A}\right\}^H_{(N)} &= \left\{\mathcal{A}^+, \mathcal{A}\right\}_{(N)}.\end{aligned}$$

Definition 8 (Pseudo-Triangular Tensor): A tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times I_1 \times \cdots \times I_N}$ is defined to be pseudo-lower triangular if

$$\begin{aligned}&\mathcal{A}_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} \\ &= \begin{cases} 0 & \text{if } (i'_1 + \sum_{k=2}^{N} (i'_k - 1) \prod_{l=1}^{k-1} I_l) \\ & \geq (i_1 + \sum_{k=2}^{N} (i_k - 1) \prod_{l=1}^{k-1} I_l) \\ a_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} & \text{otherwise} \end{cases}\end{aligned} \tag{20}$$

where $a_{i_1,\ldots,i_N,i'_1,\ldots,i'_N}$ are arbitrary scalars. Similarly, the tensor is said to be pseudo-upper triangular if the inequality in (20) is reversed. It can be readily seen that a lower triangular tensor becomes a lower triangular matrix under the tensor to matrix transformation defined in (17) and a pseudo-upper triangular tensor becomes an upper triangular matrix. In Fig. 2 we illustrate a tensor of size $J_1 \times J_2 \times I_1 \times I_2$,

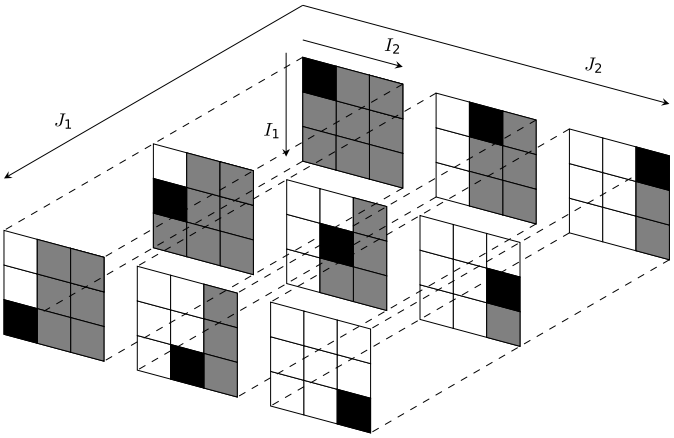

FIGURE 2. Pseudo-Lower Triangular Tensor.

$I_1 = I_2 = J_1 = J_2 = 3$ with the pseudo-lower triangular elements highlighted in gray along with the pseudo-diagonal elements shown in black.

Definition 9 (Function Tensor): A function tensor $\mathcal{A}(x)$ is a tensor whose components are functions of x. Using a third order function tensor as an example, each component of $\mathcal{A}(x)$ is written as $\mathcal{A}_{i,j,k}(x)$. A generalization of this would be the multivariate function tensor $\mathcal{A}(x_1, \ldots, x_p)$, which is a tensor whose components are functions of the variables $x_1, \ldots, x_p$. Using the same example of a third order tensor, each component can be written as $\mathcal{A}_{i,j,k}(x_1, x_2, \ldots, x_p)$. For example, a tensor $\mathcal{A}(t, u)$ is a multivariate function tensor whose components are functions of t and u.

Definition 10 (Signal Tensor and System Tensors): A *signal tensor* $\mathcal{X}(t)$ is a function tensor whose components are functions of time. A *system tensor* $\mathcal{H}(t, u)$, used to describe linear time varying multidomain systems, is a tensor of order $N + M$ that couples two signal tensors of orders N and M respectively through a contracted linear functional. For example, let $\mathcal{H}(t, u) \in \mathbb{C}_t^{Y_1 \times Y_2 \times \cdots \times Y_M \times X_1 \times X_2 \times \cdots \times X_N}$ be a system tensor that couples $\mathcal{X}(t) \in \mathbb{C}_t^{X_1 \times X_2 \times \cdots \times X_N}$ with $\mathcal{Y}(t) \in \mathbb{C}_t^{Y_1 \times Y_2 \times \cdots \times Y_M}$. Here, $\mathbb{C}_t^{A \times B}$ is used to denote the set of tensors of size $A \times B$ whose components are complex functions of t. The output tensor $\mathcal{Y}(t)$ has components

$$\mathcal{Y}_{y_1 y_2 \ldots y_M}(t) = \sum_{x_1 x_2 \ldots x_N} \int_{-\infty}^{\infty} \mathcal{H}_{y_1 y_2 \ldots y_M x_1 x_2 \ldots x_N}(t, u) \mathcal{X}_{x_1 x_2 \ldots x_N}(u) du. \quad (21)$$

Definition 11 (The Contracted Convolution and Time Invariant System Tensor): A linear time invariant system tensor $\mathcal{H}(t)$ is a tensor of order $N + M$ that couples two signal tensors of orders N and M respectively. Extending the contracted product to a contracted convolution allows us to define the coupling of the input and output signal tensors by a linear time invariant system tensor. Consider a signal tensor $\mathcal{X}(t) \in \mathbb{C}_t^{X_1 \times X_2 \times \cdots \times X_N}$ and a system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{Y_1 \times Y_2 \times \cdots \times Y_M \times X_1 \times X_2 \times \cdots \times X_N}$. The contracted convolution of tensor $\mathcal{X}(t)$ and tensor $\mathcal{H}(t)$ is a tensor $\mathcal{Y}(t) \in \mathbb{C}_t^{Y_1 \times Y_2 \times \cdots \times Y_M}$ defined as

$$\mathcal{Y}(t) = \{\mathcal{H}(t) * \mathcal{X}(t)\}_{(M+1,\ldots,M+N;1,2,\ldots N)}, \quad (22)$$

where

$$\mathcal{Y}_{y_1 \ldots y_M}(t) = \sum_{x_1 \ldots x_N} \int_{-\infty}^{\infty} \mathcal{H}_{y_1 \ldots y_M x_1 \ldots x_N}(t - \tau) \mathcal{X}_{x_1 \ldots x_N}(\tau) d\tau. \quad (23)$$

Let $\mathcal{X}(t) \in \mathbb{C}_t^{X_1 \times \cdots \times X_N}$ be a signal tensor and $\mathcal{H}(t) \in \mathbb{C}_t^{Y_1 \times \cdots \times Y_M \times X_1 \times \cdots \times X_N}$ and $\mathcal{G}(t) \in \mathbb{C}_t^{Z_1 \times \cdots \times Z_P \times Y_1 \times \cdots \times Y_M}$ be system tensors. It can readily be seen that the conditions for associativity from (4) are also valid for function tensors since the only change is that multiplications are replaced by convolutions and scalars with functions. If the output of the cascade of these two systems is denoted by $\mathcal{Z}(t) \in \mathbb{C}_t^{Z_1 \times Z_2 \times \cdots \times Z_P}$, we have

$$\begin{aligned} \mathcal{Z}(t) &= \big\{\{\mathcal{G}(t) * \mathcal{H}(t)\}_{(M)} * \mathcal{X}(t)\big\}_{(N)} \\ &= \big\{\mathcal{G}(t) * \{\mathcal{H}(t) * \mathcal{X}(t)\}_{(N)}\big\}_{(M)}. \end{aligned} \quad (24)$$

Definition 12 (The Fourier Transform of Signal Tensors): The Fourier transform of a signal tensor is a tensor of the Fourier transforms of its individual components. If the Fourier transforms of all the individual functions exist, then $\breve{\mathcal{X}}(f)$ the Fourier transform of $\mathcal{X}(t) \in \mathbb{C}_t^{I_1 \times \cdots \times I_N}$ has components

$$\breve{\mathcal{X}}_{x_1 \ldots x_N}(f) = \int_{-\infty}^{\infty} \mathcal{X}_{x_1 \ldots x_N}(t) e^{-j2\pi ft} dt = \mathcal{F}\big[\mathcal{X}_{x_1 \ldots x_N}(t)\big]. \quad (25)$$

Definition 13 (Random Tensor): A tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is said to be *random* if its components $\mathcal{X}_{i_1,\ldots,i_N}$ are random variables. Similarly, a function tensor $\mathcal{A}(x) \in \mathbb{C}_x^{I_1 \times \cdots \times I_N}$ is a random function tensor if its components are random processes.

Definition 14 (Mean): The mean of a random tensor sequence $\mathcal{X}[k] \in \mathbb{C}_k^{I_1 \times \cdots \times I_N}$ is defined as

$$\mu_{\mathcal{X}}[k] = \mathbb{E}\big[\mathcal{X}[k]\big] \quad (26)$$

with components $\mu_{\mathcal{X}_{i_1,\ldots,i_N}}[k] = \mathbb{E}\left[\mathcal{X}_{i_1,\ldots,i_N}[k]\right]$.

Definition 15 (Auto-Correlation and Cross-Correlation of a Random Tensor Sequence): The *auto-correlation* function of a random tensor sequence $\mathcal{X}[k] \in \mathbb{C}_k^{I_1 \times \cdots \times I_N}$ is a tensor $\mathcal{R}_{\mathcal{X}}[k, i] \in \mathbb{C}_{(k,i)}^{I_1 \times \cdots \times I_N \times I_1 \times \cdots \times I_N}$ defined as

$$\mathcal{R}_{\mathcal{X}}[k, i] = \mathbb{E}\big[\mathcal{X}[k] \circ \mathcal{X}^*[k - i]\big] \quad (27)$$

The pseudo-diagonal elements of $\mathcal{R}_{\mathcal{X}}[k, i]$, $\mathcal{R}_{\mathcal{X}_{i_1,\ldots,i_N,i_1,\ldots,i_N}}[k, i]$, are the auto-correlation functions of $\mathcal{X}_{i_1,\ldots,i_N}[k]$ and the cross-correlation between two different components $\mathcal{X}_{i_1,\ldots,i_N}[k]$ and $\mathcal{X}_{i'_1,\ldots,i'_N}[k]$ is $\mathcal{R}_{\mathcal{X}_{i_1,\ldots,i_N,i'_1,\ldots,i'_N}}[k, i]$. The *cross-correlation* of two random tensor sequences $\mathcal{X}[k] \in \mathbb{C}_k^{I_1 \times \cdots \times I_N}$ and $\mathcal{Y}[k] \in \mathbb{C}_k^{J_1 \times \cdots \times J_M}$ is a tensor $\mathcal{R}_{\mathcal{X},\mathcal{Y}}[k, i] \in \mathbb{C}_k^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$ defined as

$$\mathcal{R}_{\mathcal{X},\mathcal{Y}}[k, i] = \mathbb{E}\big[\mathcal{X}[k] \circ \mathcal{Y}^*[k - i]\big] \quad (28)$$

where $\mathcal{R}_{\mathcal{X},\mathcal{Y}_{i_1,\ldots,i_N,j_1,\ldots,j_M}}[k,i] = \mathbb{E}[\mathcal{X}_{i_1,\ldots,i_N}[k]\mathcal{Y}_{j_1,\ldots,j_M}[k-i]]$.

Definition 16 (Wide Sense Stationary Tensor Sequence): A random tensor sequence $\mathcal{X}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ is said to be wide sense stationary (WSS) if its mean $\mathbb{E}[\mathcal{X}[k]]$ is independent of k and its auto-correlation $\mathbb{E}[\mathcal{X}[k]\circ\mathcal{X}^*[k-i]]$ depends only on i. Two random tensor sequences $\mathcal{X}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ and $\mathcal{Y}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M}$ are jointly WSS if both $\mathcal{X}[k]$ and $\mathcal{Y}[k]$ are WSS and their cross-correlation $\mathbb{E}[\mathcal{X}[k]\circ\mathcal{Y}^*[k-i]]$ depends only on i. In the rest of this paper, auto-correlation and cross-correlation tensors of WSS and jointly WSS tensor sequences are indexed by one variable ($\mathcal{R}_{\mathcal{X}}[i]$ and $\mathcal{R}_{\mathcal{X},\mathcal{Y}}[i]$ respectively). It can be shown that if the input to a linear time invariant system tensor is a WSS tensor sequence, then the output is also WSS and the input and output tensor sequences are jointly WSS.

Definition 17 (Discrete System Tensors): A discrete system tensor is a tensor $\mathcal{H}[n] \in \mathbb{C}_n^{I_1\times\ldots I_N\times J_1\times\cdots\times J_M}$ that couples an input tensor sequence $\mathcal{X}[n] \in \mathbb{C}_n^{J_1\times\cdots\times J_M}$ with an output tensor sequence $\mathcal{Y}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ through a discrete contracted convolution defined as:

$$\mathcal{Y}[k] = \sum_n \{\mathcal{H}[n], \mathcal{X}[k-n]\}_{(M)}. \tag{29}$$

III. THE TENSOR FRAMEWORK FOR A COMMUNICATION SYSTEM

A. TENSOR MODEL

Consider a tensor communication system where $\mathcal{D}[n] \in \mathbb{C}_n^{I_1\times\cdots\times I_N}$ represents the data to be transmitted by the nth tensor symbol. The components of $\mathcal{D}[n]$ may be constellation mapped data symbols or may be precoded data symbols. Let the tensor symbol rate be T, i.e., a data tensor is transmitted at intervals of time T. Then we can represent such a data symbol by $\mathcal{D}[n]\delta(t-nT)$ where $\delta(t)$ is Dirac's delta function. Let the transmit filters, the channel and the receive filters be represented by three system tensors $\mathcal{H}_T(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_P\times I_1\times\cdots\times I_N}$, $\mathcal{H}_C(t) \in \mathbb{C}_t^{K_1\times\cdots\times K_Q\times J_1\times\cdots\times J_P}$ and $\mathcal{H}_R(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N\times K_1\times\cdots\times K_Q}$. The overall system model is presented in Figure 3. In this section, we assume that there is no noise (i.e., $\mathcal{N}(t) = \mathbf{0}_T$). The input to the transmit filter is $\sum_{n=-\infty}^{+\infty} \mathcal{D}[n]\delta(t-nT)$. The dimension of the transmit signal tensor being different from the data tensor allows a unifying representation of various schemes. For example, P would be greater than N if the same data symbol may be sent on multiple components of $\mathcal{X}(t)$. Similarly, when multiple data symbols are sent on a single component of $\mathcal{X}(t)$ then P would be smaller than N. If there is a one-to-one mapping between the symbols and waveforms then P would be equal to N.

The transmit signal tensor of order P is

$$\begin{aligned}\mathcal{X}(t) &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}_T(t) * \mathcal{D}[n]\delta(t-nT)\}_{(N)} \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}_T(t-nT), \mathcal{D}[n]\}_{(N)}\end{aligned} \tag{30}$$

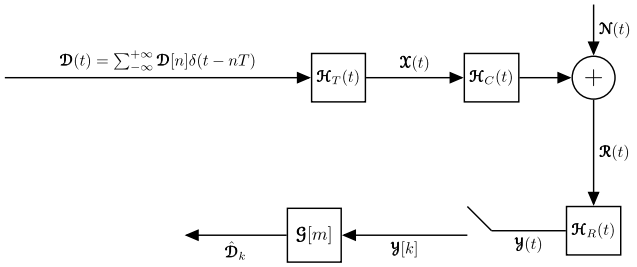

FIGURE 3. System Model.

The effects of the channel on the transmit signal tensor is represented by a contraction with a channel system tensor $\mathcal{H}_C(t)$ of order $(Q+P)$. The received signal $\mathcal{R}(t) \in \mathbb{C}_t^{K_1\times\cdots\times K_Q}$ is

$$\mathcal{R}(t) = \{\mathcal{H}_C(t) * \mathcal{X}(t)\}_{(P)} \tag{31}$$

The receive system tensor $\mathcal{H}_R(t)$ of order $(N+Q)$ transforms the received signal tensor $\mathcal{R}(t)$ into a signal tensor of the same size as the data tensor. The output of the receive filter tensor $\mathcal{Y}(t) \in \mathbb{C}_t^{I_1\times I_2\times\cdots\times I_N}$ is

$$\mathcal{Y}(t) = \{\mathcal{H}_R(t) * \mathcal{R}(t)\}_{(Q)} \tag{32}$$

From (32), (31) and (30) we get

$$\begin{aligned}&\mathcal{Y}(t) \\ &= \left\{\mathcal{H}_R(t) * \{\mathcal{H}_C(t) * \mathcal{X}(t)\}_{(P)}\right\}_{(Q)} \\ &= \left\{\mathcal{H}_R(t) * \left\{\mathcal{H}_C(t) * \sum_{n=-\infty}^{+\infty} \{\mathcal{H}_T(t) * \mathcal{D}[n]\delta(t-nT)\}_{(N)}\right\}_{(P)}\right\}_{(Q)} \\ &= \sum_{n=-\infty}^{+\infty} \left\{\mathcal{H}_R(t) * \left\{\mathcal{H}_C(t) * \{\mathcal{H}_T(t) * \mathcal{D}[n]\delta(t-nT)\}_{(N)}\right\}_{(P)}\right\}_{(Q)}\end{aligned}$$

Using the associativity property (4) we have

$$\begin{aligned}\mathcal{Y}(t) &= \sum_{n=-\infty}^{+\infty} \Big\{\{\mathcal{H}_R(t) * \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(P)}\}_{(Q)} \\ &\qquad \times \mathcal{D}[n]\delta(t-nT)\Big\}_{(N)} \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t) * \mathcal{D}[n]\delta(t-nT)\}_{(N)} \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t-nT), \mathcal{D}[n]\}_{(N)}\end{aligned} \tag{33}$$

where $\mathcal{H}(t) = \{\mathcal{H}_R(t) * \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(P)}\}_{(Q)}$ is the overall system tensor of order $2N$ that couples the input data stream with the output of the receiver. Sampling the received signal tensor $\mathcal{Y}(t)$ at a rate of $\frac{1}{T}$ gives us the estimate of the data tensor

$$\mathcal{Y}[k] = \mathcal{Y}(kT) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(kT-nT), \mathcal{D}[n]\}_{(N)}. \tag{34}$$

B. TENSOR NYQUIST CRITERION

The Nyquist criterion for distortionless transmission for the scalar case is well known. A waveform $x(t)$ is said to satisfy the Nyquist criterion for signal interval T if

$$x(nT) = \delta_n \tag{35}$$

where δ_n is the delta function defined as

$$\delta_n = \begin{cases} 0 & \text{if } n \neq 0 \\ 1 & \text{if } n = 0 \end{cases} \tag{36}$$

Denoting the Fourier transform of $x(t)$ by $X(f)$, and using the Poisson Sum Formula [40]

$$\frac{1}{T}\sum_{k=-\infty}^{+\infty} X\left(f - \frac{k}{T}\right) = \sum_{k=\infty}^{+\infty} x(nT)e^{-j2\pi fnT} \tag{37}$$

we have

$$\frac{1}{T}\sum_{k=-\infty}^{+\infty} X\left(f - \frac{k}{T}\right) = 1 \tag{38}$$

For the matrix-vector case, a generalized Nyquist criterion has been derived in the literature by authors of [41]–[43]. In this section we derive a generalization of (35) and (38), called the *Tensor Nyquist Criterion*, for the multi-domain case with higher-order signal and system tensors. We then show that the existing generalizations are specific cases of the *Tensor Nyquist Criterion*.

1) THE TENSOR POISSON SUM FORMULA AND NYQUIST'S CRITERION FOR ZERO INTER-SYMBOL INTERFERENCE

To find the multi-domain criterion for zero inter-symbol interference we begin by generalizing the ordinary Poisson sum formula [40]. Consider a signal tensor $\mathcal{A}(t) \in \mathbb{C}_t^{I_1 \times I_2 \times \cdots \times I_K}$. Define

$$\mathcal{A}_s(t) = \mathcal{A}(t)\sum_{n=-\infty}^{+\infty} \delta(t - nT) \tag{39}$$

$$= \sum_{n=-\infty}^{+\infty} \mathcal{A}(nT)\delta(t - nT) \tag{40}$$

Taking the Fourier transform of (39) we get

$$\begin{aligned} \breve{\mathcal{A}}_s(f) &= \mathcal{F}\left[\mathcal{A}(t)\sum_{n=-\infty}^{+\infty} \delta(t - nT)\right] \\ &= \breve{\mathcal{A}}(f) * \mathcal{F}\left[\sum_{n=-\infty}^{+\infty} \delta(t - nT)\right] \\ &= \breve{\mathcal{A}}(f) * \left(\frac{1}{T}\sum_{n=-\infty}^{+\infty} \delta\left(f - \frac{n}{T}\right)\right) \qquad (41) \\ &= \frac{1}{T}\sum_{n=-\infty}^{+\infty} \breve{\mathcal{A}}\left(f - \frac{n}{T}\right) \qquad (42) \end{aligned}$$

Taking the Fourier transform of (40) we get

$$\breve{\mathcal{A}}_s(f) = \mathcal{F}\left[\sum_{n=-\infty}^{+\infty} \mathcal{A}(nT)\delta(t - nT)\right] = \sum_{n=-\infty}^{+\infty} \mathcal{A}(nT)e^{-j2\pi fnT} \tag{43}$$

Equating (42) and (43) we get the Tensor Poisson Sum Formula:

$$\frac{1}{T}\sum_{n=-\infty}^{+\infty} \breve{\mathcal{A}}\left(f - \frac{n}{T}\right) = \sum_{n=-\infty}^{+\infty} \mathcal{A}(nT)e^{-j2\pi fnT} \tag{44}$$

Expanding (34) we get

$$\begin{aligned} \mathcal{Y}[k] &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(kT - nT), \mathcal{D}[n]\}_{(N)} \\ &= \{\mathcal{H}(0), \mathcal{D}[k]\}_{(N)} + \sum_{n=-\infty, n\neq k}^{+\infty} \{\mathcal{H}(kT - nT), \mathcal{D}[n]\}_{(N)} \end{aligned} \tag{45}$$

A sufficient condition to get zero interference between symbols is

$$\mathcal{H}(iT) = \begin{cases} \mathcal{H}(0) & \text{if } i = 0 \\ \mathbf{0}_T & \text{if } i \neq 0 \end{cases} \tag{46}$$

where $\mathbf{0}_T$ is the all zero tensor. Using the Tensor Poisson's sum formula (44) we obtain the *Tensor Nyquist Criterion* for zero inter-symbol interference:

$$\frac{1}{T}\sum_{n=-\infty}^{+\infty} \breve{\mathcal{H}}\left(f - \frac{n}{T}\right) = \mathcal{H}(0) = \mathcal{K} \tag{47}$$

where $\mathcal{K}$ is a non-zero tensor.

