



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Abordagem Tensorial para Estimação de Canal em Redes MIMO Cooperativas Multi-Salto

Dissertação de Mestrado

Ítalo Vitor Cavalcante

FORTALEZA – CEARÁ
JULHO 2014



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Abordagem Tensorial para Estimação de Canal em Redes MIMO Cooperativas Multi-Salto

Autor

Ítalo Vitor Cavalcante

Orientador

Prof. Dr. André Lima Férrer de Almeida

*Dissertação apresentada à
Coordenação do Programa de
Pós-graduação em Engenharia de
Teleinformática da Universidade
Federal do Ceará como parte dos
requisitos para obtenção do grau
de **Mestre em Engenharia de
Teleinformática.***

FORTALEZA – CEARÁ

JULHO 2014

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca de Pós-Graduação em Engenharia - BPGE

-
- C364a Cavalcante, Ítalo Vitor.
Abordagem tensorial para estimação de canal em Redes MIMO cooperativas multi-salto. / Ítalo Vitor Cavalcante. – 2014.
60 f. : il. color. , enc. ; 30 cm.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2014.
Área de concentração: Sinais e Sistemas.
Orientação: Prof. Dr. André Lima Férrer de Almeida.
1. Teleinformática. 2. Estimação de canal. 3. Métodos de decomposição. I. Título.



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
CAMPUS DO PICI, CAIXA POSTAL 6007 CEP 60.738-640
FORTALEZA – CEARÁ - BRASIL
FONE (+55) 85 3366-9467 – FAX (+55) 85 3366-9468

ÍTALO VITOR CAVALCANTE

**ABORDAGEM TENSORIAL PARA ESTIMAÇÃO DE CANAL EM
REDES MIMO COOPERATIVAS MULTI-SALTO**

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Teleinformática.

Área de concentração: Sinais e Sistemas.

Aprovada em: 18/07/2014.

BANCA EXAMINADORA

Prof. Dr. André Lima Férrer de Almeida (Orientador)

Universidade Federal do Ceará

Prof. Dr. Carlos Alexandre Rolim Fernandes

Universidade Federal do Ceará

Gérard Favier
CNRS Research Director

Prof. Dr. Gérard Favier

Université de Nice Sophia Antipolis

Resumo

Nesta dissertação o problema de estimação de canal em sistemas MIMO cooperativos é investigado. Mais especificamente, foram desenvolvidas técnicas para estimação de canal em um sistema de comunicação assistida por *relays* com processamento do tipo amplifica-e-encaminha (do inglês, *amplify-and-forward*) em um cenário de 3 saltos. As técnicas desenvolvidas utilizam sequência de treinamento e habilitam, no nó receptor, a estimação de todos os canais envolvidos no processo de comunicação.

Em um cenário inicial, consideramos um sistema de comunicação com N antenas transmissoras e M antenas receptoras e entre esses nós temos dois grupos de *relays* com R_1 e R_2 antenas cada um. Foram desenvolvidos protocolos de transmissão baseado em multiplexação temporal para coordenar as retransmissões dos sinais. Ao final da fase de treinamento, o nó receptor faz a estimação das matrizes de canal através da combinação dos dados recebidos. Explorando a estrutura multilinear (tensorial) dos diversos conjuntos de sinais, podemos modelar os dados recebidos através de modelos tensoriais, tais como: PARAFAC e PARATUCK2. Este trabalho propõe a utilização combinada desses modelos e de técnicas algébricas para explorar a diversidade espacial.

Em um segundo momento, consideramos que o número de antenas transmissoras e receptoras dos *relays* podem ser diferentes e ainda que os dados podem trafegar em um esquema bidirecional (do inglês, *two-way*). Para fins de validação dos algoritmos utilizamos simulações de Monte-Carlo em que comparamos os modelos propostos com outros algoritmos de estimação de canal, tais como os algoritmos baseados em PARAFAC e Fatoração de Khatri-Rao em termos de NMSE e taxa de erro de bit.

Palavras-chave: Estimação de canal, decomposições tensoriais, comunicações cooperativas.

Abstract

In this dissertation the problem of channel estimation in cooperative MIMO systems is investigated. More specifically, channel estimation techniques have been developed for a communication system assisted by relays with amplify-and-forward (AF) processing system in a three-hop scenario. The techniques developed use training sequences and enable, at the receiving node, the estimation of all the channels involved in the communication process.

In an initial scenario, we consider a communication system with N transmit antennas and M receive antennas and between these nodes we have two relay groups with R_1 and R_2 antennas each. We propose protocols based on temporal multiplexing to coordinate the retransmission of the signals. At the end of the training phase, the receiving node estimates the channel matrices by combining the received data. By exploiting the multilinear (tensorial) structure of the sets of signals, we can model the received data through tensor models, such as PARAFAC and PARATUCK2. This work proposes the combined use of these models and algebraic techniques to explore the spatial diversity.

Secondly, we consider that the number of transmit and receive antennas at the relays may be different and that the data can travel in a bidirectional scheme (two-way). In order to validate the algorithms we use Monte-Carlo simulations in which we compare our proposed models with competing channel estimation algorithms, such as, the PARAFAC and Khatri-Rao factorization based algorithms in terms of NMSE and bit error rate.

Key-words: Channel estimation, tensor decompositions, cooperative communications.

...aos meus pais Flávio e Lucineide, meus irmãos Ícaro e Igor e a todos os
meus amigos e familiares.

Agradecimentos

Agradeço a Deus, por todas as graças que recebo todos os dias, por sua misericórdia que me enche de esperança e por seu amor que me fortalece, alegra e guia para junto Dele. Aos meus pais, Flávio e Lucineide que me criaram com todo o conforto e carinho que puderam oferecer e por confiar tantas vezes em mim. Aos meus irmãos, Ícaro e Igor por serem os melhores companheiros do mundo.

Ao Prof. Dr. André Lima Férrer de Almeida, pela paciência, atenção e disponibilidade que se dispôs a ter ao longo desses dois anos de mestrado.

Aos meus companheiros do Grupo de Estudos em Telecomunicações Sem Fio (GTEL) e do Grupo de Pesquisa em Processamento de Sinais e Informação (GPSI): Gilderlan Tavares, Luíza Helena, Jordan Paiva, Wilker Lima, Paulo Ricardo, Igor Flávio, Juan Medeiros, Rafael Guimarães, Vanessa Viana, Daniel Matias e Emanuel Sena pelo apoio. A todos os professores do Departamento de Engenharia de Teleinformática da UFC pela minha formação acadêmica e auxílio.

Finalmente, agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo apoio financeiro sem o qual não seria possível a realização deste trabalho.

Sumário

Lista de Figuras	vii
Lista de Tabelas	ix
1 Introdução	1
1.1 Motivação e Objetivos	1
1.2 Estrutura da Dissertação	2
1.3 Contribuições na Forma de Artigos Publicados	3
2 Revisão Bibliográfica	4
2.1 Álgebra Multilinear e Modelos Tensoriais	4
2.1.1 Conceitos Iniciais	4
2.1.2 Decomposições Tensoriais	7
2.2 Comunicações Cooperativas	14
2.2.1 Classificação dos Nós Retransmissores (<i>Relays</i>)	15
2.3 Trabalhos Relacionados	16
2.3.1 Método Baseado em SVD para o Cenário <i>One-Way Two-Hop</i>	16
2.3.2 Método Baseado em PARAFAC para o Cenário <i>One-Way Two-Hop</i>	19
3 Receptores Propostos	21
3.1 Modelo em 4 Fases <i>One-Way</i>	22
3.1.1 Descrição do Protocolo de Transmissão	22
3.1.2 Formulação Tensorial	23
3.1.3 Algoritmo de Estimação de Canal	27
3.1.4 Modelo em 3 Fases <i>One-Way</i>	29
3.1.5 Simulações e Resultados	32
3.2 Extensão para Matrizes de Amplificação Não-diagonais	37
3.2.1 Simulações e Resultados	42

4 Conclusões e Perspectivas	45
Apêndice A Receptor para o Modelo <i>Two-Way</i>	47
A.1 Descrição do Protocolo de Transmissão	47
A.2 Formulação Tensorial	49
A.3 Algoritmo de Estimação de Canal	52
Referências Bibliográficas	57