Assuming that (46) is satisfied we have

$$\mathcal{Y}[k] = \{\mathcal{H}(0), \mathcal{D}[k]\}_{(N)} \tag{48}$$

whose elements are

$$\begin{aligned} \mathcal{Y}_{l_1,\ldots,l_N}[k] &= \sum_{i_1,\ldots,i_N} \mathcal{H}_{l_1,\ldots,l_N,i_1,\ldots,i_N}(0)\mathcal{D}_{i_1,\ldots,i_N}[k] \\ &= \mathcal{H}_{l_1,\ldots,l_N,l_1,\ldots,l_N}(0)\mathcal{D}_{l_1,\ldots,l_N}[k] \\ &\quad + \sum_{\substack{i_1,\ldots,i_N \\ i_1\neq l_1,\ldots,i_N\neq l_N}} \mathcal{H}_{l_1,\ldots,l_N,i_1,\ldots,i_N}(0)\mathcal{D}_{i_1,\ldots,i_N}[k] \end{aligned} \tag{49}$$

We see that the first term in (49) is a scaled version of the required data symbol and the second term represents intra-symbol interference from other data symbols within the same data tensor. A rather strict condition which will ensure that we are able to retrieve the transmitted data from $\mathcal{Y}[k]$ without any interference is

$$\mathcal{H}_{l_1,\ldots,l_N,i_1,\ldots,i_N}(0) = \begin{cases} 1 & \text{if } i_1 = l_1,\ \ldots,\ i_N = l_N \\ 0 & \text{otherwise} \end{cases} \tag{50}$$

This means that the tensor $\mathcal{H}(0)$ is an identity tensor. Combining (46) and (50) we get

$$\mathcal{H}_{l_1,\ldots,l_N,i_1,\ldots,i_N}(iT) = \delta_{i,0}\prod_{k=1}^{N} \delta_{i_k,l_k} \tag{51}$$

where

$$\delta_{n,m} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{otherwise} \end{cases} \tag{52}$$

Using this in (47) we get

$$\frac{1}{T}\sum_{n=-\infty}^{+\infty} \breve{\mathcal{H}}\Big(f - \frac{n}{T}\Big) = \mathcal{H}(0) = \mathcal{I}_N \tag{53}$$

Even if the strict criterion is not met, it is still possible to recover the transmitted data from $\mathcal{Y}[k]$ if certain conditions are met. Assuming that (46) holds then recovering the data reduces to solving the multi-linear tensor system (48) for $\mathcal{D}[k]$. If the inverse of $\mathcal{H}(0)$ exists, we have

$$\mathcal{D}[k] = \Big\{\mathcal{H}^{-1}(0), \mathcal{Y}[k]\Big\}_{(N)} \tag{54}$$

where $\mathcal{H}^{-1}(0)$ can be approximated by using iterative algorithms, such as the biconjugate gradient or Jacobi methods using tensor computations [24].

If the inverse does not exist, pseudo-inversion can be used to find the solution to the multi-linear system (48). The tensors $\hat{\mathcal{D}}[k]$ minimizing $\big\|\{\mathcal{H}(0), \mathcal{D}[k]\}_{(N)} - \mathcal{Y}[k]\big\|_F^2$ are called the least-square solutions of (48) and $\widetilde{\mathcal{D}}[k] = \min_{\hat{\mathcal{D}}[k]} \|\hat{\mathcal{D}}[k]\|_F^2$ is called the minimum-norm least square solution of (48) [25]. If $\{\mathcal{H}^H(0), \mathcal{H}(0)\}_{(N)}$ is invertible then the least-square solution has a unique minimizer and the multilinear system is solved as [24]

$$\mathcal{D}[k] = \Big\{\Big\{\Big\{\mathcal{H}^H(0), \mathcal{H}(0)\Big\}_{(N)}^{-1}, \mathcal{H}^H(0)\Big\}_{(N)}, \mathcal{Y}[k]\Big\}_{(N)} \tag{55}$$

Finally, if such an inversion does not exist, then the minimum-norm least square solution of (48) is

$$\mathcal{D}[k] = \Big\{\mathcal{H}^{+}(0), \mathcal{Y}[k]\Big\}_{(N)} \tag{56}$$

where $\mathcal{H}^{+}(0)$ is the Moore-Pensore pseudoinverse of $\mathcal{H}(0)$ [25].

2) COMPARISON WITH EXISTING GENERALIZATIONS OF NYQUIST'S CRITERION

This section surveys existing generalizations of the Nyquist Criterion and compares them to the Tensor Nyquist Criterion presented in this paper. The problem of interference in a multi-carrier system is considered in [41] and [42], that propose a constraint on the overall system impulse response to simultaneously eliminate both ISI and cross-talk. Based on previous work, [43] presents a multidimensional Nyquist criterion. The general system considered in [41]–[43] is a multi-carrier system specified as

$$b_r(t) = \sum_{n=-\infty}^{+\infty}\sum_{m=1}^{M} a_{nm} v_{mr}(t - nT) \quad r = 1, 2, \ldots, M \tag{57}$$

where a_{nm} is the data transmitted on the mth sub-carrier during the nth symbol and $v_{mr}(t)$ represents the overall system impulse response consisting of the mth transmit filter, the channel and the rth receive filter, $v_{mr}(t) = p_m(t) * b(t) * r_r(t)$ where $p_m(t)$ is the mth transmit filter, $b(t)$ is the channel and $r_r(t)$ is the rth receive filter. Representing (57) in vector matrix form we get

$$\mathbf{b}(t) = \sum_n \mathbf{V}(t - nT)\mathbf{a}_n \tag{58}$$

where $\mathbf{b}(t) = [b_1(t), \ldots b_M(t)]^{\mathrm{T}}$, $\mathbf{a}_n = [a_{n1}, \ldots a_{nM}]^{\mathrm{T}}$ and

$$\mathbf{V}(t) = \begin{bmatrix} v_{11}(t) & v_{21}(t) & \ldots & v_{M1}(t) \\ v_{12}(t) & v_{22}(t) & \ldots & v_{M2}(t) \\ \vdots & \vdots & \vdots & \vdots \\ v_{1M}(t) & v_{2M}(t) & \ldots & v_{MM}(t) \end{bmatrix}$$

Sampling $\mathbf{b}(t)$ at rate $\frac{1}{T}$ we get [41]

$$\mathbf{b}(kT) = \sum_n \mathbf{V}(kT - nT)\mathbf{a}_n \tag{60}$$

For no interference, it is required that $\mathbf{b}(kT) = \mathbf{a}_k$. The generalized Nyquist criterion to achieve this is [41], [42]

$$\mathbf{V}(iT) = \delta_{i,0}\mathbf{I} \tag{61}$$

where $\mathbf{I}$ is the identity matrix. Taking the Fourier transform of $\mathbf{V}(t)$ and using the Poisson's sum formula we get the frequency domain conditions for zero interference as [42]

$$\frac{1}{T}\sum_n \breve{\mathbf{V}}\Big(f - \frac{n}{T}\Big) = \frac{1}{T}\sum_n \breve{\mathbf{p}}\Big(f - \frac{n}{T}\Big)\breve{b}\Big(f - \frac{n}{T}\Big)\breve{\mathbf{r}}^T\Big(f - \frac{n}{T}\Big) = \sum_{n=-\infty}^{+\infty} \mathbf{V}(nT)e^{-j2\pi fnT} = \delta_{n,0}\mathbf{I} \tag{62}$$

where $\breve{\mathbf{p}}(f) = [\breve{p}_1(f)\ \breve{p}_2(f)\ \ldots\ \breve{p}_M(f)]^T$, $\breve{b}(f)$ and $\breve{\mathbf{r}}^T(f) = [\breve{r}_1(f)\ \breve{r}_2(f)\ \ldots\ \breve{r}_M(f)]^T$ are the Fourier transforms of the transmit filters, channel and receive filters respectively. Comparing (62) with (53) we can see that it is a special case of the generalized Nyquist criterion where the overall system tensor is of size $M \times M$ with components $\mathbf{H}_{i_1,i_2}(t) = p_{i_2}(t) * b(t) * r_{i_1}(t)$ and the data tensor is of order one (a vector of size M) with components $\mathbf{d}[n] = \mathbf{a}_n$.

IV. SYSTEMS MODELED USING THE TENSOR FRAMEWORK

This section presents models for selected waveforms using the tensor framework as examples. Representation using the tensor framework preserves the multi-domain structure of the transmitted data and allows extensions to higher domains, such as MIMO, of waveforms that were not originally designed for this.

A. OFDM

In Release 15 of 3GPP [44] it is mentioned that an Orthogonal Frequency Division Multiplexing (OFDM) based scheme will be used for the 5G New Radio (NR) uplink and downlink as the main candidate, with Discrete Fourier Transform Spread OFDM (DFT-S-OFDM) being used in

some cases. OFDM is a multi-carrier transmission technique which uses F orthogonal sub-carriers simultaneously to transmit data. By making all the sub-channels narrowband, they experience almost flat fading, making equalization very simple. The orthogonality of the sub-carriers ensures that there is no intrinsic inter-carrier interference (ISI). There are F sub-carriers with spacing $F_0 = \frac{1}{T}$ that carry data in one OFDM symbol. The transmit signal is

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d=1}^{F} d_{n,\kappa_d} w(t-nT) e^{j2\pi(\kappa_d-1)F_0 t} \\ &= \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d=1}^{F} d_{n,\kappa_d} p_{T_{\kappa_d}}(t-nT) \end{aligned} \tag{63}$$

where d_{n,κ_d} is the complex data symbol transmitted during the nth OFDM symbol on the κ_dth ($\kappa_d = 1, \ldots, F$) sub-carrier. The filter $p_{T_{\kappa_d}}(t) = w(t)e^{j2\pi(\kappa_d-1)F_0 t}$ where $w(t)$ is a rectangular window of duration T defined as

$$w(t) = \begin{cases} 1 & \text{for } -\frac{T}{2} \le t \le \frac{T}{2} \\ 0 & \text{otherwise} \end{cases} \tag{64}$$

Under the assumption of an ideal channel, the received signal $r(t)$ is the same as the transmitted signal $x(t)$. The receive filter, for the κ_yth subcarrier, $p_{R_{\kappa_y}}(t) = p^*_{T_{\kappa_y}}(-t)$ is matched to the transmit filter, i.e., $p_{R_{\kappa_y}}(t) = w(-t)e^{j2\pi(\kappa_y-1)F_0 t}$. The output of the receive filter $p_{R_{\kappa_y}}(t)$ is

$$\begin{aligned} y_{\kappa_y}(t) = x(t) * p_{R_{\kappa_y}}(t) = \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d=1}^{F} \int_{-\infty}^{+\infty} d_{n,\kappa_d} w(\tau - nT) \\ \times\ e^{j2\pi F_0(\kappa_d-1)\tau} w(-(t-\tau)) e^{j2\pi F_0(\kappa_y-1)(t-\tau)} d\tau \end{aligned} \tag{65}$$

Sampling $y_{\kappa_y}(t)$ at intervals of T and using $w(t) = w(-t)$ gives

$$\begin{aligned} & y_{\kappa_y}(kT) \\ &= \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d=1}^{F} \int_{-\infty}^{+\infty} \Big(d_{n,\kappa_d} w(\tau - nT) e^{j2\pi F_0(\kappa_d-1)(\tau)} \\ &\qquad \times w(kT-\tau) e^{j2\pi F_0(\kappa_y-1)(kT-\tau)} \Big) d\tau \\ &= \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d=1}^{F} \int_{-\infty}^{+\infty} d_{n,\kappa_d} w(\tau - nT) w(\tau - kT) e^{j2\pi F_0(\kappa_d-\kappa_y)(\tau)} d\tau \end{aligned} \tag{66}$$

Since $w(t)$ is a rectangular window of duration T we have $y_{\kappa_y}(kT) = d_{k,\kappa_y}$.

Using the tensor framework, the complex data to be transmitted on the nth OFDM symbol $\mathcal{D}[n] \in \mathbb{C}_n^F$ is

$$\mathcal{D}[n] = \left[d_{n,1}, d_{n,2}, \ldots, d_{n,F} \right]^T \tag{67}$$

with components $\mathcal{D}_i[n] = d_{n,i}$ for $i = 1, \ldots, F$. The transmit system tensor $\mathcal{H}_T(t) \in \mathbb{C}_t^{1\times F}$ is

$$\mathcal{H}_T(t) = \left[p_{T_1}(t), p_{T_2}(t), \ldots, p_{T_F}(t) \right] \tag{68}$$

and has components $\mathcal{H}_{T_{1,i}}(t) = p_{T_i}(t)$ for $i = 1, \ldots, F$. The transmit tensor $\mathcal{H}_T(t)$ converts $\mathcal{D}[n]$ into a signal $\mathcal{X}(t) \in \mathbb{C}_t$. We can write (63) using the components of (67) and (68) as $\mathcal{X}(t) = \sum_{n=-\infty}^{+\infty} \sum_{i=1}^{F} \mathcal{H}_{T_{1,i}}(t-nT)\mathcal{D}_i[n]$ which, in tensor notation, becomes

$$\mathcal{X}(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}_T(t-nT), \mathcal{D}[n]\}_{(1)} \tag{69}$$

Under the assumption of an ideal channel the channel tensor is $\mathcal{H}_C(t) \in \mathbb{C}_t = \delta(t)$ and the received signal is $\mathcal{R}(t) = \mathcal{X}(t)$. The receive system tensor $\mathcal{H}_R(t) \in \mathbb{C}_t^{F\times 1}$ is

$$\mathcal{H}_R(t) = \left[p_{R_1}(t), p_{R_2}(t), \ldots, p_{R_F}(t) \right]^T \tag{70}$$

Using (68) and (70), we get the overall system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{F\times F}$ as $\mathcal{H}(t) = \{\mathcal{H}_R(t) * \mathcal{H}_T(t)\}_{(1)}$ with components

$$\mathcal{H}_{\kappa_y,\kappa_d}(t) = \mathcal{H}_{R_{\kappa_y}}(t) * \mathcal{H}_{T_{\kappa_d}}(t) = p_{R_{\kappa_y}}(t) * p_{T_{\kappa_d}}(t). \tag{71}$$

Collecting the outputs of the received filters from (65) for all sub-carriers into a vector we may now write the received signal tensor $\mathcal{Y}(t) \in \mathbb{C}_t^F$ as $\mathcal{Y}(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t-nT), \mathcal{D}[n]\}_{(1)}$ with components

$$\mathcal{Y}_{\kappa_y}(t) = \sum_{n=-\infty}^{+\infty} \sum_{\kappa_d} \mathcal{H}_{\kappa_y,\kappa_d}(t-nT)\mathcal{D}_{\kappa_d}[n] \tag{72}$$

Comparing $\mathcal{Y}(t)$ with (33), we can see that this is a specific case when $N = 1$.

B. FBMC

Filter Bank Multi-carrier is a scheme considered for 5G. There are two types of FBMC schemes, Staggered Multitone (SMT) and Cosine-Modulated Multitone (CMT) [45]. This section describes SMT, also known as OQAM/OFDM. The number of sub-carriers K is assumed to be even ($K = 2M$) and for two consecutive sub-carriers, the time offset is applied to the imaginary part of the QAM symbol on one sub-carrier while it is applied to the real part of the QAM symbol on the other sub-carrier. The transmitted signal is [46]

$$\begin{aligned} x(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{M-1} \Bigg(& c^R_{2m,n} p(t-nT) \\ & + j c^I_{2m,n} p\left(t - \frac{T}{2} - nT\right) \Bigg) e^{j2\pi(2m)Ft} \\ & + \left(j c^I_{2m+1,n} p(t-nT) + c^R_{2m+1,n} p\left(t - \frac{T}{2} - nT\right) \right) \\ & \times\ e^{j2\pi(2m+1)Ft} \end{aligned} \tag{73}$$

where T is the signalling interval, $F = \frac{1}{T}$ is the sub-carrier spacing, $c^R_{m,n}$ and $c^I_{m,n}$ are the real and imaginary components of the QAM symbol $c_{m,n}$ to be transmitted on the mth sub-carrier during the nth multi-carrier symbol, and $p(t)$ is a real symmetric prototype filter of duration KT where K is the overlapping factor that denotes the number of multi-carrier

symbols that overlap in time. We introduce the following notations, as in [46], to simplify the expression in (73):

$$d_{2m,2n} = c^R_{2m,n},\ d_{2m,2n+1} = c^I_{2m,n},$$
$$d_{2m+1,2n} = c^I_{2m+1,n},\ d_{2m+1,2n+1} = c^R_{2m+1,n} \tag{74}$$
$$\psi_{2m,2n} = 0\ \psi_{2m,2n+1} = \frac{\pi}{2}\ \psi_{2m+1,2n} = \frac{\pi}{2}\ \psi_{2m+1,2n+1} = 0 \tag{75}$$

Using (74) and (75) in (73) we get

$$\begin{aligned} x(t) &= \sqrt{2}\sum_{n=-\infty}^{+\infty}\Bigg\{\sum_{m=0}^{M-1}\Bigg(d_{2m,2n}p\left(t-(2n)\frac{T}{2}\right)e^{\psi_{2m,2n}} \\ &\quad + d_{2m,2n+1}p\left(t-(2n+1)\frac{T}{2}\right) \\ &\quad \times e^{\psi_{2m,2n+1}}\Bigg)e^{j2\pi(2m)Ft} \\ &\quad + \Bigg(d_{2m+1,2n}p\left(t-(2n)\frac{T}{2}\right)e^{\psi_{2m+1,2n}} \\ &\quad + d_{2m+1,2n+1}p\left(t-(2n+1)\frac{T}{2}\right)e^{\psi_{2m+1,2n+1}}\Bigg) \\ &\quad \times e^{j2\pi(2m+1)Ft}\Bigg\} \end{aligned} \tag{76}$$

Defining $\lambda_{m,n}(t) = \sqrt{2}p(t-n\frac{T}{2})e^{j2\pi mFt}e^{\psi_{m,n}}$ and substituting in (76) we get

$$\begin{aligned} x(t) = \sum_{n=-\infty}^{+\infty}\sum_{m=0}^{M-1}\Big(&d_{2m,2n}\lambda_{2m,2n}(t) + d_{2m,2n+1}\lambda_{2m,2n+1}(t) \\ &+ d_{2m+1,2n}\lambda_{2m+1,2n}(t) \\ &+ d_{2m+1,2n+1}\lambda_{2m+1,2n+1}(t)\Big) \end{aligned} \tag{77}$$

Since $\sum_{k=-P}^{k=Q}(x_{2k}+x_{2k+1}) = \sum_{k=-2P}^{k=2Q+1}x_k$, we have from (77)

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{+\infty}\sum_{m=0}^{2M-1}\Big(d_{m,2n}\lambda_{m,2n}(t) + d_{m,2n+1}\lambda_{m,2n+1}(t)\Big) \\ &= \sum_{n=-\infty}^{+\infty}\sum_{m=0}^{2M-1}d_{m,n}\lambda_{m,n}(t) \\ &= \sum_{n=-\infty}^{+\infty}\sum_{m=0}^{2M-1}d_{m,n}p_m\left(t-n\frac{T}{2}\right)e^{j\psi_{m,n}} \end{aligned} \tag{78}$$

where $p_m(t) = \sqrt{2}p(t)e^{j2\pi mFt}$. The received signal is passed through an analysis filter bank (AFB) to separate the data from different sub-carriers. The receive filter for the rth sub-carrier, $p_{R_r}(t)$ is matched to the transmit filter for the rth sub-carrier $p_r(t)$, i.e., we have $p_{R_r}(t) = p^*_r(-t) = \sqrt{2}p(-t)e^{j2\pi mFt}$ ($p^*(-t) = p(-t)$ as the prototype filter $p(t)$ is real). Let

$$\begin{aligned} \langle\lambda_{m,n}(t), \lambda_{p,q}(t)\rangle = \int_{-\infty}^{+\infty}&p_m\left(t-n\frac{T}{2}\right)e^{j\psi_{m,n}}p^*_p\left(t-q\frac{T}{2}\right) \\ &\times e^{-j\psi_{p,q}}dt \end{aligned} \tag{79}$$

where $\langle,\rangle$ denotes the inner product. In a distortion free channel, perfect reconstruction of the data is obtained if the transmit and receive filters satisfy the real orthogonality condition [47]:

$$\begin{aligned} &\Re\big\{\langle\lambda_{m,n}(t), \lambda_{p,q}(t)\rangle\big\} \\ &= \Re\left\{\int_{-\infty}^{+\infty}p_m\left(t-n\frac{T}{2}\right)e^{j\psi_{m,n}}p^*_p\left(t-q\frac{T}{2}\right)e^{-j\psi_{p,q}}dt\right\} \\ &= \delta_{m,p}\delta_{n,q} \end{aligned} \tag{80}$$

In other words, this means that for $(m,n) \neq (p,q)$, $\langle\lambda_{m,n}(t), \lambda_{p,q}(t)\rangle$ can be purely imaginary or zero. The received signal after passing through the receive filter corresponding to the rth sub-carrier is

$$\begin{aligned} y_r(t) &= x(t) * p^*_r(-t) \\ &= \sum_{n=-\infty}^{+\infty}\sum_{m=0}^{2M-1}d_{m,n}e^{j\psi_{m,n}}\int_{-\infty}^{+\infty}p_m\left(\tau-n\frac{T}{2}\right)p^*_r(\tau-t)d\tau \end{aligned} \tag{81}$$

Using the tensor framework, the data to be transmitted on the nth multicarrier symbol is a tensor $\mathcal{D}[n] \in \mathbb{C}^{2M}_n$ with components $\mathcal{D}_m[n] = d_{(m-1),n}e^{j\psi_{(m-1),n}}$ and the overall channel is $\mathcal{H}(t) \in \mathbb{C}^{2M\times 2M}_t$ with components $\mathcal{H}_{r,m}(t) = \int_{-\infty}^{+\infty}p_{(m-1)}(\tau)p^*_{(r-1)}(\tau-t)d\tau$. We may now re-write (81) in tensor form as

$$\mathcal{Y}_r(t) = \sum_{n=-\infty}^{+\infty}\sum_{m=1}^{2M}\mathcal{H}_{r,m}\left(t-n\frac{T}{2}\right)\mathcal{D}_m[n] \tag{82}$$

where $\mathcal{Y}_r(t) = y_{(r-1)}(t)$. In tensor notation (82) becomes $\mathcal{Y}(t) = \sum_{n=-\infty}^{+\infty}\{\mathcal{H}(t-n\frac{T}{2}), \mathcal{D}[n]\}_{(1)}$. The signal $y_r(t)$ is sampled at intervals of $k\frac{T}{2}$ and multiplied by the phase term $e^{-j\psi_{r,k}}$. This gives

$$\begin{aligned} y_{r,k} &= y_r\left(k\frac{T}{2}\right)e^{-j\psi_{r,k}} \\ &= \sum_n\sum_m d_{m,n}e^{j\psi_{m,n}}e^{-j\psi_{r,k}}\int_{-\infty}^{+\infty}p_m\left(\tau-n\frac{T}{2}\right) \\ &\quad \times p^*_r\left(\tau-k\frac{T}{2}\right)d\tau \\ &= \sum_n\sum_m d_{m,n}\langle\lambda_{m,n}(t), \lambda_{r,k}(t)\rangle \\ &= \sum_n\sum_m d_{m,n}\Re\big\{\langle\lambda_{m,n}, (t)\lambda_{r,k}(t)\rangle\big\} \\ &\quad + \sum_n\sum_m d_{m,n}\Im\big\{\langle\lambda_{m,n}(t), \lambda_{r,k}(t)\rangle\big\} \end{aligned} \tag{83}$$

Using (80) we get

$$y_{r,k} = d_{r,k} + j\left(\sum_n\sum_m d_{m,n}\Im\big\{\langle\lambda_{m,n}(t), \lambda_{r,k}(t)\rangle\big\}\right) \tag{84}$$

Since the interference in (84) is imaginary, the estimate of the transmitted data is $\hat{d}_{r,k} = \Re\{y_{r,k}\}$. Using tensor notation, this becomes $\hat{\mathcal{D}}[k] = \Re\{\mathcal{Y}(k\frac{T}{2})\}$.