Lista de Figuras

2.1	Fibras de um tensor de ordem 3. Colunas (fibras modo-1), linhas (fibras modo-2) e tubos (fibras modo-3).	5
2.2	<i>Slices</i> de um tensor de ordem 3. <i>Slices</i> horizontais (modo-1), laterais (modo-2) e frontais (modo-3).	5
2.3	Decomposição <i>Tucker</i> mostrando as matrizes fatores e o tensor núcleo.	11
2.4	Decomposição PARATUCK2 de um tensor \mathcal{X} . Em que os k -ésimos <i>slices</i> frontais de \mathcal{D}^A e \mathcal{D}^B são $D_k(C^{(A)})$ e $D_k(C^{(B)})$ respectivamente.	12
2.5	Sistema cooperativo com um nó retransmissor (<i>relay</i>).	15
2.6	Cenário <i>One-way MIMO Two-Hop</i> com N antenas transmissoras, M antenas receptoras e <i>relay</i> AF com R antenas	17
3.1	Modelo básico da rede cooperativa. N antenas transmissoras M antenas receptoras e dois <i>relays</i> AF.	22
3.2	Tensores que modelam os dados recebidos dos enlaces Fonte- <i>Relay</i> 2-Destino, Fonte- <i>Relay</i> 1-Destino e Fonte- <i>Relay</i> 1- <i>Relay</i> 2-Destino.	26
3.3	NMSE vs. Potência de Transmissão (Comparação entre o algoritmo Comb-ALS-1 e SVD-LS). Para o canal H_{eff} em (3.47)	33
3.4	BER vs. Potência de transmissão dos dados. Potência de transmissão na fase de treinamento = 20dB. $N = M = R_1 = R_2 = 4$)	34
3.5	NMSE vs. Potência de transmissão (Comparação entre o algoritmo Comb-ALS-1.1 e PARAFAC-PARATUCK2). Para o canal H_{eff} em (3.48).	35
3.6	BER vs. SNR. Potência de transmissão = 20dB.	35
3.7	BER vs. Potência de transmissão dos dados (Comparação entre os algoritmos Comb-ALS-1 e Comb-ALS-1.1). Potência de transmissão da fase de treinamento = 20dB.	36
3.8	NMSE vs. Iterações (Comb-ALS-1).	36

3.9	NMSE vs. Iterações (Comb-ALS-1.1).	37
3.10	Tensores de amplificação do <i>Relay 1</i> (\mathcal{W}) e do <i>Relay 2</i> (\mathcal{T}).	39
3.11	NMSE vs. Potência de transmissão para $N = M = K = 4$ (Comparação entre o algoritmo Comb-ALS-2 e SVD-LS).	44
3.12	Desempenho de BER como o receptor ZF em função da potência de transmissão dos dados (Comparação entre Comb-ALS-2 e SVD-LS para $R_{in}^{(1)} = R_{out}^{(1)} = R_{in}^{(2)} = R_{out}^{(2)} = 4$). Potência de transmissão na fase de treinamento = 20dB.	44
A.1	Modelo básico da rede cooperativa. Terminal 1 com N antenas, Terminal 2 com M antenas e dois <i>relays</i> AF.	48

Lista de Tabelas

2.1	Algoritmo PARAFAC-ALS	10
2.2	Algoritmo PARATUCK2-ALS	14
2.3	Algoritmo <i>Least Squares Khatri-Rao Factorization</i>	19
3.1	Algoritmo Comb-ALS-1	29
3.2	Algoritmo Comb-ALS-1.1	31
3.3	Algoritmo SVD-LS	33
3.4	Algoritmo Comb-ALS-2	42
A.1	Algoritmo Comb-ALS-Two-Way	56

Lista de Acrônimos

ALS	Alternating Least Squares
AF	Amplify-and-Forward
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CANDECOMP	Canonical Decomposition
CANDELINC	Canonical Decomposition with Linear Constraints
CDMA	Code Division Multiple Access
CSI	Channel State Information
DEDICOM	Decomposition Into Directional Components
DF	Decode-and-Forward
DFT	Discrete Fourier Transform
INDSCAL	Individual Differences in Scaling
LS	Least Squares
MIMO	Multiple-Input and Multiple-Output
MMSE	Minimum Mean Square Error
NMSE	Normalized Mean Square Error
PARAFAC	Parallel Factor Analysis
QAM	Quadrature Amplitude Modulation
SNR	Signal-to-Noise Ratio

SVD	Singular Value Decomposition
TDMA	Time-Division Multiple Access
ZF	Zero Forcing

Notação

No decorrer deste trabalho as seguintes notações serão utilizadas:

x	Escalares
\mathbf{x}	Vetores
\mathbf{X}	Matrizes
\mathcal{X}	Tensores
$\mathbf{x} \cdot \mathbf{y}$	Produto interno entre \mathbf{x} e \mathbf{y}
$\mathbf{x} \circ \mathbf{y}$	Produto externo entre \mathbf{x} e \mathbf{y}
\mathbb{C}	Conjunto dos números complexos
\mathbb{C}^{I_1}	Conjunto dos vetores de dimensão I_1 com valores complexos
$\mathbb{C}^{I_1 \times I_2}$	Conjunto das matrizes de dimensão $I_1 \times I_2$ com valores complexos
$\mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$	Conjunto dos tensores de dimensão $I_1 \times I_2 \times \dots \times I_N$ com valores complexos
\mathbf{A}^*	Conjugado de \mathbf{A}
\mathbf{A}^T	Transposto de \mathbf{A}
\mathbf{A}^H	Conjugado transposto de \mathbf{A}
\mathbf{A}^\dagger	Pseudo-inversa de Moore-Penrose de \mathbf{A}
$\ \cdot\ _F$	Norma de Frobenius da Matriz ou Tensor
\mathbf{I}_N	Matriz identidade de tamanho $N \times N$
$\mathcal{I}_{3,N}$	Tensor identidade de ordem 3 de tamanho $N \times N \times N$
x_{i_1, i_2}	(i_1, i_2) -ésimo elemento de $\mathbf{X} \in \mathbb{C}^{I_1 \times I_2}$
x_{i_1, i_2, \dots, i_N}	(i_1, i_2, \dots, i_N) -ésimo elemento de $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$
$\mathbf{X}_{i_1 \cdot}(\mathbf{X}_{\cdot i_2})$	i_1 -ésima linha (i_2 -ésima coluna) de $\mathbf{X} \in \mathbb{C}^{I_1 \times I_2}$
$\mathbf{X}_{i_1 \cdot} \in \mathbb{C}^{I_2 \times I_3}$	i_1 -ésimo <i>slice</i> modo-1 de \mathcal{X}
$\mathbf{X}_{\cdot i_2} \in \mathbb{C}^{I_3 \times I_1}$	i_2 -ésimo <i>slice</i> modo-2 de \mathcal{X}
$\mathbf{X}_{\cdot \cdot i_3} \in \mathbb{C}^{I_1 \times I_2}$	i_3 -ésimo <i>slice</i> modo-3 de \mathcal{X}
\times_n	Produto modo- n
\sqcup_i	Concatena matrizes ao longo do modo i
$[\mathcal{X}]_{(n)}$	Matriciação modo- n de \mathcal{X}
$\mathbf{A} \otimes \mathbf{B}$	Produto de Kronecker entre as matrizes \mathbf{A} e \mathbf{B}
$\mathbf{A} \diamond \mathbf{B}$	Produto de Khatri-Rao entre as matrizes \mathbf{A} e \mathbf{B}
$\text{vec}(\mathbf{X})$	Operador vetorização. Converte \mathbf{X} em um vetor coluna \mathbf{x} empilhando suas colunas
$\text{unvec}(\mathbf{X})$	Operação inversa da vetorização. Converte o vetor \mathbf{x} de volta para uma matriz
$\text{vecd}(\mathbf{X})$	Converte os elementos da diagonal principal de \mathbf{X} num vetor
$D_j(\mathbf{A})$	Representa uma matriz diagonal construída a partir da j -ésima linha de \mathbf{A}
$\text{diag}(\mathbf{x})$	Forma uma matriz diagonal a partir do vetor \mathbf{x}
k_A	k -rank da matriz \mathbf{A}

Introdução

1.1 Motivação e Objetivos

Nos últimos anos o uso de sistemas com múltiplas antenas (sistemas **MIMO**, do inglês, *Multiple-Input and Multiple-Output*) têm sido utilizados extensivamente como um mecanismo para o aumento da capacidade de transmissão dos sistemas de comunicação sem fio devido ao ganho de diversidade e de multiplexação que esta tecnologia propicia. Entretanto, é fato notável que para que haja ganhos máximos nos sistemas **MIMO** as matrizes de canal devem possuir *rank* completo [1]. Em sistemas de comunicação muitas vezes essa condição de *rank* das matrizes de canal não é satisfeita devido a alta correlação entre os diversos enlaces antena-antena. Neste contexto, um novo paradigma cooperativo tem sido usado como solução alternativa para prover ganhos de diversidade/multiplexação bem como aumentar a área de cobertura da rede e **SNR** nos receptores [2], [3], [4].

Em sistemas cooperativos práticos, um ou mais nós retransmissores (*relays*) são adicionados à rede provendo enlaces alternativos ao enlace direto fonte-destino. Neste contexto, técnicas que usam *relays* do tipo amplifica-e-encaminha (**AF**, do inglês *Amplify-and-Forward*) são soluções atrativas quando a decodificação dos dados nos *relays* quer ser evitada ou quando se quer limitar a latência do sistema [5], [6].