Next, we consider the MIMO extension of FBMC. In each FBMC symbol, let there be P independent streams of data transmitted using $2M$ sub-carriers, N_T transmit and N_R receive antennas. There are P synthesis filter banks, one for each stream of data. Denote the filter for the mth sub-carrier for the pth synthesis filter bank by $p_{p,m}(t)$. A weight $w_{n_t,p}$ is assigned to the pth SFB output for the n_tth antenna. The weights $w_{n_t,p}$ are the coefficients of a linear precoder that couples the P streams of data with the N_T transmit antennas. Denote the data symbol for the pth stream, mth sub-carrier and nth FBMC symbol by $d_{p,m,n}$. As for the scalar case in (78), each data symbol $d_{p,m,n}$ is multiplied by a phase term $e^{j\psi_{m,n}}$. We get the transmit signal from the n_tth antenna as

$$x_{n_t}(t) = \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{2M-1} \sum_{p=0}^{P-1} w_{n_t,p} p_{p,m}\left(t - n\frac{T}{2}\right) e^{j\psi_{m,n}} d_{p,m,n} \tag{85}$$

Denoting the channel between the n_tth transmit and the n_rth receive antenna by $h_{n_r,n_t}(t)$, the received signal on the n_rth antenna is

$$r_{n_r}(t) = \sum_{n_t} h_{n_r,n_t}(t) * x_{n_t}(t) = \sum_{n_t} h_{n_r,n_t}(t) \times \left(\sum_{n=-\infty}^{+\infty} \sum_{m=0}^{2M-1} \sum_{p=0}^{P-1} w_{n_t,p} p_{p,m}\left(t - n\frac{T}{2}\right) e^{j\psi_{m,n}} d_{p,m,n} \right) \tag{86}$$

Let $c_{n_r,p,m}(t) = \sum_{n_t} h_{n_r,n_t}(t) * w_{n_t,p} p_{p,m}(t)$. We can then re-write (86) as

$$r_{n_r}(t) = \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{2M-1} \sum_{p=0}^{P-1} c_{n_r,p,m}\left(t - n\frac{T}{2}\right) e^{j\psi_{m,n}} d_{p,m,n} \tag{87}$$

There are P analysis filter banks (AFB) at the receiver, one corresponding to each transmit SFB, that filter the N_R received signals $r_{n_r}(t)$ and produce outputs $y_{p,m}(t)$. Denote the filter for the pth AFB, n_rth receive antenna and mth sub-carrier by $p_{R_{p,m,n_r}}(t)$. The output of the AFB is

$$\begin{aligned} y_{p,m}(t) &= \sum_{n_r} p_{R_{p,m,n_r}}(t) * r_{n_r}(t) \\ &= \sum_{n_r} \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{2M-1} \sum_{p=0}^{P-1} p_{R_{p,m,n_r}}(t) * \\ &\quad c_{n_r,p,m}\left(t - n\frac{T}{2}\right) e^{j\psi_{m,n}} d_{p,m,n} \end{aligned} \tag{88}$$

If the AFB is designed to be matched to the combined channel and transmit filter banks, then we have $p_{R_{p,m,n_r}}(t) = c^*_{n_r,p,m}(-t)$. Using the Tensor Framework, the data to be transmitted on the nth multi-carrier symbol is $\mathcal{D}[n] \in \mathbb{C}_n^{P\times 2M}$ with components

$$\mathcal{D}_{p,m}[n] = e^{\psi_{(m-1),n}} d_{(p-1),(m-1),n} \tag{89}$$

where $p = 1,\ldots,P; m = 1,\ldots,2M$. As discussed in Section III, the data tensor in the tensor framework may be constellation mapped data symbols or data symbols with some form of precoding. In this case, the data tensor is the latter due to the multiplication by the phase term. The transmit system tensor $\mathcal{H}_T(t) \in \mathbb{C}_t^{N_T\times P\times 2M}$ converts $\mathcal{D}[n]$ into a signal tensor $\mathcal{X}(t) \in \mathbb{C}_t^{N_T}$. The components of the transmit system tensor are

$$\mathcal{H}_{T_{n_t,p,m}}(t) = w_{n_t,(p-1)} p_{(p-1),(m-1)}(t) \tag{90}$$

where $p = 1,\ldots,P; m = 1,\ldots,2M; n_t = 1,\ldots,N_T$. We may now re-write (85) in tensor notation as $\mathcal{X}(t) = \{\mathcal{H}_T(t - n\frac{T}{2}), \mathcal{D}[n]\}_{(2)}$. The channel system tensor is $\mathcal{H}_C(t) \in \mathbb{C}_t^{N_R\times N_T}$ whose components

$$\mathcal{H}_{C_{n_r,n_t}}(t) = h_{n_r,n_t}(t) \quad \text{for } n_r = 1,\ldots,N_R; n_t = 1,\ldots,N_T \tag{91}$$

are the channel between the n_tth transmit and n_rth receive antenna. The output of the channel tensor $\mathcal{R}(t) \in \mathbb{C}_t^{N_R}$ is $\mathcal{R}(t) = \{\mathcal{H}_C(t) * \mathcal{X}(t)\}_{(1)}$ and the combined channel and transmit system tensor is $\mathcal{C}(t) = \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(1)} \in \mathbb{C}_t^{N_R\times P\times 2M}$ with components $\mathcal{C}_{n_r,p,m}(t) = \sum_{n_t} \mathcal{H}_{C_{n_r,n_t}}(t) * \mathcal{H}_{T_{n_t,p,m}}(t) = \sum_{n_t} h_{n_r,n_t}(t) * w_{n_t,(p-1)} p_{(p-1),(m-1)}(t) = c_{n_r,(p-1),(m-1)}(t)$. If a system matched to the combined channel and synthesis filter bank is used, then the receive system tensor $\mathcal{H}_R(t) \in \mathbb{C}_t^{P\times 2M\times N_R}$ is $\mathcal{H}_R(t) = \mathcal{C}^H(-t)$ and converts $\mathcal{R}(t)$ into a signal tensor $\mathcal{Y}(t) \in \mathbb{C}_t^{P\times 2M}$. The overall system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{P\times 2M\times P\times 2M}$ is

$$\begin{aligned} \mathcal{H}(t) &= \left\{\mathcal{H}_R(t) * \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(1)}\right\}_{(1)} \\ &= \left\{\mathcal{C}^H(-t) * \mathcal{C}(t)\right\}_{(1)} \end{aligned} \tag{92}$$

with components $\mathcal{H}_{p',m',p,m}(t) = \sum_{n_r=1}^{N_R} \mathcal{C}^*_{p',m',n_r}(-t) * \mathcal{C}_{n_r,p,m}(t) = c^*_{(p'-1),(m'-1),n_r}(-t) * c_{n_r,(p-1),(m-1)}(t)$. We may now re-write (88), for the case when $p_{R_{p,m,n_r}}(t) = c^*_{p,m,n_r}(-t)$, as

$$\begin{aligned} &y_{(p'-1),(m'-1)}(t) \\ &= \sum_{n=-\infty}^{+\infty} \sum_{n_r=1}^{N_R} \sum_{m=1}^{2M} \sum_{p=1}^{P} \Big[c^*_{(p'-1),(m'-1),n_r}(-t) \\ &\qquad \times c_{nr,(p-1),(m-1)}\left(t - n\frac{T}{2}\right) e^{j\psi_{(m-1),n}} d_{(p-1),(m-1),n} \Big] \\ &= \sum_{n=-\infty}^{+\infty} \sum_{m=1}^{2M} \sum_{p=1}^{P} \mathcal{H}_{p',m',p,m}\left(t - n\frac{T}{2}\right) \mathcal{D}_{p,m}[n] = \mathcal{Y}_{p',m'}(t) \end{aligned} \tag{93}$$

Writing $\mathcal{Y}_{p',m'}(t) = \sum_{n=-\infty}^{+\infty} \sum_{m=1}^{2M} \sum_{p=1}^{P} \mathcal{H}_{p',m',p,m}(t - n\frac{T}{2}) \mathcal{D}_{p,m}[n]$ in tensor notation gives

$$\mathcal{Y}(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t - nT_0), \mathcal{D}[n]\}_{(2)} \tag{94}$$

where $T_0 = \frac{T}{2}$. Sampling (94) at intervals kT_0 we get

$$\mathcal{Y}[k] = \mathcal{Y}(kT_0) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}((k-n)T_0), \mathcal{D}[n]\}_{(2)} \quad (95)$$

An estimate of the data tensor $\hat{\mathcal{D}}[k] \in \mathbb{C}_k^{P\times 2M}$ is found by feeding $\mathcal{Y}[k]$ into a tensor detector $\mathcal{G}[k] \in \mathbb{C}_k^{P\times 2M\times P\times 2M}$, some of which are derived in the next section, such that $\hat{\mathcal{D}}[k] = \{\mathcal{G}[k], \mathcal{Y}[k]\}_{(2)}$.

C. GENERALIZED FREQUENCY DIVISION MULTIPLEXING (GFDM)

Generalized frequency division multiplexing (GFDM) [48] is a block-based multicarrier modulation scheme that employs circular filtering. Consider a time-frequency resource block of duration T and bandwidth B. The available bandwidth is divided into K equally-spaced subcarriers with subcarrier spacing $\Delta_f = \frac{B}{K}$ [48], and the time slot is divided into M subsymbols with subsymbol spacing $T_{\text{sub}} = \frac{T}{M}$. The relation between the subcarrier spacing and the subsymbol spacing is given by $\Delta_f T_{\text{sub}} = 1$. The data symbol transmitted on the mth subsymbol, and kth sub-carrier is modulated by a pulse $p_{k,m}(t)$ given by

$$p_{k,m}(t) = w_T(t) p_T(t - mT_{\text{sub}}) e^{j2\pi\,\Delta_f kt} \quad (96)$$

where $m = 0, \ldots, M-1; k = 0, \ldots, K-1$ and $p_T(t)$ is a prototype pulse shape of period T, T_{sub} is the duration of one sub-symbol, T is the duration of the entire GFDM symbol, and $w_T(t)$ is a rectangular window of duration T such that

$$w_T(t) = \begin{cases} 1 & \text{for } t \in [0,\ T] \\ 0 & \text{elsewhere} \end{cases} \quad (97)$$

The rectangular window $w_T(t)$ is used to limit the final modulating pulse $p_{k,m}(t)$ to the interval $t \in [0,\ T]$. A GFDM block hence comprises of pulse shapes generated by time and frequency shifts of a prototype pulse shape followed by multiplication by a finite time window. The transmit signal is given by

$$x(t) = \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} p_{k,m}(t - nT) d_{k,m,n} \quad (98)$$

where $d_{k,m,n}$ is the complex data transmitted during the mth subsymbol on the kth subcarrier and the nth GFDM symbol. Denoting the channel by $c(t)$ and the additive white Gaussian noise (AWGN) process by $v(t)$, the received signal is

$$\begin{aligned} r(t) &= c(t) * x(t) + v(t) \\ &= \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \big(c(t) * p_{k,m}(t - nT)\big) d_{k,m,n} + v(t) \end{aligned} \quad (99)$$

Defining second order tensors $\mathcal{D}[n] \in \mathbb{C}_n^{K\times M}$ and $\mathcal{H}(t) \in \mathbb{C}_t^{K\times M}$ with components $\mathcal{D}_{k,m}[n] = d_{k,m,n}$ and $\mathcal{H}_{k,m}(t) = c(t) * p_{k,m}(t)$, we re-write (99) using the tensor representation as

$$r(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t - nT), \mathcal{D}\}_{(2)} + v(t) \quad (100)$$

The signal $r(t)$ is the input to a system tensor $\mathcal{H}_R(t) \in \mathbb{C}_t^{K\times M\times 1}$ where the singleton dimension is used to indicate that the input to this system is a scalar function ($r(t)$). If there is a bank of filters matched to the transmit filters at the receiver then we have $\mathcal{H}_{R_{k,m}}(t) = p_{k,m}^*(-t)$ whose output $\mathcal{Y} \in \mathbb{C}^{K\times M}$ has components $\mathcal{Y}_{k,m}(t) = p_{k,m}^*(-t) * r(t)$.

Extending this to the MIMO case, let P independent streams of data be transmitted using K sub-carriers, M sub-symbols and N_T transmit antennas. Let there be N_R receive antennas. Assuming that different banks of filters are used at each transmit and receive antenna, we get the signal transmitted by the n_tth antenna as

$$x_{n_t}(t) = \sum_{n=-\infty}^{+\infty} \sum_{p=1}^{P} \sum_{k=1}^{K} \sum_{m=1}^{M} w_{n_t,p} d_{n,p,k,m} p_{T_{n_t,k,m}}(t - nT) \quad (101)$$

where $d_{n,p,k,m}$ is the data transmitted on the nth GFDM symbol, during the mth sub-symbol, on the kth sub-carrier and on the pth stream. Using the Tensor Framework, the complex data to be transmitted on the nth GFDM symbol is $\mathcal{D}[n] \in \mathbb{C}_n^{P\times K\times M}$ with components $\mathcal{D}_{p,k,m}[n] = d_{n,p,k,m}$. The transmit tensor $\mathcal{H}_T(t) \in \mathbb{C}_t^{N_T\times P\times K\times M}$ whose (n_t, p, k, m)th component is $\mathcal{H}_{T_{n_t,p,k,m}} = w_{n_t,p} p_{T_{n_t,k,m}}(t)$ converts the data tensor into a signal tensor $\mathcal{X}(t) \in \mathbb{C}_t^{N_T}$. We write (98) using tensor notation as $\mathcal{X}(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}_T(t - nT), \mathcal{D}[n]\}_{(3)}$. The channel system tensor is $\mathcal{H}_C(t) \in \mathbb{C}_t^{N_R\times N_T}$ whose components $\mathcal{H}_{C_{n_r,n_t}}(t)$ are the channel between the n_tth transmit and n_rth receive antenna. The output of the channel $\mathcal{R}(t) \in \mathbb{C}_t^{N_R}$ is $\mathcal{R}(t) = \{\mathcal{H}_C(t) * \mathcal{X}(t)\}_{(1)}$ and the combined channel and transmit tensor $\mathcal{C}(t) \in \mathbb{C}_t^{N_R\times P\times K\times M}$ is $\mathcal{C}(t) = \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(1)}$. If a matched filter is used at the receiver, the receive system tensor $\mathcal{H}_R(t) = \mathcal{C}^H(-t) \in \mathbb{C}_t^{P\times K\times M\times N_R}$ converts $\mathcal{R}(t)$ into a signal tensor $\mathcal{Y}(t) \in \mathbb{C}_t^{P\times F\times K}$. The overall system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{P\times K\times M\times P\times K\times M}$ is

$$\begin{aligned} \mathcal{H}(t) &= \Big\{\mathcal{H}_R(t) * \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(1)}\Big\}_{(1)} \\ &= \Big\{\mathcal{C}^H(-t) * \mathcal{C}(t)\Big\}_{(1)} \end{aligned} \quad (102)$$

and

$$\mathcal{Y}(t) = \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t - nT), \mathcal{D}[n]\}_{(3)} \quad (103)$$

Sampling (103) at intervals of T gives $\mathcal{Y}[k] = \mathcal{Y}(kT) = \{\mathcal{H}(kT - nT), \mathcal{D}[n]\}_{(3)}$. An estimate of the data tensor $\hat{\mathcal{D}}[k] \in \mathbb{C}_k^{P\times K\times M}$ is found by feeding $\mathcal{Y}[k]$ into a tensor detector $\mathcal{G}[k] \in \mathbb{C}_k^{P\times K\times M\times P\times K\times M}$, some of which are derived in the next section, such that $\hat{\mathcal{D}}[k] = \{\mathcal{G}[k], \mathcal{Y}[k]\}_{(3)}$.

V. MULTI-DOMAIN EQUALIZATION

A. EQUIVALENT DISCRETE TIME MODEL

Consider the system in Figure 3. The data transmitted on the nth tensor symbol is $\mathcal{D}[n] \in \mathbb{C}_n^{I_1\times\cdots\times I_N}$. The transmit, channel and receive system tensors are $\mathcal{H}_T(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, $\mathcal{H}_C(t) \in \mathbb{C}_t^{K_1\times\cdots\times K_O\times J_1\times\cdots\times J_M}$, $\mathcal{H}_R(t) \in \mathbb{C}_t^{L_1\times\cdots\times L_P\times K_1\times\cdots\times K_O}$ respectively, $\mathcal{N}(t) \in \mathbb{C}_t^{K_1\times\cdots\times K_O}$ is an additive noise tensor, $\mathcal{X}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M}$ is the output of the transmit system tensor and $\mathcal{Y}(t) \in \mathbb{C}_t^{L_1\times\cdots\times L_P}$ is the output of the receive system tensor. Furthermore, $\mathcal{G}[n] \in \mathbb{C}_n^{I_1\times\cdots\times I_N\times L_1\times\cdots\times L_P}$ is a discrete system tensor whose output $\hat{\mathcal{D}}[n]$ is the estimate of the data $\mathcal{D}[n]$. The input to the receive system tensor is

$$\begin{aligned}\mathcal{R}(t) &= \left\{\{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(M)} * \sum_n \mathcal{D}[n]\delta(t-nT)\right\}_{(N)} \\ &\quad + \mathcal{N}(t) \\ &= \{\mathcal{C}(t-nT), \mathcal{D}[n]\}_{(N)} + \mathcal{N}(t) \end{aligned} \tag{104}$$

where the cascade of the transmit tensor and the channel tensor is represented by $\mathcal{C}(t) = \{\mathcal{H}_C(t) * \mathcal{H}_T(t)\}_{(M)} \in \mathbb{C}_t^{K_1\times\cdots\times K_O\times I_1\times\cdots\times I_N}$. The output of the receive system tensor is

$$\begin{aligned}\mathcal{Y}(t) &= \{\mathcal{H}_R(t) * \mathcal{R}(t)\}_{(O)} + \{\mathcal{H}_R(t) * \mathcal{N}(t)\}_{(O)} \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(t-nT), \mathcal{D}[n]\}_{(N)} + \mathcal{V}(t) \end{aligned} \tag{105}$$

where $\mathcal{H}(t) \in \mathbb{C}_t^{L_1\times\cdots\times L_P\times I_1\times\cdots\times I_N}$ is the overall system tensor comprising the transmit, channel and receive system tensors, and $\mathcal{V}(t) = \{\mathcal{H}_R(t) * \mathcal{N}(t)\}_{(O)} \in \mathbb{C}_t^{L_1\times\cdots\times L_P}$. Sampling the output $\mathcal{Y}(t)$ at a rate of $\frac{1}{T}$, we get

$$\begin{aligned}\mathcal{Y}[k] = \mathcal{Y}(kT) &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}(kT-nT), \mathcal{D}[n]\}_{(N)} + \mathcal{V}[kT] \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{H}[k-n], \mathcal{D}[n]\}_{(N)} + \mathcal{V}[k] \end{aligned} \tag{106}$$

where $\mathcal{H}[k] = \mathcal{H}(kT)$ is the kth sample of $\mathcal{H}(t)$ and $\mathcal{V}[k] = \{\mathcal{H}_R(t) * \mathcal{N}(t)\}_{(O)}|_{(t=kT)}$ is the sampled noise with autocorrelation

$$\mathcal{R}_{\mathcal{V}}[i] = \mathbb{E}\big[\mathcal{V}[k] \circ \mathcal{V}^*[k-i]\big] \tag{107}$$

The output of $\mathcal{G}[n]$, which is the estimate of the data tensor $\mathcal{D}[k]$, is

$$\begin{aligned}\hat{\mathcal{D}}[k] &= \sum_m \{\mathcal{G}[m], \mathcal{Y}[k-m]\}_{(P)} \\ &= \sum_m \sum_n \big\{\mathcal{G}[m], \{\mathcal{H}[n], \mathcal{D}[k-m-n]\}_{(N)}\big\}_{(P)} \\ &\quad + \sum_m \{\mathcal{G}[m], \mathcal{V}[k-m]\}_{(P)} \\ &= \sum_m \sum_n \big\{\mathcal{G}[m], \{\mathcal{H}[n], \mathcal{D}[k-m-n]\}_{(N)}\big\}_{(P)} + \tilde{\mathcal{V}}[k] \end{aligned} \tag{108}$$

resulting in the equivalent discrete time system model.

Theorem 1: Consider an input $\mathcal{X}(t) = \sum_{n=-\infty}^{+\infty} \mathcal{X}[n]\delta(t-nT) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$ to a system tensor $\mathcal{A}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$. The output of this filter, corrupted by additive white Gaussian noise, is $\mathcal{R}(t) = \{\mathcal{A}(t) * \mathcal{X}(t)\}_{(N)} + \mathcal{N}(t)$ where $\mathcal{N}(t)$ is a tensor whose components are white Gaussian noise processes. Let $\mathcal{R}(t)$ be the input to a system tensor $\mathcal{B}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ with an output $\mathcal{Y}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$. The per component SNR of the samples $\mathcal{Y}(kT) = \mathcal{Y}[k]$ is maximized when $\mathcal{B}(t) = \mathcal{A}^H(-t)$.

Proof: The tensor $\mathcal{Y}(t)$ can be written as

$$\begin{aligned}\mathcal{Y}(t) &= \big\{\mathcal{B}(t) * \{\mathcal{A}(t) * \mathcal{X}(t)\}_{(N)}\big\}_{(M)} + \{\mathcal{B}(t) * \mathcal{N}(t)\}_{(M)} \\ &= \sum_{n=-\infty}^{+\infty} \{\mathcal{C}(t-nT), \mathcal{X}[n]\}_{(N)} + \mathcal{V}(t) \end{aligned} \tag{109}$$

where $\mathcal{C}(t) = \{\mathcal{B}(t) * \mathcal{A}(t)\}_{(M)}$ and $\mathcal{V}(t) = \{\mathcal{B}(t) * \mathcal{N}(t)\}_{(M)}$. Sampling at intervals of kT gives

$$\mathcal{Y}[k] = \sum_{n=-\infty}^{+\infty} \{\mathcal{C}[k-n], \mathcal{X}[n]\}_{(N)} + \mathcal{V}[k] \tag{110}$$

with components

$$\mathcal{Y}_{i_1\ldots i_N}[k] = \mathcal{C}_{i_1\ldots i_N i_1\ldots i_N}[0]\mathcal{X}_{i_1\ldots i_N}[k] + \mathcal{I}_{i_1\ldots i_N}[k] + \mathcal{V}_{i_1\ldots i_N}[k] \tag{111}$$

where

$$\begin{aligned}\mathcal{I}_{i_1\ldots i_N}[k] &= \sum_{i'_1\ldots i'_N, i'_q\neq i_q, q=1,\ldots,N} \mathcal{C}_{i_1\ldots i_N i'_1\ldots i'_N}[0]\mathcal{X}_{i'_1\ldots i'_N}[k] \\ &\quad + \sum_{\substack{n=-\infty \\ n\neq k}}^{+\infty} \sum_{i'_1\ldots i'_N} \mathcal{C}_{i_1\ldots i_N i'_1\ldots i'_N}[k-n]\mathcal{X}_{i'_1\ldots i'_N}[n] \end{aligned} \tag{112}$$

We define the per component SNR $\gamma_{i_1,\ldots,i_N}$ at the sampled output of the system $\mathcal{B}(t)$ as the ratio of the power of the desired symbol $\mathcal{X}_{i_1,\ldots,i_N}[k]$ and the power of the noise $\mathcal{V}_{i_1,\ldots,i_N}[k]$. The intra-tensor and inter-tensor interference is contained in $\mathcal{I}_{i_1,\ldots,i_N}[k]$ and is not considered in this definition of the SNR. We have

$$\gamma_{i_1,\ldots,i_N} = \frac{\big|\mathcal{C}_{i_1,\ldots,i_N,i_1,\ldots,i_N}[0]\big|^2 E_{i_1,\ldots,i_N}}{\mathbb{E}\big[\mathcal{V}_{i_1,\ldots,i_N}[k]\mathcal{V}^*_{i_1,\ldots,i_N}[k]\big]} \tag{113}$$

where $E_{i_1,\ldots,i_N} = \mathbb{E}\,[\mathcal{X}_{i_1,\ldots,i_N}\mathcal{X}^*_{i_1,\ldots,i_N}]$. We may expand $|\mathcal{C}_{i_1,\ldots,i_N,i_1,\ldots,i_N}[0]|^2$ as

$$\begin{aligned}&\big|\mathcal{C}_{i_1,\ldots,i_N,i_1,\ldots,i_N}[0]\big|^2 \\ &= \left|\sum_{j_1,\ldots,j_M} \int_{-\infty}^{+\infty} \mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t)\mathcal{A}_{j_1,\ldots,j_M,i_1,\ldots,i_N}(-t)dt\right|^2 \end{aligned} \tag{114}$$

The denominator of (113) can be expanded, using $\mathbb{E}\left[N_{j_1,\ldots,j_M}(t)N_{j'_1,\ldots,j'_M}(p)\right] = N_0\delta(t-p)$ when $(j_1,\ldots,j_M) = (j'_1,\ldots,j'_M)$ and 0 otherwise, as

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{V}_{i_1,\ldots,i_N}[k]\mathcal{V}^*_{i_1,\ldots,i_N}[k]\right] \\
&= \mathbb{E}\Bigg[\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\mathcal{B}_{i_1\ldots i_N j_1\ldots j_M}(t)\mathcal{N}_{j_1\ldots j_M}(kT-t)dt \\
&\qquad\times\sum_{j'_1\ldots j'_M}\int_{-\infty}^{+\infty}\mathcal{B}^*_{i_1\ldots i_N,j'_1\ldots j'_M}(p)\mathcal{N}^*_{j'_1\ldots j'_M}(kT-p)dp\Bigg] \\
&= N_0\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathcal{B}_{i_1\ldots i_N j_1\ldots j_M}(t)\mathcal{B}^*_{i_1\ldots i_N j_1\ldots j_M}(p)\delta(t-p)dtdp \\
&= N_0\sum_{j_1,\ldots,j_M}\int_{-\infty}^{+\infty}\left|\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t)\right|^2dt
\end{aligned} \tag{115}$$

Using (114) and (115) in (113) we get

$$\begin{aligned}
&\gamma_{i_1,\ldots,i_N} \\
&= \frac{\left|\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\mathcal{B}_{i_1\ldots i_N j_1\ldots j_M}(t)\mathcal{A}_{j_1\ldots j_M i_1\ldots i_N}(-t)dt\right|^2 E_{i_1\ldots i_N}}{N_0\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\left|\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t)\right|^2dt}
\end{aligned} \tag{116}$$

Using the Cauchy-Schwartz inequality from Appendix A this becomes

$$\begin{aligned}
&\gamma_{i_1,\ldots,i_N} \\
&\le \frac{\left(\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\left|\mathcal{A}_{i_1\ldots i_N j_1\ldots j_M}(t)\right|^2dt\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\left|\mathcal{B}_{j_1\ldots j_M k_1\ldots k_P}(t)\right|^2dt\right)E_{i_1\ldots i_N}}{N_0\sum_{j_1,\ldots,j_M}\int_{-\infty}^{+\infty}\left|\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t)\right|^2dt} \\
&= \frac{\sum_{j_1\ldots j_M}\int_{-\infty}^{+\infty}\left|\mathcal{A}_{i_1\ldots i_N j_1\ldots j_M}(t)\right|^2dtE_{i_1,\ldots,i_N}}{N_0}
\end{aligned} \tag{117}$$

with equality when $\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t) = \mathcal{A}^*_{j_1,\ldots,j_M,i_1,\ldots,i_N}(-t)$. This means that the SNR attains its maximum value when $\mathcal{B}(t) = \mathcal{A}^H(-t)$. ■

If the receiver employs a tensor matched filter, that is matched to the combined channel and transmit tensors, we have $\mathcal{H}_R(t) = \mathcal{C}^H(-t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N\times K_1\times\cdots\times K_O}$ and the overall system tensor is $\mathcal{H}(t) = \{\mathcal{C}^H(-t) * \mathcal{C}(t)\}_{(O)} \in \mathbb{C}_t^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$.

B. LINEAR EQUALIZATION

We now consider linear equalization schemes where the equalizer has a finite number (M) of tensor taps $\{\mathcal{G}[i] \in \mathbb{C}_i^{I_1\times\cdots\times I_N\times L_1\times\cdots\times L_P}\}$, $i = 0,1,\ldots,M-1$. Assume that the overall channel contains $v+1$ tensor taps $\{\mathcal{H}[i] \in \mathbb{C}_i^{L_1\times\cdots\times L_P\times I_1\times\cdots\times I_N}\}$, $i = 0,1,\ldots,v$ such that the estimate of the data tensor is given by

$$\hat{\mathcal{D}}[k] = \sum_{i=0}^{M-1}\{\mathcal{G}[i],\mathcal{Y}[k-i]\}_{(P)} \tag{118}$$

There is a decision delay Δ, such that $0 \le \Delta \le M+v-1$, to ensure causality. This delay is important when designing finite-length equalizers as non-causal filters cannot be implemented in practice. Hence, the tensor $\hat{\mathcal{D}}[k]$ is an estimate of $\mathcal{D}[k-\Delta]$. Collecting the received tensor for different delays $\mathcal{Y}[k], \mathcal{Y}[k-1],\ldots,\mathcal{Y}[k-(M-1)]$ into an extended tensor $\bar{\mathcal{Y}}[k] \in \mathbb{C}_k^{M\times L_1\times\cdots\times L_P}$, where the additional domain is the delay domain, we have

$$\bar{\mathcal{Y}}_{m,l_1,\ldots,l_P}[k] = \mathcal{Y}_{l_1,\ldots,l_P}[k-(m-1)] \quad \text{for } m = 1,\ldots,M \tag{119}$$

Similarly, define extended noise tensor $\bar{\mathcal{V}}[k] \in \mathbb{C}_k^{M\times L_1\times\cdots\times L_P}$ and extended data tensor $\bar{\mathcal{D}}[k] \in \mathbb{C}_k^{(M+v)\times I_1\times\cdots\times I_N}$ such that

$$\begin{aligned}
\bar{\mathcal{D}}_{q,i_1,\ldots,i_N}[k] &= \mathcal{D}_{i_1,\ldots,i_N}\left[k-(q-1)\right] \quad \text{for } q = 1,\ldots,M+v \\
\bar{\mathcal{V}}_{m,l_1,\ldots,l_P}[k] &= \mathcal{V}_{l_1,\ldots,l_P}[k-(m-1)] \quad \text{for } m = 1,\ldots,M
\end{aligned} \tag{120}$$

The *slice* of a tensor is defined as a two-dimensional section of a tensor obtained by fixing all but two indices [8]. For example, a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times I_3}$ has three slices denoted by $\mathcal{A}_{i_1,:,:}$,$\mathcal{A}_{:,i_2,:}$ and $\mathcal{A}_{:,:,i_3}$. Define $\bar{\mathcal{H}} \in \mathbb{C}_k^{M\times L_1\times\cdots\times L_P\times(M+v)\times I_1\times\cdots\times I_N}$ with two additional domains corresponding to the delays at the receiver and the transmitter such that the slice $\bar{\mathcal{H}}_{:,l_1,\ldots,l_P,:,i_1,\ldots,i_N}$ has components as shown in (121) at the bottom of the next page.

For a channel $\mathcal{H}[k]$ with $v+1$ non-zero taps, (106) becomes $\mathcal{Y}[k] = \sum_{n=0}^{v}\{\mathcal{H}[k],\mathcal{D}[k-n]\}_{(N)} + \mathcal{V}[k]$ with components

$$\mathcal{Y}_{l_1\ldots l_P}[k] = \sum_{n=0}^{v}\sum_{i_1\ldots i_N}\mathcal{H}_{l_1\ldots l_P i_1\ldots I_N}[n]\mathcal{D}_{i_1\ldots i_N}[k-n] + \mathcal{V}_{l_1\ldots l_P}[k] \tag{122}$$

Using (119) and (122) we have

$$\begin{aligned}
&\bar{\mathcal{Y}}_{m,l_1,\ldots,l_P}[k] \\
&= \mathcal{Y}_{l_1,\ldots,l_P}[k-(m-1)] \\
&= \sum_{n=0}^{v}\sum_{i_1\ldots i_N}\mathcal{H}_{l_1\ldots l_P i_1\ldots I_N}[n]\mathcal{D}_{i_1\ldots i_N}[k-(m-1)-n] \\
&\quad + \mathcal{V}_{l_1\ldots l_P}[k-(m-1)]
\end{aligned}$$

and with (121) this becomes

$$\begin{aligned}
&\bar{\mathcal{Y}}_{m,l_1\ldots l_P}[k] \\
&= \sum_{q=1}^{M+v}\sum_{i_1\ldots i_N}\bar{\mathcal{H}}_{m,l_1\ldots l_P,q,i_1\ldots I_N}\bar{\mathcal{D}}_{q,i_1\ldots i_N}[k] + \bar{\mathcal{V}}_{m,l_1\ldots l_P}[k]
\end{aligned} \tag{123}$$

which in tensor notation gives

$$\bar{\mathcal{Y}}[k] = \left\{\bar{\mathcal{H}},\bar{\mathcal{D}}[k]\right\}_{(N+1)} + \bar{\mathcal{V}}[k] \tag{124}$$

Further, define an augmented tensor $\bar{\mathcal{G}} \in \mathbb{C}^{I_1\times\cdots\times I_N\times M\times L_1\times\cdots\times L_P}$, which is a collection of the tensor equalizer taps $\mathcal{G}[0],\ldots,\mathcal{G}[M-1]$, whose components are

$$\bar{\mathcal{G}}_{i_1\ldots i_N m l_1\ldots l_P} = \mathcal{G}_{i_1\ldots i_N l_1\ldots l_P}[(m-1)] \quad \text{for } m = 1\ldots M \tag{125}$$

The components of $\hat{\mathcal{D}}[k]$ from (118) can be written as

$$\begin{aligned}\hat{\mathcal{D}}_{i_1\ldots i_N}[k] &= \sum_{m=0}^{M-1}\left(\sum_{l_1\ldots l_P}\mathcal{G}_{i_1\ldots i_N l_1\ldots l_P}[m]\mathcal{Y}_{l_1\ldots l_P}[k-m]\right)\\ &= \sum_{l_1\ldots l_P}\mathcal{G}_{i_1,\ldots,i_N,l_1,\ldots,l_P}[0]\mathcal{Y}_{l_1,\ldots,l_P}[k]+\cdots\\ &\quad+\sum_{l_1\ldots l_P}\mathcal{G}_{i_1,\ldots,i_N,l_1,\ldots,l_P}[M-1]\mathcal{Y}_{l_1,\ldots,l_P}[k-M+1]\\ &= \sum_{l_1\ldots l_P}\bar{\mathcal{G}}_{i_1,\ldots,i_N,1,l_1,\ldots,l_P}\bar{\mathcal{Y}}_{1,l_1,\ldots,l_P}[k]\\ &\quad+\cdots+\sum_{l_1\ldots l_P}\bar{\mathcal{G}}_{i_1,\ldots,i_N,M,l_1,\ldots,l_P}\bar{\mathcal{Y}}_{M,l_1,\ldots,l_P}[k]\\ &= \left(\left\{\bar{\mathcal{G}},\bar{\mathcal{Y}}[k]\right\}_{(P+1)}\right)_{i_1,\ldots,i_N}\end{aligned}\tag{126}$$

and (118) becomes

$$\hat{\mathcal{D}}[k]=\left\{\bar{\mathcal{G}},\bar{\mathcal{Y}}[k]\right\}_{(P+1)}\tag{127}$$

The choice of the equalizer taps depends on the criterion used. A *zero forcing* equalizer aims to remove the interference introduced by the channel. Define an augmented identity tensor $\mathcal{P}\in\mathbb{C}^{I_1\times\cdots\times I_N\times(M+v)\times I_1\times\cdots\times I_N}$ such that

$$\mathcal{P}_{:,\ldots,:,m,:,\ldots,:}=\begin{cases}\mathbf{0}_T & \text{if } m\neq\Delta+1\\ \mathcal{I}_N & \text{if } m=\Delta+1\end{cases}\tag{128}$$

where $0_T\in\mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is the all zero tensor and $\mathcal{I}_N\in\mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is the identity tensor. Notice that $\{\mathcal{P},\bar{\mathcal{D}}[k]\}_{(N+1)}=\mathcal{D}[k-\Delta]$. Substituting for $\bar{\mathcal{Y}}[k]$ from (124) in (127) gives

$$\hat{\mathcal{D}}[k]=\left\{\left\{\bar{\mathcal{G}},\bar{\mathcal{H}}\right\}_{(P+1)},\bar{\mathcal{D}}[k]\right\}_{(N+1)}+\left\{\bar{\mathcal{G}},\bar{\mathcal{V}}[k]\right\}_{(P+1)}\tag{129}$$

This means that if

$$\left\{\bar{\mathcal{G}},\bar{\mathcal{H}}[k]\right\}_{(P+1)}=\mathcal{P}\tag{130}$$

then the effects of the interference are completely removed, and in the absence of noise, $\hat{\mathcal{D}}[k]=\mathcal{D}[k-\Delta]$. The solution to (130) gives

$$\bar{\mathcal{G}}_{\mathrm{ZF}}=\left\{\mathcal{P},\bar{\mathcal{H}}^{+}[k]\right\}_{(N+1)}\tag{131}$$

For an equalizer that minimizes the error between the data and the estimate, we define an error tensor

$$\mathcal{E}[k]=\hat{\mathcal{D}}[k]-\mathcal{D}[k-\Delta]=\left\{\bar{\mathcal{G}},\bar{\mathcal{Y}}[k]\right\}_{(P+1)}-\mathcal{D}[k-\Delta]\tag{132}$$

The mean squared error is defined as

$$\begin{aligned}\mathrm{MSE}&=\sum_{i_1,\ldots,i_N}\mathbb{E}\left[\left|\mathcal{E}_{i_1,\ldots,i_N}[k]\right|^2\right]\\ &=\sum_{i_1,\ldots,i_N}\mathbb{E}\left[\left|\mathcal{D}_{i_1,\ldots,i_N}[k]-\hat{\mathcal{D}}_{i_1,\ldots,i_N}[k]\right|^2\right]\qquad(133)\\ &=\mathbb{E}\left[\left\|\mathcal{D}[k]-\hat{\mathcal{D}}[k]\right\|_F^2\right].\qquad(134)\end{aligned}$$

Theorem 2: The mean squared error between a tensor $\mathcal{D}[k]\in\mathbb{C}_k^{I_1\times\cdots\times I_N}$ and its estimate $\hat{\mathcal{D}}[k]\in\mathbb{C}_k^{I_1\times\cdots\times I_N}$ is minimized if and only if the error is uncorrelated with all the observed tensors $\mathcal{Y}[k]\in\mathbb{C}_k^{L_1\times\cdots\times L_M}$

$$\mathbb{E}\left[\mathcal{E}[k]\circ\mathcal{Y}^*[k-i]\right]=\mathbf{0}_T\quad\forall i\tag{135}$$

where $\mathcal{E}[k]=\mathcal{D}[k]-\hat{\mathcal{D}}[k]$.