Em sistemas de comunicação sem fio, a detecção dos sinais nos receptores depende de maneira crucial da informação de estado do canal (**CSI**, do inglês *Channel State Information*) de cada enlace envolvido no processo de comunicação. Além disso, o uso de técnicas de pré-codificação na fonte e/ou destino [7], [8] frequentemente requer uma estimação de **CSI** instantânea a fim de melhorar a taxa de erro de bits (**BER**, do inglês *Bit Error Rate*) ou a taxa de transmissão. Em [9], [10] um algoritmo baseado na decomposição em valores singulares (**SVD**, do inglês *Singular Value Decomposition*) foi

proposto para a estimação das matrizes de canal envolvidas em um cenário de comunicação cooperativa de dois saltos (do inglês, *two-hop*). Alguns outros trabalhos fazem uso de técnicas baseadas em tensores para solucionar o problema de estimação de canal. Em [11], um esquema de estimação de canal baseada em modelagem tensorial foi proposto para um sistema de comunicação bidirecional (do inglês, *two-way*) com múltiplas antenas nos *relays*. Em [12], receptores para um sistema multiusuário baseado em um modelo trilinear foram propostos, mais especificamente, para o *uplink* de um sistema de comunicação com diversidade cooperativa empregando um *array* de antenas no nó receptor (estação base). Um trabalho mais recente [13] propõe um receptor cego para o *uplink* de um sistema cooperativo multiusuário empregando múltiplo acesso por divisão de código (CDMA, do inglês *Code Division Multiple Access*). Todos estes trabalhos consideram um modelo *two-hop* com *relays*, e o problema de estimação de canal consiste em estimar as matrizes de canal de ambos os saltos. Para melhorar a cobertura e mitigar a perda de percurso e sombreamento, pode ser vantajoso introduzir mais saltos explorando múltiplos enlaces, sempre que possível [1].

Neste trabalho, novos métodos são propostos para fazer a estimação conjunta das matrizes de canal de um sistema de comunicação multi-relay MIMO. Considerando cenários com três saltos (do inglês, *three-hop*) e utilizando protocolos de transmissão baseados em TDMA (do inglês, *Time-Division Multiple Access*), estes métodos exploram de maneira rica a estrutura algébrica (tensorial) de todos os enlaces MIMO disponíveis, combinando os modelos PARAFAC (do inglês, *Parallel Factor Analysis*) [14], PARATUCK2 e Tucker [15] para estimar iterativamente as matrizes de canal por meio de algoritmos do tipo mínimos quadrados alternado (ALS, do inglês *Alternating Least Squares*). Os métodos propostos proveem soluções efetivas para o problema de estimação de canal devido ao uso eficiente da diversidade cooperativa.

1.2 Estrutura da Dissertação

Este trabalho está organizado da seguinte forma:

Capítulo 2 – Neste capítulo serão expostos os conceitos básicos para a compreensão do restante do texto. Serão apresentados de forma breve alguns conceitos chave sobre álgebra multilinear que serão recorrentes ao longo do texto. Em seguida, será feita uma breve introdução aos sistemas cooperativos. Finalmente, serão apresentados os modelos matemáticos de alguns trabalhos anteriores que serviram de base para o

desenvolvimento desta dissertação bem como parâmetros de comparação e análise de desempenho dos algoritmos desenvolvidos aqui.

Capítulo 3 – Neste capítulo, será apresentada a modelagem de sinal (considerações sobre perda de percurso, modelagem dos sinais em banda base, número de antenas nos nós da rede, etc) do modelo em estudo. Também serão descritos os protocolos de transmissão e algoritmos de estimação de canal para o caso *one-way* deste cenário. Em seguida serão descritas extensões deste modelo para o caso em que o número de antenas receptoras e transmissoras nos *relays* são diferentes. Ao final de cada seção, faremos uma discussão dos resultados obtidos incluindo comparações com outras abordagens na literatura (descritos no capítulo 2) através de simulações computacionais a fim de validar a eficiência dos métodos propostos em termos de BER e erro médio quadrático normalizado (NMSE, do inglês, *Normalized Mean Square Error*) dos canais MIMO estimados.

Capítulo 4 – Neste capítulo faremos conclusões sobre os resultados obtidos no Capítulo 3 e também abordaremos perspectivas para trabalhos futuros.

Apêndice A – Neste apêndice será descrita uma extensão do algoritmo de estimação de canal para o cenário em estudo onde os dados trafegam de forma bidirecional (do inglês, *two-way*).

1.3 Contribuições na Forma de Artigos Publicados

Tivemos a oportunidade de publicar um artigo em um evento sobre *array* de sensores e de submeter um trabalho em um periódico. Os trabalhos são os seguintes:

► Congresso

Ítalo Vitor Cavalcante, A. L. F. de Almeida, M. Haardt, *Tensor-Based Approach to Channel Estimation in Amplify-and-Forward MIMO Relaying Systems*. Aceito no VIII IEEE Sensor Array and Multichannel Signal Processing Workshop - SAM 2014, 22-25 de junho, Corunha, Espanha.

► Periódico

Ítalo Vitor Cavalcante, A. L. F. de Almeida, M. Haardt, *Channel Estimation for One-Way Three-Hop MIMO Multi-Relay Systems*. Submetido em IEEE Signal Processing Letters.