The proof of this theorem can be found in Appendix B. Using Theorem 2, we have that the optimal multi-linear equalizer must satisfy

$$\mathcal{R}_{\mathcal{E},\mathcal{Y}}[i]=\mathbb{E}\left[\left(\hat{\mathcal{D}}[k]-\mathcal{D}[k-\Delta]\right)\circ\mathcal{Y}^*[k-i]\right]=\mathbf{0}_T\quad\text{for } |i|\leq M\tag{136}$$

where $\mathbf{0}_T\in\mathbb{C}^{I_1\times\cdots\times I_N\times L_1\times\cdots\times L_P}$, which is equivalent to

$$\mathbb{E}\left[\left(\hat{\mathcal{D}}[k]-\mathcal{D}[k-\Delta]\right)\circ\bar{\mathcal{Y}}^*[k]\right]=\mathbf{0}_T\tag{137}$$

$$\implies\mathbb{E}\left[\hat{\mathcal{D}}[k]\circ\bar{\mathcal{Y}}^*[k]\right]=\mathbb{E}\left[\mathcal{D}[k-\Delta])\circ\bar{\mathcal{Y}}^*[k]\right.\tag{138}$$

where $\mathbf{0}_T\in\mathbb{C}^{I_1\times\cdots\times I_N\times M\times L_1\times\cdots\times L_P}$. Substituting the value of $\hat{\mathcal{D}}[k]$ from (127), we get

$$\mathbb{E}\left[\left\{\bar{\mathcal{G}},\bar{\mathcal{Y}}[k]\right\}_{(P+1)}\circ\bar{\mathcal{Y}}^*[k]\right]=\mathbb{E}\left[\mathcal{D}[k-\Delta]\circ\bar{\mathcal{Y}}^*[k]\right]\tag{139}$$

Denoting $\mathcal{R}_{\bar{\mathcal{Y}}}[0]=\mathbb{E}\left[\bar{\mathcal{Y}}[k]\circ\bar{\mathcal{Y}}^*[k]\right]=\mathcal{R}_{\bar{\mathcal{Y}}}$, we have

$$\begin{aligned}\mathcal{R}_{\bar{\mathcal{Y}}}&=\mathbb{E}\left[\bar{\mathcal{Y}}[k]\circ\bar{\mathcal{Y}}^*[k]\right]\\ &=\mathbb{E}\left[\left(\left\{\bar{\mathcal{H}},\bar{\mathcal{D}}[k]\right\}_{(N+1)}+\bar{\mathcal{V}}[k]\right)\circ\left(\left\{\bar{\mathcal{H}},\bar{\mathcal{D}}[k]\right\}_{(N+1)}+\bar{\mathcal{V}}[k]\right)^*\right]\\ &=\mathbb{E}\left[\left\{\bar{\mathcal{H}},\bar{\mathcal{D}}[k]\right\}_{(N+1)}\circ\left\{\bar{\mathcal{H}}^*,\bar{\mathcal{D}}^*[k]\right\}_{(N+1)}\right]\\ &\quad+\mathbb{E}\left[\left\{\bar{\mathcal{H}},\bar{\mathcal{D}}[k]\right\}_{(N+1)}\circ\bar{\mathcal{V}}^*[k]\right]\\ &\quad+\mathbb{E}\left[\bar{\mathcal{V}}[k]\circ\left\{\bar{\mathcal{H}}^*,\bar{\mathcal{D}}^*[k]\right\}_{(N+1)}\right]+\mathbb{E}\left[\bar{\mathcal{V}}[k]\circ\bar{\mathcal{V}}^*[k]\right]\\ &=\left\{\left\{\bar{\mathcal{H}},\mathcal{R}_{\bar{\mathcal{D}}}\right\}_{(N+1)},\bar{\mathcal{H}}^H\right\}_{(N+1)}+\mathcal{R}_{\bar{\mathcal{V}}}\end{aligned}\tag{140}$$

$$\bar{\mathcal{H}}_{:,l_1,\ldots,l_P,:,i_1,\ldots,i_N}=\begin{bmatrix}\mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[0] & \cdots & \mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[v] & 0 & \cdots & \cdots & 0\\ 0 & \mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[0] & \cdots & \mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[v] & 0 & \cdots & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & \cdots & 0 & \mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[0] & \cdots & \mathcal{H}_{l_1\ldots l_P i_1\ldots i_N}[v]\end{bmatrix}\tag{121}$$

where $\mathcal{R}_{\bar{\mathcal{V}}} = \mathbb{E}[\bar{\mathcal{V}}[k] \circ \bar{\mathcal{V}}^*[k]]$ and $\mathcal{R}_{\bar{\mathcal{D}}} = \mathbb{E}[\bar{\mathcal{D}}[k] \circ \bar{\mathcal{D}}^*[k]]$. Let $\mathbb{E}[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{Y}}^*[k]] = \mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}$. Using (124) gives

$$\begin{aligned}\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} &= \mathbb{E}\Big[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{Y}}^*[k]\Big] \\ &= \mathbb{E}\Big[\mathcal{D}[k-\Delta] \circ \Big(\{\bar{\mathcal{H}}, \bar{\mathcal{D}}[k]\}_{(N+1)} + \bar{\mathcal{V}}[k]\Big)^*\Big] \\ &= \Big\{\mathbb{E}\big[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{D}}[k]\big], \bar{\mathcal{H}}^H\Big\}_{(N+1)} \\ &= \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}}, \bar{\mathcal{H}}^H\Big\}_{(N+1)} \end{aligned} \tag{141}$$

where $\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}} = \mathbb{E}[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{D}}[k]]$ and has components

$$\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}_{i_1 \ldots i_N m i'_1 \ldots i'_N}} = \mathbb{E}\Big[\mathcal{D}_{i_1 \ldots i_N}[k-\Delta]\mathcal{D}^*_{i'_1 \ldots i'_N}[k-m]\Big]$$

Assuming the data tensors are uncorrelated, we get

$$\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}_{i_1 \ldots i_N m i'_1 \ldots i'_N}} = \begin{cases} \mathbb{E}\big[\mathcal{D}_{i_1 \ldots i_N}[k-m] \circ \mathcal{D}^*_{i_1 \ldots i_N}[k-m]\big] & \text{if } m = \Delta \\ & \text{and } i_1 = i'_1 \ldots i_N = i'_N \\ 0 & \text{otherwise} \end{cases} \tag{142}$$

Further, we have

$$\begin{aligned}\mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}} &= \mathbb{E}\big[\bar{\mathcal{Y}}[k] \circ \mathcal{D}^*[k-\Delta]\big] \\ &= \mathbb{E}\Big[\Big(\{\bar{\mathcal{H}}, \bar{\mathcal{D}}[k]\}_{(N+1)} + \bar{\mathcal{V}}[k]\Big) \circ \mathcal{D}^*[k-\Delta]\Big] \\ &= \big\{\bar{\mathcal{H}}, \mathbb{E}\big[\bar{\mathcal{D}}[k] \circ \mathcal{D}^*[k-\Delta]\big]\big\}_{(N+1)} \\ &= \Big\{\bar{\mathcal{H}}, \Big(\mathbb{E}\Big[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{D}}^*[k]\Big]\Big)^H\Big\}_{(N+1)} \\ &= \Big(\Big\{\mathbb{E}\big[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{D}}[k]\big], \bar{\mathcal{H}}^H\Big\}_{(N+1)}\Big)^H \\ &= \mathcal{R}^H_{\mathcal{D},\bar{\mathcal{Y}}} \end{aligned} \tag{143}$$

The LHS of (139) becomes

$$\begin{aligned}\mathbb{E}\Big[\{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} \circ \bar{\mathcal{Y}}^*[k]\Big] &= \mathbb{E}\Big[\Big\{\bar{\mathcal{G}}, \Big(\bar{\mathcal{Y}}[k] \circ \bar{\mathcal{Y}}^*[k]\Big)\Big\}_{(P+1)}\Big] \\ &= \Big\{\bar{\mathcal{G}}, \Big(\mathbb{E}\Big[\bar{\mathcal{Y}}[k] \circ \bar{\mathcal{Y}}^*[k]\Big]\Big)\Big\}_{(P+1)} \\ &= \{\bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}}}\}_{(P+1)} \end{aligned} \tag{144}$$

and the RHS of (139) becomes

$$\mathbb{E}\Big[\mathcal{D}[k-\Delta] \circ \bar{\mathcal{Y}}^*[k]\Big] = \mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} \tag{145}$$

Using (144) and (145), (139) becomes

$$\{\bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}}}\}_{(P+1)} = \mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} \tag{146}$$

To find the optimal tap co-efficients, we solve (146) by contracting both sides of (146) with $\mathcal{R}^{-1}_{\bar{\mathcal{Y}}}$. This gives

$$\Big\{\{\bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}}}\}_{(P+1)}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)} = \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)} \tag{147}$$

Using the associativity property we get

$$\Big\{\bar{\mathcal{G}}, \Big\{\mathcal{R}_{\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}\Big\}_{(P+1)} = \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)} \tag{148}$$

Since $\{\mathcal{R}_{\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\}_{(P+1)} = \mathcal{I}_{P+1}$, we have the optimal tap coefficients

$$\bar{\mathcal{G}}_{\text{opt}} = \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}. \tag{149}$$

C. PERFORMANCE ANALYSIS

We begin by deriving the mean squared error for the linear equalizers. The auto-correlation of the error tensor can be written as

$$\begin{aligned}\mathcal{R}_{\mathcal{E}} &= \mathbb{E}\big[\mathcal{E}[k] \circ \mathcal{E}^*[k]\big] \\ &= \mathbb{E}\Big[\Big(\{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} - \mathcal{D}[k]\Big) \circ \Big(\{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} - \mathcal{D}[k]\Big)^*\Big] \\ &= \mathbb{E}\Big[\{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} \circ \Big\{\bar{\mathcal{G}}^*, \bar{\mathcal{Y}}^*[k]\Big\}_{(P+1)} - \{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} \circ \mathcal{D}^*[k] \\ &\quad - \mathcal{D}[k] \circ \Big\{\bar{\mathcal{G}}^*, \bar{\mathcal{Y}}^*[k]\Big\}_{(P+1)} + \mathcal{D}[k] \circ \mathcal{D}[k]^*\Big] \end{aligned} \tag{150}$$

Using (16) in (150) we get

$$\begin{aligned}\mathcal{R}_{\mathcal{E}} = \mathbb{E}\Big[&\{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} \circ \Big\{\bar{\mathcal{Y}}^*, \bar{\mathcal{G}}^H\Big\}_{(P+1)} + \mathcal{D}[k] \circ \mathcal{D}[k]^* \\ &- \{\bar{\mathcal{G}}, \bar{\mathcal{Y}}[k]\}_{(P+1)} \circ \mathcal{D}^*[k] - \mathcal{D}[k] \circ \Big\{\bar{\mathcal{Y}}^*[k], \bar{\mathcal{G}}^H\Big\}_{(P+1)}\Big] \end{aligned} \tag{151}$$

Using the associativitiy property, (151) becomes

$$\begin{aligned}\mathcal{R}_{\mathcal{E}} = &\Big\{\bar{\mathcal{G}}, \Big\{\mathcal{R}_{\bar{\mathcal{Y}}}, \bar{\mathcal{G}}^H\Big\}_{(P+1)}\Big\}_{(P+1)} - \Big\{\bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}\Big\}_{(P+1)} \\ &- \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \bar{\mathcal{G}}^H\Big\}_{(P+1)} + \mathcal{R}_{\mathcal{D}} \end{aligned} \tag{152}$$

Substituting $\bar{\mathcal{G}}_{\text{ZF}}$ from (131) in (152), we get

$$\begin{aligned}\mathcal{R}_{\mathcal{E}_{\text{ZF}}} = &\Big\{\Big\{\mathcal{P}, \bar{\mathcal{H}}^+[k]\Big\}_{(N+1)}, \Big\{\mathcal{R}_{\bar{\mathcal{Y}}}, \Big\{\mathcal{P}, \bar{\mathcal{H}}^+[k]\Big\}^H_{(N+1)}\Big\}_{(P+1)}\Big\}_{(P+1)} \\ &- \Big\{\Big\{\mathcal{P}, \bar{\mathcal{H}}^+[k]\Big\}_{(N+1)}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}\Big\}_{(P+1)} \\ &- \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \Big\{\mathcal{P}, \bar{\mathcal{H}}^+[k]\Big\}^H_{(N+1)}\Big\}_{(P+1)} + \mathcal{R}_{\mathcal{D}} \end{aligned} \tag{153}$$

and the mean squared error is $\text{MSE}_{\text{ZF}} = \textit{trace}(\mathcal{R}_{\mathcal{E}_{\text{ZF}}})$. For the MMSE equalizer, substituting the optimal system tensor from (149) in (152), we get

$$\begin{aligned}\mathcal{R}_{\mathcal{E}_{\text{MMSE}}} = &\Big\{\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}, \Big\{\mathcal{R}_{\bar{\mathcal{Y}}}, \Big(\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}\Big)^H\Big\}_{(P+1)}\Big\}_{(P+1)} \\ &- \Big\{\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}\Big\}_{(P+1)} \\ &- \Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \Big(\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \mathcal{R}^{-1}_{\bar{\mathcal{Y}}}\Big\}_{(P+1)}\Big)^H\Big\}_{(P+1)} + \mathcal{R}_{\mathcal{D}} \end{aligned} \tag{154}$$

$$\begin{aligned}= &\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \Big\{\mathcal{R}^{-1}_{\bar{\mathcal{Y}}}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}\Big\}_{(P+1)}\Big\}_{(P+1)} \\ &- 2\Big\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \Big\{\mathcal{R}^{-1}_{\bar{\mathcal{Y}}}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}\Big\}_{(P+1)}\Big\}_{(P+1)} + \mathcal{R}_{\mathcal{D}} \end{aligned} \tag{155}$$

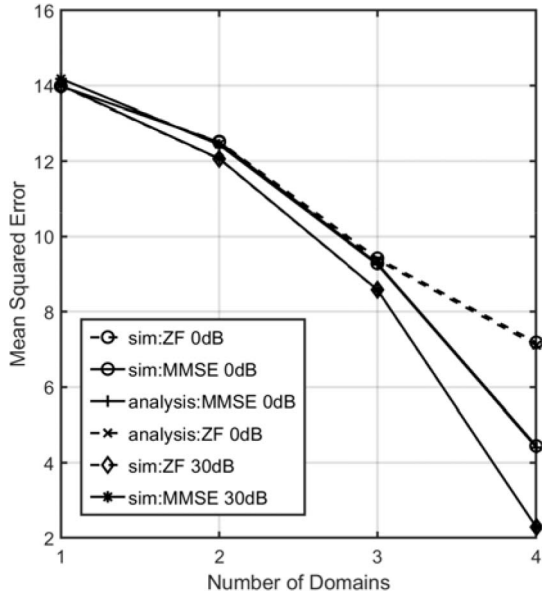

FIGURE 4. Minimum Mean Squared Error versus number of receive domains for a system with 4 transmit domains of size 2 each and $E_S = 1$.

$$= \mathcal{R}_{\mathcal{D}} - \left\{ \mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}, \left\{ \mathcal{R}_{\bar{\mathcal{Y}}}^{-1}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}} \right\}_{(P+1)} \right\}_{(P+1)} \tag{156}$$

$$= \mathcal{R}_{\mathcal{D}} - \left\{ \bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}} \right\}_{(P+1)} \tag{157}$$

and the minimum mean squared error becomes

$$\begin{aligned} \text{MSE}_{\text{MMSE}} &= trace\left(\mathcal{R}_{\mathcal{E}_{\text{MMSE}}}\right) \\ &= trace\left(\mathcal{R}_{\mathcal{D}} - \left\{ \bar{\mathcal{G}}, \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}} \right\}_{(P+1)} \right) \end{aligned} \tag{158}$$

Next, we present some simulation results for different systems using our tensor framework. The energy per tensor component is denoted by E_s and the total energy of a tensor symbol E_T is the sum of the energies of its components. The SNR is defined as $\frac{E_s}{N_0}$. The signal to noise ratio (SNR) is varied by changing N_0 for a fixed signal energy. In Fig. 4 we present the mean squared error vs P, the number of domains at the receiver, for the MMSE and ZF equalizers. The simulations were carried out at two fixed SNRs (0dB and 30dB). The results are averaged over 1000 channel realizations. The input $\mathcal{D}[k] \in \mathbb{C}_k^{2\times 2\times 2\times 2}$ is a fourth order tensor with components drawn from an i.i.d source, and 16-QAM is used for modulation. Hence we have $N = 4$, implying $\mathcal{R}_{\mathcal{D}} = \mathcal{I}_4$ and $\mathcal{R}_{\bar{\mathcal{D}}} = \mathcal{I}_5$. The noise tensor $\mathcal{V}[k] \in \mathbb{C}_k^{\overbrace{2\times\cdots\times 2}^{P}}$, whose components are complex Gaussian with zero mean and variance N_0, has the same size as the received tensor $\mathcal{Y}[k] \in \mathbb{C}_k^{\overbrace{2\times\cdots\times 2}^{P}}$. The channels used for the simulation contain three tensor taps ($v = 2$) $\mathcal{H}[0]$, $\mathcal{H}[1]$ and $\mathcal{H}[2]$, whose components are randomly generated complex zero-mean uncorrelated Gaussian random variables with unit variance per complex sample. The equalizer used for the simulations contains $M = 7$ taps. Hence, we have $\bar{\mathcal{V}}[k] \in \mathbb{C}_k^{7\times 2\times\cdots\times 2}$, $\bar{\mathcal{Y}}[k] \in \mathbb{C}_k^{7\times 2\times\cdots\times 2}$ and $\bar{\mathcal{H}} \in \mathbb{C}_k^{7\times 2\times\cdots\times 2\times 9\times 2\times\cdots\times 2}$, as defined in (120) and (121), such that (140) and (141) become

$$\mathcal{R}_{\bar{\mathcal{Y}}} = \left\{ \bar{\mathcal{H}}, \bar{\mathcal{H}}^H \right\}_{(5)} + N_0 \mathcal{I}_{P+1} \tag{159}$$

and

$$\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} = \left\{ \mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}}, \bar{\mathcal{H}}^H \right\}_{(5)} \tag{160}$$

Substituting the values of $\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}}$ and $\mathcal{R}_{\bar{\mathcal{Y}}}$ from (159) and (160) in (153) and (154), and using $\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} = \mathcal{R}_{\bar{\mathcal{Y}},\mathcal{D}}^H$, gives the analytical mean squared error for the ZF and MMSE equalizers. The analytical mean squared error for the MMSE equalizer is shown in (161) at the bottom of the next page. The expression for the ZF equalizer is omitted here for the sake of brevity. The simulations we carried out consider four different number of domains at the receiver which are summarized in Table 1. The simulated mean squared error is consistent with the mean squared error values obtained analytically from (153) and (161). It can be seen that the ZF and MMSE equalizers perform similarly at the high SNR value of $30dB$ regardless of the number of receiver domains. At the lower SNR value of $0dB$, the two equalizers perform similarly for a small number of domains, but the MMSE out performs the zero forcing equalizer as the number of receiver domains is increased. Further, the mean squared error decreases as the number of domains at the receiver is increased. Hence performance improvements can be made to a system by the addition of domains rather than having to increase the size of the individual domains themselves. However, the MMSE equalizer harnesses an increased number of domains better than the ZF equalizer does. We use the MMSE equalizer at an SNR of 30dB as an example and compute the mean squared error for a system with a channel of size $\mathbb{C}^{2\times 2\times 2\times 2\times 8}$. This system contains only one domain at the receiver of size 8. The mean squared error for this case, at SNR = $30dB$, was found to be $MSE = 8.16$. From Fig. 4, we can see that this value of mean squared error is reached when there are 3 domains of size 2 each at the receiver. Similarly, the mean squared error for a system of size $\mathbb{C}^{2\times 2\times 2\times 2\times 16}$ at SNR = $30dB$ was found to be $MSE = 1.35$. From Fig. 4, we can see that this value of mean squared error is reached when there are 4 domains of size 2 each at the receiver. In summary, the MSE for a system with 3 equal domains of size 2 reaches the same value as a system with one large receive domain of size 8, and a system with 4 equal domains of size 2 has the performance of a system with one large receive domain of size 16. This is useful when the size of a domain is constrained. For example in certain cases, performance improvements through the addition of frequency or time domains might be more desirable as compared to the addition of more antennas.

In Fig. 5, we compare the performance of tensor based multi-domain equalization with a Per Domain Equalizer (PDE). The input $\mathcal{D}[k] \in \mathbb{C}_k^{2\times 2}$ is a second order tensor with components drawn from i.i.d source, with $16QAM$ used for

TABLE 1. Dimension sizes of the receiver.