Revisão Bibliográfica

Neste capítulo serão introduzidos alguns conceitos de álgebra multilinear e modelagem tensorial que serão utilizados extensivamente ao longo do texto. Em seguida, será feita uma breve revisão sobre sistemas cooperativos e, finalmente, serão expostos alguns trabalhos que serviram como ponto de partida para este trabalho.

2.1 Álgebra Multilinear e Modelos Tensoriais

Nesta seção serão introduzidos alguns conceitos chave sobre álgebra multilinear bem como os modelos tensoriais utilizados nesta dissertação. Primeiramente, serão descritas as principais notações, definições e operações matemáticas envolvendo tensores. Em seguida, serão descritas as 3 decomposições tensoriais das quais faremos uso no decorrer do capítulo 3: [PARAFAC](#), [PARATUCK2](#), e *Tucker*.

2.1.1 Conceitos Iniciais

Tensores são adotados nos mais diversos ramos do conhecimento: processamento de imagens, bioinformática, sistemas de comunicação sem fio e muitas outras áreas em que o conjunto de dados em estudo ou as características que se quer analisar possuem natureza multilinear. Para os propósitos desta dissertação, um tensor pode ser visto como uma extensão multilinear do conceito de matriz, ou seja, um *array* numérico multi-dimensional [16], [17], [18]. A seguir temos algumas definições importantes das quais faremos uso recorrente ao longo desta dissertação.

Definição 1. O (i_1, i_2, \dots, i_N) -ésimo elemento do tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ de ordem N é dado por

$$x_{i_1, i_2, \dots, i_N} = [\mathcal{X}]_{i_1, i_2, \dots, i_N}. \quad (2.1)$$

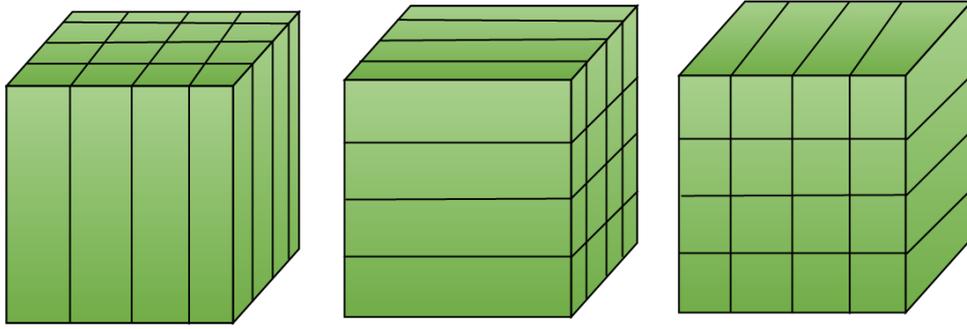


Figura 2.1: Fibras de um tensor de ordem 3. Colunas (fibras modo-1), linhas (fibras modo-2) e tubos (fibras modo-3).

Definição 2. Uma fibra modo- n é uma secção unidimensional (vetor) de um tensor. Uma fibra é obtida fixando-se todos os índices de um tensor com exceção de um (Figura 2.1)

$$\mathbf{x}_{\dots i_j \dots} = [\mathcal{X}]_{\dots i_j \dots} \quad (2.2)$$

Uma fibra modo- n possui dimensão $I_n \times 1$. Em que I_n é a dimensão correspondente ao n -ésimo modo.

Definição 3. Um *slice* é uma secção bidimensional (matriz) de um tensor de ordem 3 (ou terceira ordem). Neste caso, um *slice* é obtido fixando-se todos os índices de um tensor com exceção de dois (Figura 2.2).

$$\mathbf{X}_{\dots i_j \dots i_k \dots} = [\mathcal{X}]_{\dots i_j \dots i_k \dots} \quad (2.3)$$

Definição 4. Um tensor de ordem N $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ é dito de *rank-1* se ele poder ser escrito como o produto externo de N vetores, i. e.,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \dots \circ \mathbf{a}^{(N)}, \quad (2.4)$$

em que $x_{i_1, i_2, \dots, i_N} = a_{i_1}^{(1)} a_{i_2}^{(2)} \dots a_{i_N}^{(N)}$, ou seja, cada elemento do tensor é o produto

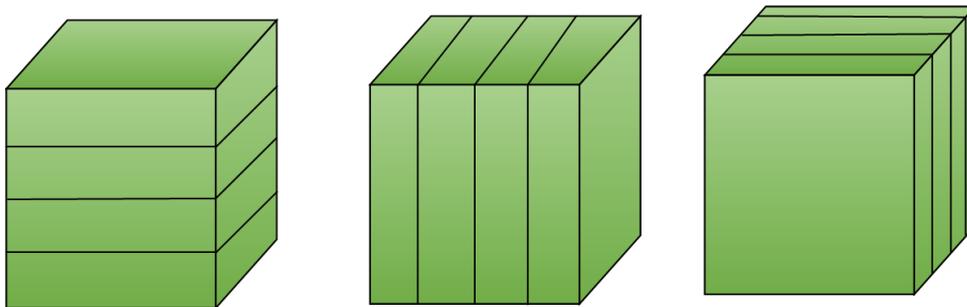


Figura 2.2: Slices de um tensor de ordem 3. Slices horizontais (modo-1), laterais (modo-2) e frontais (modo-3).

dos elementos correspondentes nos vetores fatores.

Definição 5. *Rank* de um tensor \mathcal{X} é o número mínimo de tensores de *rank*-1 necessários para gerar \mathcal{X} em uma combinação linear. Por exemplo, seja \mathcal{X} um tensor de ordem 3 que possui *rank* R , R é o menor inteiro tal que:

$$\mathcal{X} = \sum_{r=1}^R \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r. \quad (2.5)$$

Definição 6. Um tensor identidade $\mathcal{I} \in \mathbb{C}^{N \times N \times \dots \times N}$ é aquele em que todos os seus elementos são iguais a zero, exceto aqueles situados na superdiagonal (elementos em que $i_1 = i_2 = \dots = i_N$, para $i_n = 1, \dots, N$), a qual possui todos os seus elementos iguais a 1. Ou seja, para um tensor identidade, temos:

$$x_{i_1, i_2, \dots, i_N} = \begin{cases} 1, & \text{se } i_1 = i_2 = \dots = i_N \\ 0, & \text{caso contrário.} \end{cases} \quad (2.6)$$

Definição 7. A forma matriciada ou *unfolding* modo- n de um tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ de ordem N denotada por $[\mathcal{X}]_{(n)}$ consiste no rearranjo das fibras modo- n deste tensor nas colunas de uma matriz. Como cada fibra modo- n possui dimensão $I_n \times 1$, a forma *unfolding* modo- n do tensor terá dimensão $I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N$.

Matematicamente, o processo de matriciação mapeia o (i_1, i_2, \dots, i_n) -ésimo elemento do tensor no (i_n, j) -ésimo elemento de uma matriz, em que

$$j = 1 + \sum_{k=1, k \neq n}^N (i_k - 1) J_k, \quad \text{com } J_k = \prod_{m=1, m \neq n}^{k-1} I_m. \quad (2.7)$$

Definição 8. Seja $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ um tensor de ordem N . O produto modo- n de \mathcal{X} por uma matriz $\mathbf{A} \in \mathbb{C}^{J \times I_n}$, denotado por $\mathcal{X} \times_n \mathbf{A}$, é um tensor \mathcal{Y} também de ordem N , mas com dimensões $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$. Cada fibra modo- n de \mathcal{Y} consiste da correspondente fibra modo- n de \mathcal{X} pré-multiplicada por \mathbf{A} . Podemos expressar o produto modo- n em termos das formas *unfolding*

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{A} \Leftrightarrow [\mathcal{Y}]_{(n)} = \mathbf{A} [\mathcal{X}]_{(n)}. \quad (2.8)$$

Observe que o produto modo- n pressupõe compatibilidade entre as dimensões da matriz e a forma *unfolding* modo- n do tensor. Caso queiramos realizar o produto do tensor (2.8) por outra matriz \mathbf{B} no mesmo modo, basta

pré-multiplicar a matriz A por B e realizar o produto modo- n do tensor \mathcal{X} pela matriz resultante, ou seja,

$$\mathcal{Z} = \mathcal{X} \times_n A \times_n B = \mathcal{X} \times_n (BA) \Leftrightarrow [\mathcal{Z}]_{(n)} = (BA)[\mathcal{X}]_{(n)}.$$

Definição 9. O produto de Kronecker entre duas matrizes $A \in \mathbb{C}^{I \times J}$ e $B \in \mathbb{C}^{K \times L}$ é denotado por $A \otimes B$. A matriz resultante tem dimensões $IK \times JL$ e é definida como

$$A \otimes B = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \cdots & a_{IJ}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{IK \times JL}. \quad (2.9)$$

Definição 10. O produto de Khatri-Rao entre duas matrizes $A \in \mathbb{C}^{I \times K}$ e $B \in \mathbb{C}^{J \times K}$ é denotado por $A \diamond B$. A matriz resultante tem dimensões $IJ \times K$ e é definida como

$$A \diamond B = \begin{bmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \cdots & \mathbf{a}_K \otimes \mathbf{b}_K \end{bmatrix} \in \mathbb{C}^{IJ \times K}, \quad (2.10)$$

ou seja, o produto de Khatri-Rao consiste em se fazer o produto de Kronecker coluna a coluna.