No. of Domains	Size of $\mathcal{Y}[k]$	Size of $\mathcal{H}[k]$
1	$\mathbb{C}^{2}$	$\mathbb{C}^{2\times2\times2\times2}$
2	$\mathbb{C}^{2\times2}$	$\mathbb{C}^{2\times2\times2\times2\times2\times2}$
3	$\mathbb{C}^{2\times2\times2}$	$\mathbb{C}^{2\times2\times2\times2\times2\times2\times2}$
4	$\mathbb{C}^{2\times2\times2\times2}$	$\mathbb{C}^{2\times2\times2\times2\times2\times2\times2\times2}$

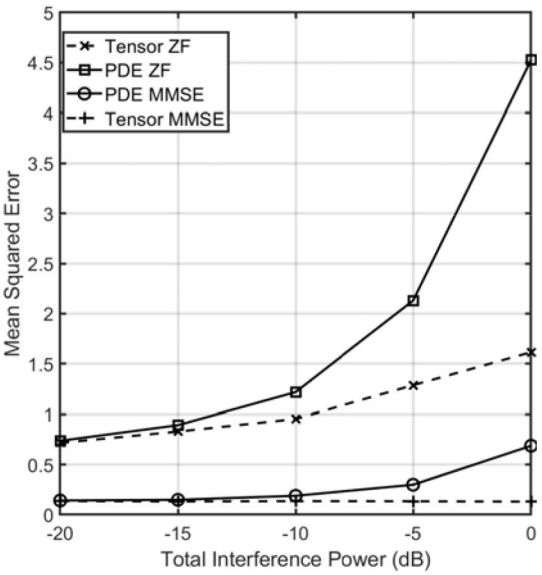

FIGURE 5. Mean Squared Error versus Interference Power for single domain and multi domain equalization for SNR = 10*dB*.

modulation. The energy per tensor component $E_s = 1$. The channel $\mathcal{H}[k] \in \mathbb{C}_k^{2\times2\times2\times2}$ is a fourth order tensor with 1 tensor tap (i.e., $\mathcal{H}[k] = \mathbf{0}$ if $k \neq 0$) and its components are randomly generated complex zero-mean Gaussian random variables. We define the Total Interference Power I_T as the combined energy of those components of the channel that contribute to inter-domain interference. For example, consider the component $\mathcal{Y}_{1,1}[k]$. We have

$$\mathcal{Y}_{1,1}[k] = \mathcal{H}_{1,1,1,1}\mathcal{D}_{1,1}[k] + \overbrace{\mathcal{H}_{1,1,1,2}\mathcal{D}_{1,2}[k]}^{\text{intra-domain interference}} + \overbrace{\mathcal{H}_{1,1,2,1}\mathcal{D}_{2,1}[k] + \mathcal{H}_{1,1,2,2}\mathcal{D}_{2,2}[k]}^{\text{inter-domain interference}} + \mathcal{N}_{1,1}[k] \tag{162}$$

The inter-domain interference in this case is from the third and fourth component of (162) and summing the energy of all such components gives us the total interference power I_T. We compare the performance of multi-domain equalization, where the tensor MMSE and ZF equalizers are used considering both inter-domain and intra-domain interference vs PDE where only intra-domain interference is considered for different values of I_T in dB. The PDE ignores the inter-domain interference and hence the third term in the RHS of (162) is ignored. The equalizers for this case are hence designed using

$$\mathcal{Y}_{1,1}[k] = \mathcal{H}_{1,1,1,1}\mathcal{D}_{1,1}[k] + \mathcal{H}_{1,1,1,2}\mathcal{D}_{1,2}[k] + \mathcal{N}_{1,1}[k] \tag{163}$$

which generalizes, for component $\mathcal{Y}_{i,j}$, to

$$\mathcal{Y}_{i,j}[k] = \sum_{m=1}^{2} \mathcal{H}_{i,j,i,m}\mathcal{D}_{i,m}[k] + \mathcal{N}_{i,j}[k] \tag{164}$$

Define $\mathbf{y}^{(i)} = [\mathcal{Y}_{i,1}, \mathcal{Y}_{i,2}]^{\mathrm{T}}$, $\mathbf{d}^{(i)} = [\mathcal{D}_{i,1}, \mathcal{D}_{i,2}]^{\mathrm{T}}$, $\mathbf{n}^{(i)} = [\mathcal{N}_{i,1}, \mathcal{N}_{i,2}]^{\mathrm{T}}$ and

$$\mathbf{H}^{(i)} = \begin{bmatrix} \mathcal{H}_{i,1,i,1} & \mathcal{H}_{i,1,i,2} \\ \mathcal{H}_{i,2,i,1} & \mathcal{H}_{i,2,i,2} \end{bmatrix} \tag{165}$$

Using (164) we get

$$\mathbf{y}^{(i)} = \mathbf{H}^{(i)}\mathbf{d}^{(i)} + \mathbf{n}^{(i)} \tag{166}$$

and the MMSE and ZF equalizers for the PDE are

$$\mathcal{G}_{\text{PDMMSE}}^{(i)} = \left(\mathbf{H}^{(i)}\right)^{\mathrm{H}} \left(\mathbf{H}^{(i)}\left(\mathbf{H}^{(i)}\right)^{\mathrm{H}} + N_0\mathbf{I}\right)^{-1} \tag{167}$$

$$\mathcal{G}_{\text{PDZF}}^{(i)} = \left(\mathbf{H}^{(i)}\right)^{+} \tag{168}$$

Hence, each PDE consists of two separate equalizers for $i = 1$ and $i = 2$. Further, to allow us to compare the mean squared errors of the two types of equalizers we define the MMSE as the total minimum mean squared error per domain, i.e., we have

$$\begin{aligned} MMSE &= \frac{1}{2} trace\left(\mathbb{E}\left[\left(\mathcal{D} - \hat{\mathcal{D}}\right) \circ \left(\mathcal{D} - \hat{\mathcal{D}}\right)^{*}\right]\right) \\ MMSEPD &= \frac{1}{2}\sum_{i=1}^{2} trace\left(\mathbb{E}\left[\left(\mathbf{d}^{(i)} - \hat{\mathbf{d}}^{(i)}\right) \circ \left(\mathbf{d}^{(i)} - \hat{\mathbf{d}}^{(i)}\right)^{*}\right]\right) \end{aligned} \tag{169}$$

The SNR used for the simulation is 10dB. It can be seen that the two types of equalizers perform similarly at low values of I_T but there is an aggressive degradation in relative performance of the single domain equalizers as the power of the interference is increased. The tensor framework provides an intuitive way to combat interference from multiple domains simultaneously. Hence, in scenarios where additional domains are introduced into a system, the tensor framework can be used to ensure acceptable performance.

In Fig. 6 we present the error rate for different equalizer tap lengths $M = 2K + 1$ ($K = 3, 5, 7$) plotted against the averaged receive SNR for the finite tap MMSE equalizer.

$$\mathcal{R}_{\mathcal{E}_{\text{MMSE}}} = \mathcal{I}_4 - \left\{\left(\left\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}}, \bar{\mathcal{H}}^{H}\right\}_{(5)}\right), \left\{\left(\left\{\bar{\mathcal{H}}, \bar{\mathcal{H}}^{H}\right\}_{(5)} + N_0\mathcal{I}_{(P+1)}\right)^{-1}, \left(\left\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}}, \bar{\mathcal{H}}^{H}\right\}_{(5)}\right)^{H}\right\}_{(P+1)}\right\}_{(P+1)} \tag{161}$$

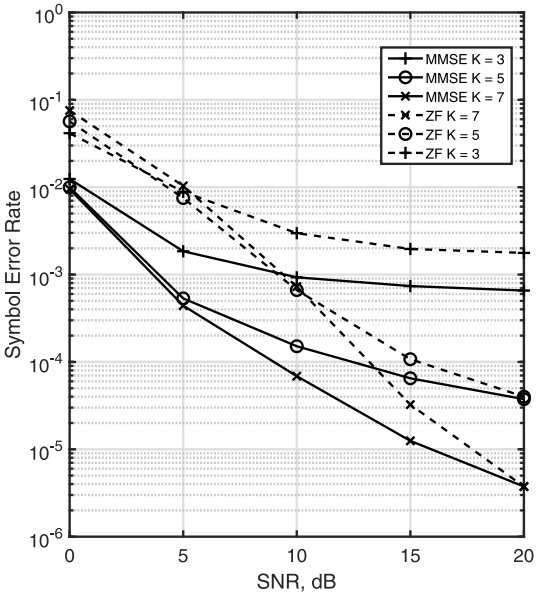

FIGURE 6. Finite Tap MMSE Equaliser for $L = 2$.

Also shown for comparison is the performance for of zero forcing equalizers with the same number of taps. The input data tensor $\mathcal{D}[k]$ is of size $2 \times 2 \times 2$ with components drawn from an i.i.d source, with $\mathcal{R}_{\mathcal{D}}[i] = \mathcal{I}_3$, and 4-QAM is used for modulation. The channel used consists of two taps ($v = 1$), i.e., the received tensor $\mathcal{Y}[k] \in \mathbb{C}_k^{2\times2\times2}$ only contains inter-tensor interference from $\mathcal{D}[k-1]$. The channel is assumed to be time-invariant and known at the receiver. For each realization of a test channel, the components of $\mathcal{H}[k]$ are drawn from a complex Gaussian distribution such that each complex sample has zero mean and unit variance. In this case (140) and (141) become

$$\mathcal{R}_{\bar{\mathcal{Y}}} = \left\{\bar{\mathcal{H}}, \bar{\mathcal{H}}^H\right\}_{(4)} + N_0\mathcal{I}_4 \tag{170}$$

and

$$\mathcal{R}_{\mathcal{D},\bar{\mathcal{Y}}} = \left\{\mathcal{R}_{\mathcal{D},\bar{\mathcal{D}}}, \bar{\mathcal{H}}^H\right\}_{(4)} \tag{171}$$

The coefficients of the equalizer are calculated using (149) and the error rate is found by averaging MATLAB simulation results over 100 channel realizations, accumulating 250 errors at each SNR. We can see that there is an improvement in error performance as the number of tensor taps in the equalizer increases. Further, the performance of the zero forcing equalizer for the same number of taps is worse than the MMSE equalizer but the difference between their performance decreases with increased SNR. In Fig. 7 we present results for a channel with three taps ($v = 2$). We can see that the error floors are formed as the SNR is increased due to the fact that the interference is not completely removed. The error floors are higher for three channel taps when compared to the two-tap channel scenario of Fig. 6. Once again, the zero forcing equalizer performs worse than the MMSE equalizer and yields a higher error floor. Table 2 summarizes the simulation parameters for Figs. 6 and 7.

TABLE 2. Simulation parameters.

Description	Fig. 6	Fig. 7
Number of channel taps	2	3
Size of data	$2 \times 2 \times 2$	$2 \times 2 \times 2$
Size of received tensor	$2 \times 2 \times 2$	$2 \times 2 \times 2$
Modulation	4-QAM	4-QAM

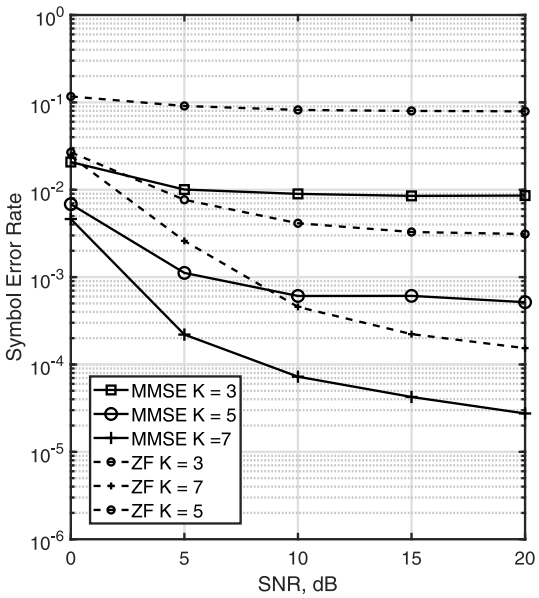

FIGURE 7. Finite Tap MMSE Equaliser for $L = 3$.

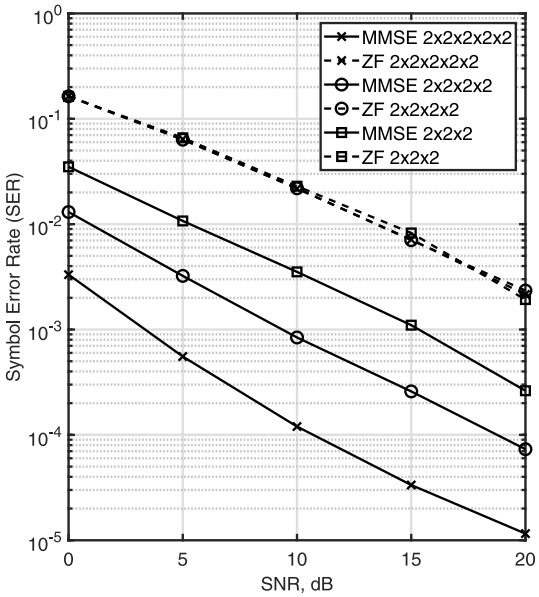

FIGURE 8. ZF and MMSE Equalizers for zero inter-tensor interference. The size in the legend corresponds to the size of the transmit tensor.

The error performance when there is no inter-tensor interference is illustrated in Fig. 8. We consider the case when there is only intra-tensor interference, and hence each

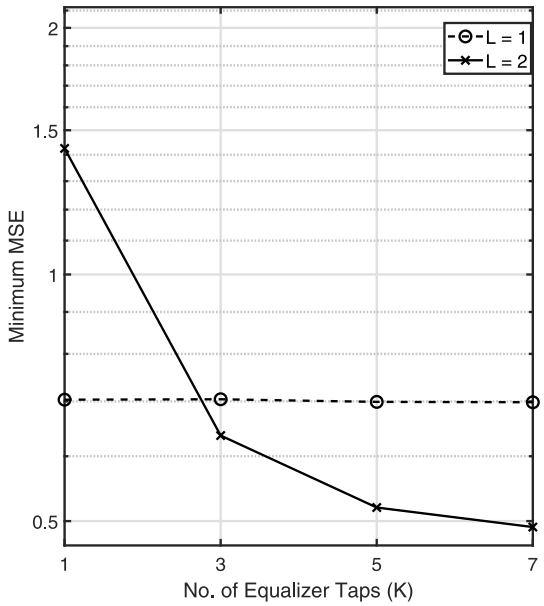

FIGURE 9. MSE for different equalizer number of taps, $E_S = 1$.

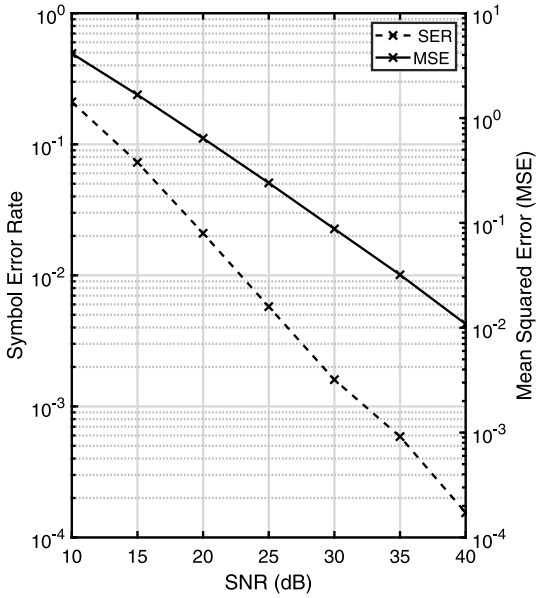

FIGURE 10. MSE (right) and SER (left) vs SNR for MIMO GFDM, $E_S = 1$.

component of $\mathcal{Y}[k]$ is a linear combination of the components of $\mathcal{D}[k]$. It is assumed that $\mathcal{R}_{\mathcal{D}}[0] = \mathcal{I}_{(N)}$ and $\mathcal{R}_{\mathcal{V}}[0] = N_0\mathcal{I}_{(N)}$. The channel contains only one non-zero tap ($v = 0$), $\mathcal{H} = \mathcal{H}[0]$ and the equalizer also contains one tap $\mathcal{G}$. The simulations are carried out for channels of three different sizes: $(2\times2\times2\times2\times2\times2)$, $(2\times2\times2\times2\times2\times2\times2\times2)$ and $(2\times2\times2\times2\times2\times2\times2\times2\times2\times2)$ with corresponding sizes of transmit tensors: $(2\times2\times2)$, $(2\times2\times2\times2)$ and $(2\times2\times2\times2\times2)$. We can see that the performance of the MMSE equalizer is better than the zero forcing equalizer. This is attributed to the fact that there is significant noise enhancement due to the contraction of the noise tensor $\mathcal{V}[k]$ with $\mathcal{H}^{-1}$. Further, as the size of the channels, and hence the number of domains in the receiver, increases, the MMSE equalizer performs better. This is because increasing the number of domains, and hence the number of samples, results in better averaging of the noise. Unlike the MMSE equalizer that optimizes the mean squared error, the zero forcing equalizer simply eliminates the interference from the other components of the data tensor at the expense of noise enhancement. In the three systems of Fig. 8, the number of transmit domains increases at the same rate as the number of receiver domains and any gain in performance from the additional receive domains is nullified by the additional data being transmitted on the added transmit domains that increases the noise enhancement due to channel inversion.

Comparing Fig. 8 with Fig. 6 we can see that the multi-tap equalizers perform better even though there is more interference in the system of Fig. 6. The variation of the mean squared error with the number of equalizer taps at SNR $= 5dB$ is shown in Fig. 9. As we can see, the equalizer for the case where there is no inter-tensor interference ($L = 1$) does not benefit from increasing the number of taps while the equalizer with $L = 2$ performs worse for $K = 1$ but improves as the number of taps is increased. As illustrated in Fig. 9 we see that as the number of equalizer taps is increased, the performance improves in the case of the system with inter-tensor interference. This is not true for the case where there is only intra-tensor interference.

VI. APPLICATIONS OF THE TENSOR FRAMEWORK TO MULTI-USER MIMO GFDM

In this section we explore applications of the tensor framework by studying higher order extensions of GFDM in the form of MIMO GFDM and Multi-user MIMO GFDM.

A. MIMO GFDM

The GFDM system model and its extension to MIMO GFDM was considered in Section IV-C. In this section we consider the performance of a MIMO GFDM system using a linear MMSE equalizer. The data being transmitted in the nth MIMO GFDM symbol (which contains P streams of K subcarriers and M subsymbols) is a third order tensor $\mathcal{D}[n] \in \mathbb{C}_n^{K\times M\times P}$ with $\mathcal{R}_{\mathcal{D}} = \mathcal{I}_3$. There are 2 transmit and 2 receive antennas. We have assumed that there is no interference between successive MIMO GFDM symbols due to the use of a cyclic prefix and hence there is only intra-tensor interference. Hence, the discrete time channel model is given by (106), where the overall channel is a sixth order tensor $\mathcal{H}[n] \in \mathbb{C}_n^{K\times M\times P\times K\times M\times P}$ that couples the input with an output, which is another third order tensor $\mathcal{Y}[n] \in \mathbb{C}_n^{K\times M\times P}$. The transmit and receive system tensors are as described in Section IV and the components of the channel tensor are $\mathcal{H}_{C_{n_r,n_t}}(t) = \sum_{q=0}^{15} a_{q,nr,nt}\delta(t - q\frac{T}{KM})$ where $a_{q,nr,nt}$ are independent complex Gaussian random variables with zero mean and unit variance. The equalizer consists of one tap and the delay used is $\Delta = 0$. Monte-Carlo simulation results assessing performance are presented in Fig. 10. Results from

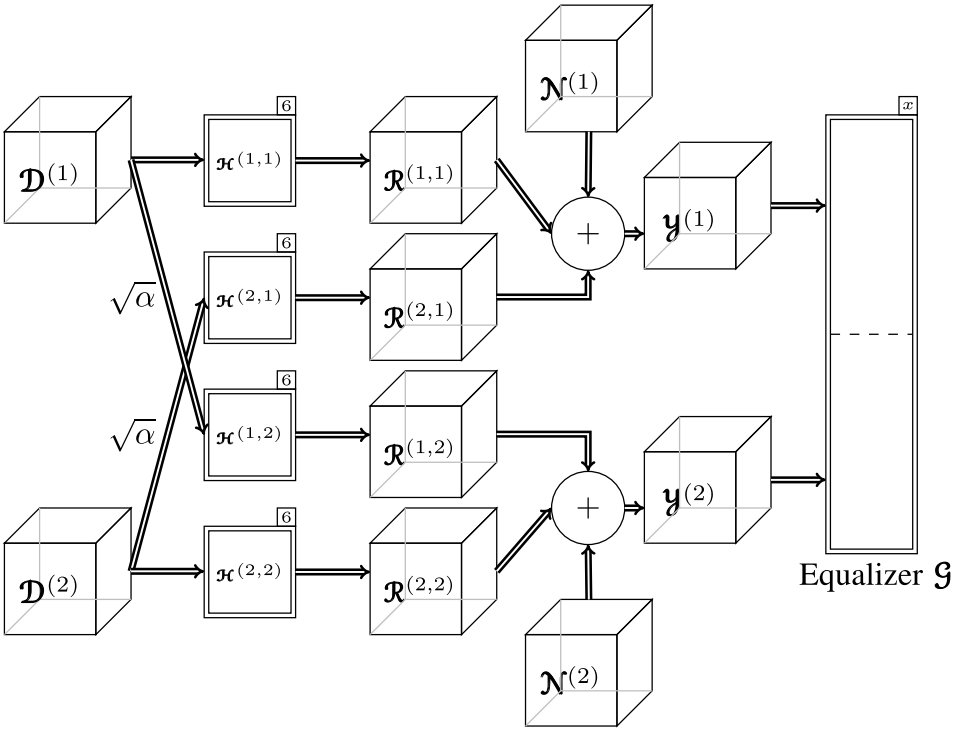

FIGURE 11. System Model for two user MIMO GFDM.

TABLE 3. Simulation parameters for MIMO GFDM.

Description	parameter	value
No. of subcarriers	K	16
No. of time slots	M	5
pulse shaping filter	g	Raised Cosine (RC)
roll-off factor	α	1
No. of streams	P	2
No. of transmit,receive antennas	T, R	2,2
modulation order	μ	6 (64-QAM)

1000 different channel realizations were averaged to find the mean squared error. The error rate is found by averaging MATLAB simulation results over 100 channel realizations, accumulating 250 errors at each SNR. The other parameters used for the simulations are defined in Table 3. Our results show that the SER decreases as the SNR increases, and there is no error floor. This is because there is no inter-tensor interference.

Using the tensor framework allows us to maintain the structure of the system and this becomes important as the number of domains increases. For instance, as we shall see in the next sub-section, if we add an additional user domain then the transmitted data becomes four dimensional and representing this using matrices becomes increasingly more difficult to comprehend. On the other hand, as seen before, the addition of domains using the tensor framework is achieved more naturally and intuitively.

B. MULTI-USER MIMO-GFDM

Next, we explore the addition for a fourth domain to the MIMO-GFDM system discussed previously by considering

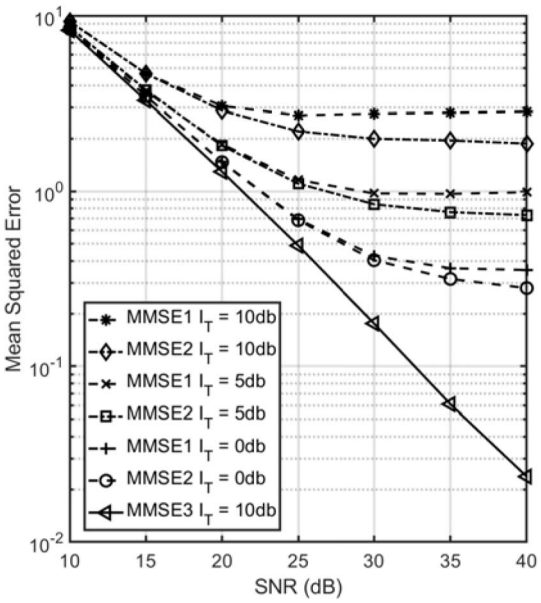

FIGURE 12. MSE vs SNR for the three different equalizers.

Multi-User MIMO-GFDM. Assume U users with data transmitted by the ith user represented by a third order tensor $\mathcal{D}^{(i)} \in \mathbb{C}^{P\times M\times K}$ where P is the number of GFDM streams per user, M is the number of time-slots and K is the number of sub-carriers. There is no interference between two successive data tensors and hence we have dropped the index $[n]$ for ease of representation. The components of the data tensor are i.i.d and $\mathcal{R}_{\mathcal{D}^{(i)}} = \mathcal{I}_{(3)}$. The channel tensor that couples

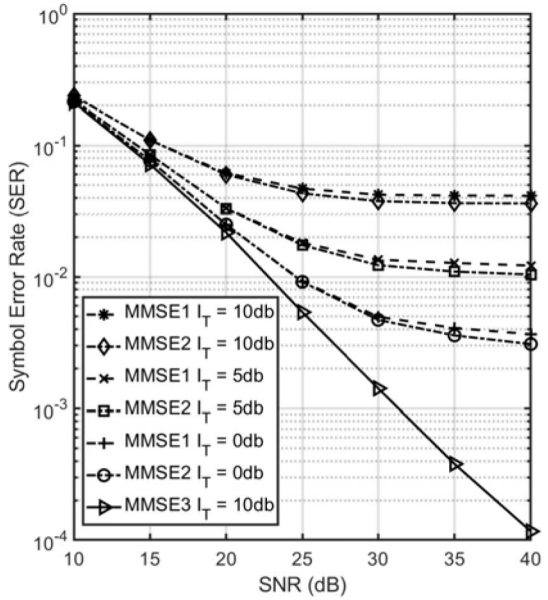

FIGURE 13. SER vs SNR for the three different equalizers.

the data of user i with the received tensor corresponding to user j is $\mathcal{H}^{(i,j)} \in \mathbb{C}^{P\times K\times M\times P\times K\times M}$. Each component of $\mathcal{H}^{(i,j)}$ is an independent realization of a complex Gaussian random variable with zero mean and unit variance. To quantify the interference between users we have assumed that the channel tensor $\mathcal{H}^{(i,j)}, i \neq j$ is scaled by $\sqrt{\alpha}$. The received tensor corresponding to user j is

$$\mathcal{Y}^{(j)} = \left\{\mathcal{H}^{(j,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \sqrt{\alpha}\sum_{\substack{i=1\\ i\neq j}}^{U}\left\{\mathcal{H}^{(j,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \mathcal{N}^{(j)} \tag{172}$$

where the first term on the RHS is the desired data corrupted by intra-user interference due to the contraction with $\mathcal{H}^{(j,j)}$, the second term is the interference from other users and the third term is additive thermal noise whose components are complex Gaussian random variables with zero mean and variance N_0. The system model for 2 users is illustrated in Fig. 11.

We study the performance of three types of equalizers for the above system. The first type of equalizer is a per-user equalizer that equalizes only the intra-tensor interference without considering the interference from other users. The equalizer for this case is designed by considering only the first and third terms on the RHS of (172). Hence the equalizer for user i becomes $\mathcal{G}^{(i)}_{\text{MMSE}} = \{\mathcal{R}_{\mathcal{D}^{(i)},\mathcal{Y}^{(i)}}, \mathcal{R}^{-1}_{\mathcal{Y}^{(i)}}\}_{(3)}$. We can write

$$\begin{aligned}
&\mathcal{R}_{\mathcal{D}^{(i)},\mathcal{Y}^{(i)}}\\
&= \mathbb{E}\left[\mathcal{D}^{(i)} \circ \left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)^{*}\right]\\
&= \left(\mathcal{H}^{(i,i)}\right)^{H} \qquad (173)\\
&\mathcal{R}_{\mathcal{Y}^{(i)}}\\
&= \mathbb{E}\left[\left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \mathcal{N}^{(i)}\right) \circ \left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)^{*}\right]\\
&= \left\{\mathcal{H}^{(i,i)}, \left(\mathcal{H}^{(i,i)}\right)^{H}\right\}_{(3)} + N_0\mathcal{I}_3 \qquad (174)
\end{aligned}$$

Hence

$$\mathcal{G}^{(i)}_{\text{MMSE1}} = \left\{\left(\mathcal{H}^{(i,i)}\right)^{\text{H}}, \left[\left\{\mathcal{H}^{(i,i)}, \left(\mathcal{H}^{(i,i)}\right)^{H}\right\}_{(3)} + N_0\mathcal{I}_3\right]^{-1}\right\}_{(3)} \tag{175}$$

The second type of equalizer considers the inter-user interference as noise and uses the second order statistics of the inter-user interference to aid in the equalization. Hence we have

$$\begin{aligned}
&\mathcal{R}_{\mathcal{D}^{(i)},\mathcal{Y}^{(i)}}\\
&= \mathbb{E}\left[\mathcal{D}^{(i)} \circ \left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \sqrt{\alpha}\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)^{*}\right]\\
&= \left(\mathcal{H}^{(i,i)}\right)^{H} \qquad (176)\\
&\mathcal{R}_{\mathcal{Y}^{(i)}}\\
&= \mathbb{E}\left[\left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \sqrt{\alpha}\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)\right.\\
&\qquad \left.\circ \left(\left\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\right\}_{(3)} + \sqrt{\alpha}\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)^{*}\right]\\
&= \left\{\mathcal{H}^{(i,i)}, \left(\mathcal{H}^{(i,i)}\right)^{H}\right\}_{(3)} + \mathcal{R}_{\mathcal{V}} \qquad (177)
\end{aligned}$$

where we have used the fact that $\mathbb{E}[\{\mathcal{H}^{(i,i)}, \mathcal{D}^{(i)}\}_{(3)} \circ (\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\}_{(3)})^{*}] = \mathbf{0}$ when $i \neq j$, and

$$\begin{aligned}
\mathcal{R}_{\mathcal{V}} &= \mathbb{E}\left[\left(\sqrt{\alpha}\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)\right.\\
&\qquad \left.\circ \left(\sqrt{\alpha}\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)} + \mathcal{N}^{(i)}\right)^{*}\right]\\
&= \alpha\,\mathbb{E}\left[\left(\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)}\right) \circ \left(\sum_{\substack{j=1\\ j\neq i}}^{U}\left\{\mathcal{H}^{(i,j)}, \mathcal{D}^{(j)}\right\}_{(3)}\right)^{*}\right]
\end{aligned}$$

$$+ N_0\mathfrak{I}_3$$
$$= \alpha \sum_{\substack{j=1 \\ j\neq i}}^{U} \mathbb{E}\left[\left\{\mathcal{H}^{(i,j)}, \left(\mathcal{H}^{(i,j)}\right)^H\right\}_{(3)}\right] + N_0\mathfrak{I}_3 \tag{178}$$

The components of $\mathbb{E}\left[\{\mathcal{H}^{(i,j)}, (\mathcal{H}^{(i,j)})^H\}_{(3)}\right]$ are

$$\left(\mathbb{E}\left[\left\{\mathcal{H}^{(i,j)}, \left(\mathcal{H}^{(i,j)}\right)^H\right\}_{(3)}\right]\right)_{p',k',m',p,k,m}$$
$$= \sum_{x=1,y=1,z=1}^{P,K,M} \mathbb{E}\left[\mathcal{H}^{(i,j)}_{p',k',m',x,y,z}\left(\mathcal{H}^{(i,j)}\right)^*_{p,k,m,x,y,z}\right]$$
$$= \begin{cases} \sum_{x=1,y=1,z=1}^{P,K,M} 1 & \text{if } p'=p, k'=k, m'=m \\ 0 & \text{otherwise} \end{cases} \tag{179}$$

which in tensor form becomes $\mathbb{E}\left[\{\mathcal{H}^{(i,j)}, (\mathcal{H}^{(i,j)})^H\}_{(3)}\right] = PMK \cdot \mathfrak{I}_3$. Substituting this in (178) gives

$$\mathcal{R}_{\mathcal{V}} = ((U-1)PMK\alpha + N_0)\mathfrak{I}_3 \tag{180}$$

Hence, the equalizer for user i becomes

$$\mathcal{G}^{(i)}_{\text{MMSE2}} = \left\{\left(\mathcal{H}^{(i,i)}\right)^H, \left[\left\{\mathcal{H}^{(i,i)}, \left(\mathcal{H}^{(i,i)}\right)^H\right\}_{(3)} + ((U-1)PMK\alpha + N_0)\mathfrak{I}_3\right]^{-1}\right\}_{(3)} \tag{181}$$

Notice that the second type of equalizer improves on the first type by accounting for the inter-user interference as noise, but does not harness the information included in cross-talk to aid in the detection of the other users data. The third type of equalizer is designed to simultaneously consider both the inter-user interference and the intra-user interference. Here we define a fourth order data tensor $\mathcal{D} \in \mathbb{C}^{P\times M\times K\times U}$ such that $\mathcal{D}_{:,:,:,i} = \mathcal{D}^{(i)}$ and a channel tensor of order 8 such that

$$\mathcal{H}_{:,:,:,i,:,:,:,j} = \mathcal{H}^{(i,j)} \tag{182}$$

We can write the received tensor as

$$\mathcal{Y} = \{\mathcal{H}, \mathcal{D}\}_{(4)} + \mathcal{N} \tag{183}$$

where $\mathcal{N}$ is a noise tensor of order 4 such that $\mathcal{N}_{:,:,:,i} = \mathcal{N}^{(i)}$. The equalizer for this case becomes

$$\mathcal{G}_{\text{MMSE3}} = \left\{\mathcal{H}^H, \left[\left\{\mathcal{H}, \mathcal{H}^H\right\}_{(4)} + N_0\mathfrak{I}_4\right]^{-1}\right\}_{(4)} \tag{184}$$

To illustrate the influence of α on the multi-user equalizer we write (184), for the two user case, in terms of its sub-tensors as shown in (185) on bottom the page.

where $\mathcal{G}^{(i,j)} = (\mathcal{G}_{\text{MMSE3}})_{:,:,:,i,:,:,:,j}$. Further, to allow us to compare the mean squared errors of the three types of equalizers we define the MMSE as the total minimum mean squared error per user, i.e., we have

$$MMSE = \frac{1}{U} trace\left(\mathbb{E}\left[\left(\mathcal{D} - \hat{\mathcal{D}}\right) \circ \left(\mathcal{D} - \hat{\mathcal{D}}\right)^*\right]\right)$$
$$= \frac{1}{U}\sum_{i=1}^{U} trace\left(\mathbb{E}\left[\left(\mathcal{D}^{(i)} - \hat{\mathcal{D}}^{(i)}\right) \circ \left(\mathcal{D}^{(i)} - \hat{\mathcal{D}}^{(i)}\right)^*\right]\right) \tag{186}$$

The MSE and SER versus SNR curves for different values of Total Interference Power (I_T) in dB for the equalizers discussed in this section are illustrated in Fig. 12 and Fig. 13 respectively. For the simulations we have used $U = 2$, $K = 16$, $M = 5$ and $P = 2$. The Total interference power is defined as the energy of the inter-user interference in each component of the received tensor. Consider the received components of $\mathcal{Y}^{(1)}$.

$$\mathcal{Y}^{(1)}_{p'k'm'} = \sum_{p,k,m} \mathcal{H}^{(1,1)}_{p'k'm'pkm}\mathcal{D}^{(1)}_{pkm} + \sum_{p,k,m} \sqrt{\alpha}\mathcal{H}^{(1,2)}_{p'k'm'pkm}\mathcal{D}^{(2)}_{pkm} + \mathcal{N}^{(1)}_{p'k'm'} \tag{187}$$

The inter-user interference is the second term on the RHS of (187) and the total interference power I_T is defined as the power of $(\sum_{p,k,m} \sqrt{\alpha}\mathcal{H}^{(1,2)}_{p'k'm'pkm}\mathcal{D}^{(2)}_{pkm})$. Hence we have

$$I_T = \alpha\,\mathbb{E}\left[\sum_{p,k,m} \mathcal{H}^{(1,2)}_{p',k',m',p,k,m}\mathcal{D}^{(2)}_{p,k,m} \times \left(\sum_{x,y,z} \mathcal{H}^{(1,2)}_{p',k',m',x,y,z}\mathcal{D}^{(2)}_{x,y,z}\right)^*\right]$$
$$= \alpha \sum_{p,k,m} \mathbb{E}\left[\mathcal{H}^{(1,2)}_{p',k',m',p,k,m} \cdot \mathcal{H}^{(1,2)*}_{p',k',m',p,k,m}\right] \times \mathbb{E}\left[\mathcal{D}^{(2)}_{p,k,m} \cdot \mathcal{D}^{(2)*}_{p,k,m}\right]$$
$$= PMK\alpha \tag{188}$$

The first per-user equalizer (MMSE1) does not consider the inter user interference and equalizes each user separately while the second per-user equalizer (MMSE2) uses the second order statistics of the inter-user interference. It can be

$$\begin{bmatrix} \mathcal{G}^{(1,1)} & \mathcal{G}^{(1,2)} \\ \mathcal{G}^{(2,1)} & \mathcal{G}^{(2,2)} \end{bmatrix} = \left\{\begin{bmatrix} \mathcal{H}^{(1,1)} & \sqrt{\alpha}\mathcal{H}^{(1,2)} \\ \sqrt{\alpha}\mathcal{H}^{(2,1)} & \mathcal{H}^{(2,2)} \end{bmatrix}, \left(\left\{\begin{bmatrix} \mathcal{H}^{(1,1)} & \sqrt{\alpha}\mathcal{H}^{(1,2)} \\ \sqrt{\alpha}\mathcal{H}^{(2,1)} & \mathcal{H}^{(2,2)} \end{bmatrix}, \begin{bmatrix} \mathcal{H}^{(1,1)} & \sqrt{\alpha}\mathcal{H}^{(1,2)} \\ \sqrt{\alpha}\mathcal{H}^{(2,1)} & \mathcal{H}^{(2,2)} \end{bmatrix}^H\right\}_{(4)} + N_0\mathfrak{I}_4\right)^{-1}\right\}_{(4)} \tag{185}$$

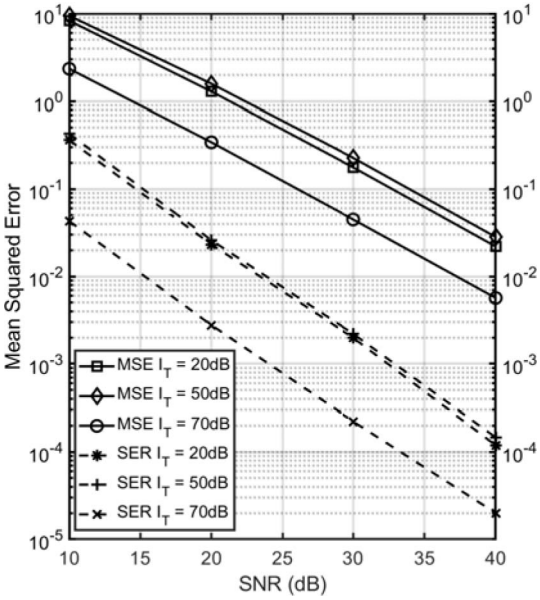

FIGURE 14. MSE and SER vs SNR for the multi-user equalizer with different values of interference power.

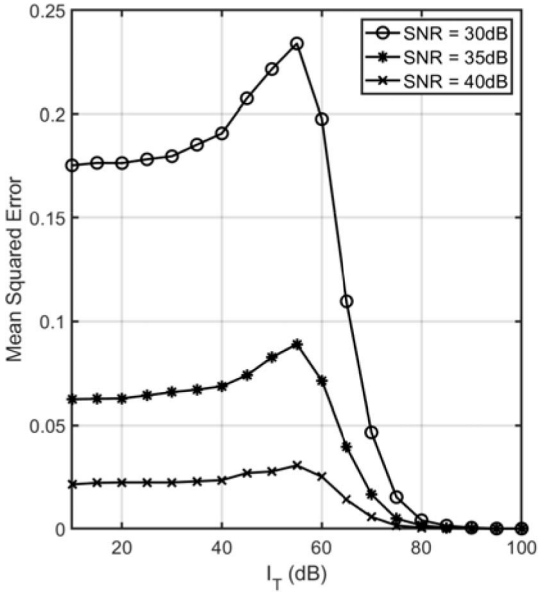

FIGURE 15. MSE of multi-user equalizer vs I_T for different values of SNR.

seen that the latter has improved performance in terms of the mean squared error (MSE) and the symbol error rate (SER). The multi-user equalizer (MMSE3) allows for simultaneous equalization of all the users at once and hence efficiently counteracts the effects of inter user interference. At low values of interference power I_T, the difference between the performance of the three equalizers is small. This implies that for systems where the inter-user interference is small, it might be more suitable to use per-user equalization as the performance gain in performing multi-domain equalization is small. However, as the value of I_T increases and the inter-user interference becomes more pronounced, the performance of the per-user equalizer deteriorates. The change in the performance of the multi-user MMSE equalizer is negligible when compared to the per-user equalizer for the values of interference shown and hence only the curve for $I_T = 10dB$ is presented for the Multi-domain MMSE.

To better understand the variation in performance of the multi-user equalizer, the MSE and SER curves for three different values of interference power I_T are illustrated in Fig. 14. It can be seen that the multi-user equalizer at $I_T = 50dB$ has a higher MSE and SER as compared to $I_T = 20dB$. However, when the interference power is increased further to $70dB$, it can be seen that the performance actually improves. This is because, beyond a certain interference power, the equalizer stars using the cross-talk for detection and the data from the direct link is treated as interference instead. This effectively improves the signal to noise ratio as signal power is now proportional to the cross-talk which is much stronger than the noise. To this end, the MSE vs interference power curves for different values of SNR are illustrated in Fig. 15. As we can see, the mean squared error initially increases with the interference power. However, it reaches a tipping point beyond which the cross-talk takes over as the main signal and hence causes the mean squared error to decrease with increasing interference power. This is an important distinction between the multi-user equalizer and the per-user equalizers. The former is able to make efficient use of the inter-user interference unlike the latter which either ignores it completely or treats it as noise.

VII. CONCLUSION

This paper presented a unified tensor framework, which can be used to represent, design and analyse communication systems that span several domains. The transmitted signals are represented by Nth order time function tensors which are coupled, using a system tensor of order $N + M$, with the received signals which are represented by another signal tensor of order M through the contracted convolution. The notion of a tensor of functions forms the basis for the definition of signal and system tensors. A generalization of the Nyquist's criterion for zero inter symbol interference was derived which allows unifying treatment of interference from several domains, dubbed multi domain interference (MDI). It is shown that known generalizations of the Nyquist's criterion are special cases of this tensor based Nyquist criterion. Furthermore, for the tensor case, a relaxation of the Nyquist criterion is possible, allowing recovery of data symbols even in the presence of intra-tensor interference. As examples, this tensor framework was used to model existing systems such as OFDM, GFDM and FBMC. Using the tensor framework, an example of higher domain extension for GFDM and FBMC was derived where different filters are used at the analysis and synthesis filter banks of each antenna.

As examples of using our tensor framework we derived joint domain linear tensor based equalizers for ZF and

MMSE criteria. The performance of the MMSE linear equalizer was found to be better than that of the ZF equalizer, which is consistent with the scalar case. For such linear tensor equalizers the mean squared error decreases as the number of domains at the receiver is increased while keeping the size of each domain the same. This implies that domain diversity can be leveraged to improve performance in cases where there are restrictions of the size of any particular domain, as is often the case in practice due to bandwidth limitations (frequency domain) or restrictions on the number of antennas (space domain). Further, it was found that the performance of the MMSE and ZF equalizers for the case when there is no inter-tensor interference can be worse than the case when there is inter-tensor and intra-tensor interference. Simulations were performed to investigate this phenomenon, and it was found that the performance of the equalizers in the latter improve as we increase the number of equalizer taps. This is because the residual interference and hence the error floor decreases as the number of tensor taps increases, and the performance tends towards that of an infinite tap equalizer. This is not true for the case when there is no inter-tensor interference as the performance of an equalizer with a single tap is the same as that of an infinite tap equalizer.

As a specific example of an application, we studied the use of the tensor framework for equalization in Multi-User MIMO GFDM for three different cases where the interference was treated differently. The first type of equalizer is a per domain equalizer (PDE) that equalizes only the interference from a single user while ignoring the inter-user interference completely. The second type of equalizer utilizes the second order statistics of the inter-user interference to achieve improvements in performance. The third type of equalizer, called the multi-domain equalizer, operates jointly in all domains and accounts for both inter-user interference and intra-user interference, and provides the best performance. Joint multi-domain equalization allows for efficient use of inter-user interference to aid in data detection as it is able to harness the information contained in cross-talk. The tensor framework can be used to efficiently design multi-user equalizers that are able to utilize the information contained in the interference between users. This concept can easily be extended to systems with any number of domains to design efficient multi-domain equalizers.

In order to keep the length of this paper reasonable, we present only examples of using the tensor framework for designing various linear equalizer/detector types. However in our work we have also considered decision-feedback type of tensor equalization, using this framework. Furthermore, it is known that significant improvement in performance can be achieved using iterative techniques instead of linear equalization, such as in [49]. The tensor framework of this paper can be used to derive such techniques capable of jointly operating in several domains, a subject that forms our following up work.

APPENDIX A
A CAUCHY-SCHWARTZ INEQUALITY

For two tensors $\mathcal{A}(t) \in \mathbb{C}_t^{J_1 \times \cdots \times J_M}$ and $\mathcal{B}(t) \in \mathbb{C}_t^{J_1 \times \cdots \times J_M}$ the following inequality holds:

$$\left| \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}_{j_1 \ldots j_M}(t) dt \right|^2 \leq \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{A}_{j_1 \ldots j_M}(t) \right|^2 dt \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{B}_{j_1 \ldots j_M}(t) \right|^2 dt. \tag{189}$$

Proof: Let λ be a complex scalar. We have

$$\begin{aligned} 0 &\leq \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{A}_{j_1 \ldots j_M}(t) + \lambda \mathcal{B}_{j_1 \ldots j_M}(t) \right|^2 dt \\ &= \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{A}_{j_1 \ldots j_M}(t) \right|^2 dt + |\lambda|^2 \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{B}_{j_1 \ldots j_M}(t) \right|^2 dt \\ &\quad + 2\Re \left\{ \lambda^* \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}^*_{j_1 \ldots j_M}(t) dt \right\} \end{aligned} \tag{190}$$

Choose $\lambda = y \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}^*_{j_1 \ldots j_M}(t) dt$ where $y \in \mathbb{R}$. This gives

$$0 \leq y^2 \left| \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}^*_{j_1 \ldots j_M}(t) dt \right|^2 \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{B}_{j_1 \ldots j_M}(t) \right|^2 dt \tag{191}$$

$$+ 2y \left| \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}^*_{j_1 \ldots j_M}(t) dt \right|^2 + \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{A}_{j_1 \ldots j_M}(t) \right|^2 dt \tag{192}$$

Define

$$\begin{aligned} p &= \left| \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \mathcal{A}_{j_1 \ldots j_M}(t) \mathcal{B}^*_{j_1 \ldots j_M}(t) dt \right|^2 \\ r &= \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{B}_{j_1 \ldots j_M}(t) \right|^2 dt \\ q &= \sum_{j_1 \ldots j_M} \int_{-\infty}^{+\infty} \left| \mathcal{A}_{j_1 \ldots j_M}(t) \right|^2 dt. \end{aligned} \tag{193}$$

Hence from (192)

$$0 \leq pry^2 + 2py + q \tag{194}$$

Since (194) is a non-negative quadratic polynomial, the discriminant is non-positive, i.e., $(2p)^2 - 4prq \leq 0$ or $p \leq rq$. Substituting for p,q and r gives (189), with equality when $\mathcal{B}_{i_1,\ldots,i_N,j_1,\ldots,j_M}(t) = \mathcal{A}^*_{j_1,\ldots,j_M,i_1,\ldots,i_N}(-t)$. ■

APPENDIX B
PROOF OF THEOREM 2

Consider the general system where a data sequence $\mathcal{D}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ is the input to a channel $\mathcal{H}[k] \in \mathbb{C}_k^{L_1\times\cdots\times L_M\times I_1\times\cdots\times I_N}$ and is corrupted by additive noise $\mathcal{V}[k] \in \mathbb{C}_k^{L_1\times\cdots\times L_M}$. The observation $\mathcal{Y}[k] \in \mathbb{C}_k^{L_1\times\cdots\times L_M}$ is

$$\mathcal{Y}[k] = \sum_{m=-\infty}^{+\infty} \{\mathcal{H}[m], \mathcal{D}[k-m]\}_{(N)} + \mathcal{V}[k] \tag{195}$$

The estimate of the data sequence is

$$\hat{\mathcal{D}}[k] = \sum_{m=0}^{+\infty} \{\mathcal{G}[m], \mathcal{Y}[k-m]\}_{(M)} \tag{196}$$

where $\mathcal{G}[m] \in \mathbb{C}^{I_1\times\cdots\times I_N\times L_1\times\cdots\times L_M}$. Denote the error by $\mathcal{E}[k] = \hat{\mathcal{D}}[k] - \mathcal{D}[k]$. We wish to prove that for the $\mathcal{G}[m]$ that minimizes the mean squared error between the estimate and the data, the error is uncorrelated with the observation, i.e., we wish to show that if

$$\mathbb{E}\big[\mathcal{E}[k] \circ \mathcal{Y}^*[k-i]\big] = \mathbf{0}_T \quad \text{for all } i \tag{197}$$

then the error is minimized in the mean squared sense. *Proof:* Assuming that the equalizer co-efficients are complex, the equalizer tensor may be written as $\mathcal{G}[m] = \mathcal{A}[m] + j\mathcal{B}[m]$. Extending the gradient vector in [50], we define a corresponding tensor gradient operator ∇, with components

$$\nabla_{i_1,\ldots,i_N,m,l_1,\ldots,l_M} = \frac{\partial}{\partial \mathcal{A}_{i_1,\ldots,i_N,l_1,\ldots,l_M}[m']} + j\frac{\partial}{\partial \mathcal{B}_{i_1,\ldots,i_N,l_1,\ldots,l_M}[m']} \tag{198}$$

Define the cost function

$$\begin{aligned} J(\mathcal{G}) &= \operatorname{trace}\big(\mathbb{E}\big[\mathcal{E}[k] \circ \mathcal{E}^*[k]\big]\big) \\ &= \sum_{i_1\ldots i_N} \mathbb{E}\big[\mathcal{E}_{i_1\ldots i_N}[k]\mathcal{E}^*_{i_1\ldots i_N}[k]\big] \end{aligned} \tag{199}$$

Due to the quadratic nature of the error surface, finding a stationary point assures global optimization of the cost function [50]. Minimizing the cost function is thus a convex unconstrained optimization problem [51] and can be solved be equating each component of the gradient tensor of the cost function to zero:

$$\nabla_{i_1,\ldots,i_N,m,l_1,\ldots,l_M} J\big(\mathcal{G}^{opt}\big) = 0 \tag{200}$$

Let us consider one particular component of the gradient tensor where the indices have values $i'_1, \ldots, i'_N, m', l'_1, \ldots, l'_M$. From (199) we get

$$\begin{aligned} &\nabla_{i'_1,\ldots,i'_N,m',l'_1,\ldots,l'_M} J(\mathcal{G}) \\ &= \nabla_{i'_1,\ldots,i'_N,m',l'_1,\ldots,l'_M} \mathbb{E}\left[\sum_{i_1,\ldots,i_N} \mathcal{E}_{i_1,\ldots,i_N}[k]\mathcal{E}^*_{i_1\ldots i_N}[k]\right] \\ &= \frac{\partial\, \mathbb{E}\big[\sum_{i_1,\ldots,i_N} \mathcal{E}_{i_1,\ldots,i_N}[k]\mathcal{E}^*_{i_1\ldots i_N}[k]\big]}{\partial \mathcal{A}_{i'_1,\ldots,i'_N,l'_1,\ldots,l'_M}[m']} \\ &\quad + j\frac{\partial\, \mathbb{E}\big[\sum_{i_1,\ldots,i_N} \mathcal{E}_{i_1,\ldots,i_N}[k]\mathcal{E}^*_{i_1\ldots i_N}[k]\big]}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']} \\ &= \mathbb{E}\left[\sum_{i_1,\ldots,i_N} \frac{\partial \mathcal{E}_{i_1,\ldots,i_N}[k]\mathcal{E}^*_{i_1\ldots i_N}[k]}{\partial \mathcal{A}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']} + j\frac{\partial \mathcal{E}_{i_1\ldots i_N}\mathcal{E}^*_{i_1\ldots i_N}}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}\right] \\ &= \mathbb{E}\Bigg[\sum_{i_1,\ldots,i_N} \frac{\partial \mathcal{E}_{i_1,\ldots,i_N}[k]}{\partial \mathcal{A}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}\mathcal{E}^*_{i_1\ldots i_N}[k] \\ &\qquad + \frac{\partial \mathcal{E}^*_{i_1\ldots i_N}[k]}{\partial \mathcal{A}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}\mathcal{E}_{i_1\ldots i_N}[k] \\ &\qquad + \frac{\partial \mathcal{E}_{i_1\ldots i_N}[k]}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}j\mathcal{E}^*_{i_1\ldots i_N}[k] \\ &\qquad + \frac{\partial \mathcal{E}^*_{i_1\ldots i_N}[k]}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}j\mathcal{E}_{i_1\ldots i_N}[k]\Bigg] \end{aligned} \tag{201}$$

The first term on the right hand side of (201) can be expanded as shown in (202) at the bottom of the page.

Similarly we have

$$\frac{\partial \mathcal{E}^*_{i_1\ldots i_N}[k]}{\partial \mathcal{A}_{i'_1,\ldots,i'_N,l'_1,\ldots,l'_M}[m']}\mathcal{E}_{i_1\ldots i_N}[k] = \mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k] \tag{203}$$

$$\frac{\partial \mathcal{E}_{i_1\ldots i_N}[k]}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}j\mathcal{E}^*_{i_1\ldots i_N}[k] = -\mathcal{Y}_{l'_1\ldots l'_M}[k-m']\mathcal{E}^*_{i'_1\ldots i'_N}[k] \tag{204}$$

$$\sum_{i_1\ldots i_N} \frac{\partial \mathcal{E}^*_{i_1\ldots i_N}[k]}{\partial \mathcal{B}_{i'_1\ldots i'_N l'_1\ldots l'_M}[m']}j\mathcal{E}_{i_1\ldots i_N}[k] = \mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k] \tag{205}$$

$$\begin{aligned} &\sum_{i_1,\ldots,i_N} \frac{\partial \mathcal{E}_{i_1,\ldots,i_N}[k]}{\partial \mathcal{A}_{i'_1,\ldots,i'_N,l'_1,\ldots,l'_M}[m']}\mathcal{E}^*_{i_1\ldots i_N}[k] \\ &= \sum_{i_1,\ldots,i_N} \frac{\partial\big\{\sum_m \sum_{l_1\ldots l_M} \mathcal{G}_{i_1\ldots i_N l_1\ldots l_M}[m]\mathcal{Y}_{l_1\ldots l_M}[k-m] - \mathcal{D}_{i_1\ldots i_N}[k]\big\}}{\partial \mathcal{A}_{i'_1,\ldots,i'_N,l'_1,\ldots,l'_M}[m']}\mathcal{E}^*_{i_1\ldots i_N}[k] \\ &= \sum_{i_1,\ldots,i_N}\sum_m\sum_{l_1\ldots l_M} \frac{\partial\big\{\big(\mathcal{A}_{i_1\ldots i_N l_1\ldots l_M}[m] + j\mathcal{B}_{i_1\ldots i_N l_1\ldots l_M}[m]\big)\mathcal{Y}_{l_1\ldots l_M}[k-m] - \mathcal{D}_{i_1\ldots i_N}[k]\big\}}{\partial \mathcal{A}_{i'_1,\ldots,i'_N,l'_1,\ldots,l'_M}[m']}\mathcal{E}^*_{i_1\ldots i_N}[k] \\ &= \mathcal{Y}_{l'_1\ldots l'_M}[k-m']\mathcal{E}^*_{i'_1\ldots i'_N}[k] \end{aligned} \tag{202}$$

Substituting these in (201) we get

$$\begin{aligned}
&\nabla_{i'_1,\ldots,i'_N,m',l'_1,\ldots,l'_M} J(\mathcal{G}) \\
&= \mathbb{E}\Big[\mathcal{Y}_{l'_1\ldots l'_M}[k-m']\mathcal{E}^*_{i'_1\ldots i'_N}[k] + \mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k] \\
&\qquad - \mathcal{Y}_{l'_1\ldots l'_M}[k-m']\mathcal{E}^*_{i'_1\ldots i'_N}[k] + \mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k]\Big] \\
&= 2\,\mathbb{E}\Big[\mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k]\Big] \qquad (206)
\end{aligned}$$

The optimal $\mathcal{G}[m]$ is found by equating (206) to 0 for all values of $i'_1, \ldots, i'_N, m', l'_1, \ldots, l'_M$. This gives

$$\mathbb{E}\Big[\mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\mathcal{E}_{i'_1\ldots i'_N}[k]\Big] = 0 \quad \forall\ i'_1 \ldots i'_N, m', l'_1, \ldots, l'_M \qquad (207)$$

We can see that the LHS of (207) is the auto-correlation between the error and the observation. We have

$$\begin{aligned}
\mathcal{R}_{\mathcal{E},\mathcal{Y}_{i'_1\ldots i'_N m' l'_1\ldots l'_M}}[-m'] &= \mathbb{E}\Big[\mathcal{E}_{l'_1\ldots l'_M}[k]\mathcal{Y}^*_{l'_1\ldots l'_M}[k-m']\Big] \\
&= 0 \quad \forall\ i'_1 \ldots i'_N, m', l'_1 \ldots l'_M \qquad (208)
\end{aligned}$$

Since (208) holds for all values of m', we can write this in tensor notation as

$$\begin{aligned}
\mathcal{R}_{\mathcal{E},\mathcal{Y}}[m] &= \mathbb{E}\big[\mathcal{E}[k] \circ \mathcal{Y}^*[k-m]\big] \\
&= \mathbf{0}_T \qquad (209)
\end{aligned}$$

showing that for the optimal $\mathcal{G}[m]$, the error is uncorrelated with the observation. This can be considered as a tensor orthogonality condition. ■

REFERENCES

[1] *Visual Networking Index*, Cisco Syst. Inc., San Jose, CA, USA, 2017, [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html

[2] G. Fettweis and S. Alamouti, "5G: Personal mobile Internet beyond what cellular did to telephony," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 140–145, Feb. 2014.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[4] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[5] P. Yang, Y. Xiao, Y. L. Guan, M. Di Renzo, S. Li, and L. Hanzo, "Multidomain index modulation for vehicular and railway communications: A survey of novel techniques," *IEEE Veh. Technol. Mag.*, vol. 13, no. 3, pp. 124–134, Sep. 2018.

[6] H. Gacanin, "Inter-domain bi-directional access in G.hn with network coding at the physical-layer," in *Proc. IEEE Int. Symp. Power Line Commun. Appl.*, Mar. 2012, pp. 144–149.

[7] V. Oksman and S. Galli, "G.hn: The new ITU-T home networking standard," *IEEE Commun. Mag.*, vol. 47, no. 10, pp. 138–145, Oct. 2009.

[8] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[9] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, nos. 1–4, pp. 164–189, 1927.

[10] R. B. Cattell, "Parallel proportional profiles and other principles for determining the choice of factors by rotation," *Psychometrika*, vol. 9, no. 4, pp. 267–283, Dec. 1944.

[11] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.

[12] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, Sep. 1970.

[13] R. Bro, "Review on multiway analysis in chemistry," *Critical Rev. Anal. Chem.*, vol. 36, nos. 3–4, pp. 279–293, 2006.

[14] C. J. Appellof and E. R. Davidson, "Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents," *Anal. Chem.*, vol. 53, no. 13, pp. 2053–2056, 1981.

[15] R. Bro, "PARAFAC. Tutorial and applications," *Chemometr. Intell. Lab. Syst.*, vol. 38, no. 2, pp. 149–171, 1997.

[16] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[17] B. Chen, A. P. Petropulu, and L. De Lathauwer, "Blind identification of convolutive MIMO systems with 3 sources and 2 sensors," *EURASIP J. Adv. Signal Process.*, vol. 2002, no. 5, pp. 487–496, May 2002.

[18] E. Acar, S. A. Çamtepe, M. S. Krishnamoorthy, and B. Yener, "Modeling and multiway analysis of chatroom tensors," in *Intelligence and Security Informatics*, P. Kantor *et al.*, Eds. Heidelberg, Germany: Springer, 2005, pp. 256–268.

[19] B. W. Bader, R. A. Harshman, and T. G. Kolda, "Temporal analysis of semantic graphs using ASALSAN," in *Proc. 7th IEEE Int. Conf. Data Mining (ICDM)*, Oct. 2007, pp. 33–42.

[20] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *Computer Vision—ECCV 2002*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. Heidelberg, Germany: Springer, 2002, pp. 447–460.

[21] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," in *Proc. Object Recognit. Supported User Interaction Serv. Robots*, vol. 2, Aug. 2002, pp. 511–514.

[22] D. Muti and S. Bourennane, "Multidimensional filtering based on a tensor approach," *Signal Process.*, vol. 85, no. 12, pp. 2338–2353, 2005.

[23] L. D. Lathauwer, J. Castaing, and J. F. Cardoso, "Fourth-order cumulant-based blind identification of underdetermined mixtures," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2965–2973, Jun. 2007.

[24] M. Brazell, N. Li, C. Navasca, and C. Tamon, "Solving multilinear systems via tensor inversion," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 2, pp. 542–570, 2013.

[25] L. Sun, B. Zheng, C. Bu, and Y. Wei, "Moore–Penrose inverse of tensors via Einstein product," *Linear Multilinear Algebra*, vol. 64, no. 4, pp. 686–698, 2016.

[26] R. Pan, "Tensor transpose and its properties," 2014. [Online]. Available: arXiv:1411.1503.

[27] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.

[28] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometr.*, vol. 14, no. 3, pp. 229–239, 2000.

[29] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 810–823, Mar. 2000.

[30] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[31] T. Jiang and N. D. Sidiropoulos, "A direct blind receiver for SIMO and MIMO OFDM systems subject to unknown frequency offset and multipath," in *Proc. 4th IEEE Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2003, pp. 358–362.

[32] N. D. Sidiropoulos and R. S. Budampati, "Khatri-Rao space-time codes," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2396–2407, Oct. 2002.

[33] A. L. F. de Almeida, G. Favier, and L. R. Ximenes, "Space-time-frequency (STF) MIMO communication systems with blind receiver based on a generalized PARATUCK2 Model," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1895–1909, Apr. 2013.

[34] K. Liu, J. P. C. L. da Costa, H. C. So, and A. L. F. de Almeida, "Semi-blind receivers for joint symbol and channel estimation in space-time-frequency MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5444–5457, Nov. 2013.

[35] A. L. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Process.*, vol. 87, no. 2, pp. 337–351, 2007.
[36] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Space-time multiplexing codes: A tensor modeling approach," in *Proc. IEEE 7th Workshop Signal Process. Adv. Wireless Commun.*, Jul. 2006, pp. 1–5.
[37] G. Favier and A. L. F. de Almeida, "Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5987–6002, Nov. 2014.
[38] M. N. da Costa, G. Favier, and J. Romano, "Tensor modelling of MIMO communication systems with performance analysis and Kronecker receivers," *Signal Process.*, vol. 145, pp. 304–316, Apr. 2018.
[39] M.-L. Liang, B. Zheng, and R.-J. Zhao, "Tensor inversion and its application to the tensor equations with Einstein product," *Linear Multilinear Algebra*, vol. 67, no. 4, pp. 843–870, 2019.
[40] A. Papoulis, *The Fourier Integral and Its Applications*. New York, NY, USA: McGraw-Hill, 1962.
[41] D. A. Shnidman, "A generalized Nyquist criterion and an optimum linear receiver for a pulse modulation system," *Bell Syst. Tech. J.*, vol. 46, no. 9, pp. 2163–2177, Nov. 1967.
[42] J. W. Smith, "A unified view of synchronous data transmission system design," *Bell Syst. Tech. J.*, vol. 47, no. 3, pp. 273–300, Mar. 1968.
[43] W. van Etten, "An optimum linear receiver for multiple channel digital transmission systems," *IEEE Trans. Commun.*, vol. COM-23, no. 8, pp. 828–834, Aug. 1975.
[44] *Physical Channels and Modulation*, ETSI Standard TS 138 211, 2018.
[45] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, May 2011.
[46] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1170–1183, May 2002.
[47] C. Lélé, P. Siohan, and R. Legouable, "The alamouti scheme with CDMA-OFDM/OQAM," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, Jan. 2010.
[48] N. Michailow *et al.*, "Generalized frequency division multiplexing for 5th generation cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3045–3061, Sep. 2014.
[49] M. Matthe, I. Gaspar, D. Zhang, and G. Fettweis, "Near-ML detection for MIMO-GFDM," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC)*, Sep. 2015, pp. 1–2.
[50] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[51] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

ADITHYA VENUGOPAL received the B.Tech. degree in electronics and communications engineering from the Manipal Institute of Technology, India, in 2016, and the M.S. degree in electrical and computer engineering from McGill University, Montreal, QC, Canada, in 2019. He worked as a Network Engineer with Cisco Systems, Inc., Bengaluru, India, from January 2016 to August 2016. His research interests include digital communications and tensor algebra with applications to communications and signal processing.

HARRY LEIB received the B.Sc. (*cum laude*) and M.Sc. degrees in electrical engineering from the Technion-Israel Institute of Technology, Israel, in 1977 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Toronto, Canada, in 1987.

From 1977 to 1984, he was with the Israel Ministry of Defense, working in communication systems. He was with the University of Toronto as a Postdoctoral Research Associate and as an Assistant Professor. Since September 1989, he has been with the Department of Electrical and Computer Engineering, McGill University, where he is currently a Full Professor. His current research activities are in the areas of digital communications, wireless communication systems, global navigation satellite systems, detection, estimation, and information theory.

Prof. Leib was an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2000 to 2013, and an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2001 to 2007. He was a Guest Co-Editor for special issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATION on "Differential and Noncoherent Wireless Communication" from 2003 to 2005, and on "Spectrum and Energy Efficient Design of Wireless Communication Networks" from 2012 to 2013. Since 2017, he has been the Founding Editor-in-Chief of *AIMS Electronics and Electrical Engineering*.