Definição 11. Definimos k -rank de uma matriz $A \in \mathbb{C}^{I \times R}$, denotado por k_A , como o número máximo k tal que qualquer conjunto de k colunas de A é linearmente independente. O k -rank de uma matriz é sempre limitado superiormente pelo seu $rank$, i.e., $k_A \leq r \leq \min(I, R)$, em que r é o $rank$ de A .

Definição 12. Uma matriz $A \in \mathbb{C}^{I \times R}$ tem k -rank completo se $k_A = r = \min(I, R)$. Note que, se $A \in \mathbb{C}^{I \times R}$ contiver um par de colunas proporcionais, então temos $k_A = 1$.

2.1.2 Decomposições Tensoriais

As primeiras contribuições para o estudo de dados multilineares datam do século 19 quando matemáticos proeminentes tais como Gauss, Cayley e Hilbert estudavam polinômios homogêneos [19]. Em 1927 com o trabalho de Hitchcock [20] deu-se início o estudo moderno de tensores. A decomposição *Tucker* para tensores foi desenvolvida no contexto da psicometria [21], [15]. A decomposição **PARAFAC** foi independentemente descoberta quando

se faziam estudos sobre linguística [14] e estudos sobre psicometria [22] (aqui com o nome de **CANDECOMP**, do inglês, *Canonical Decomposition*). As decomposições tensoriais apresentam grande vantagem quando se está trabalhando com modelos de dados inerentemente multilineares. Por exemplo, em sistemas de comunicação que usam diversidades espacial, temporal e frequencial. Neste caso, o sinal transmitido (ou recebido) pode ser mapeado em um tensor de terceira ordem, cujos modos estão associados aos domínios "espaço", "tempo" e "frequência", nos quais a informação é veiculada. Além disso, diversas propriedades algébricas dos tensores, tais como a unicidade [23], tornam o seu uso vantajoso em relação aos métodos matriciais clássicos. A seguir as decomposições **PARAFAC**, **PARATUCK2** e *Tucker* serão apresentadas. As mesmas serão utilizadas nas descrições dos algoritmos propostos ao longo do Capítulo 3.

2.1.2.1 Decomposição PARAFAC

A ideia de se expressar um tensor como uma soma finita de tensores de *rank*-1 vem de Hitchcock [20], [24], e posteriormente foi aprimorada por Cattell [25], [26]. A decomposição **CANDECOMP/PARAFAC** (do inglês, *Canonical Decomposition* e *Parallel Factor Analysis*, respectivamente) foi desenvolvida independentemente por Carroll [22] e Harshman [14]. Nesta dissertação ela será referida como decomposição **PARAFAC**. Para um tensor de ordem 3 a decomposição **PARAFAC** expressa seus elementos em termos de suas componentes trilineares (para os fins desta dissertação iremos trabalhar com tensores de ordem 3, para ordem superiores o tensor seria expresso em termos de 4 ou mais componentes), ou seja, o (i, j, k) -ésimo elemento de um tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ é expresso como

$$x_{i,j,k} = \sum_{r=1}^R a_{ir} b_{jr} c_{kr}, \text{ para } i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K, \quad (2.11)$$

em que $a_{ir} = [\mathbf{A}]_{ir}$, $b_{jr} = [\mathbf{B}]_{jr}$ e $c_{kr} = [\mathbf{C}]_{kr}$, são os elementos da r -ésima coluna e linhas i , j e k das matrizes fatores $\mathbf{A} \in \mathbb{C}^{I \times R}$, $\mathbf{B} \in \mathbb{C}^{J \times R}$ e $\mathbf{C} \in \mathbb{C}^{K \times R}$, respectivamente. Dizemos que \mathbf{A} , \mathbf{B} e \mathbf{C} são matrizes fatores pois podemos expressar o tensor \mathcal{X} em termos de um somatório de tensores de *rank*-1 que por sua vez são expressos em termos do produto externo das colunas de \mathbf{A} , \mathbf{B} e \mathbf{C} , i.e.:

$$\mathcal{X} = \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r. \quad (2.12)$$

Podemos expressar os *slices* de um tensor em termos de suas matrizes fatores [16]. Os slices horizontais, laterais e frontais são dados por [16]

$$\mathbf{X}_{i..} = \mathbf{B}D_i(\mathbf{A})\mathbf{C}^T, i = 1, \dots, I, \quad (2.13a)$$

$$\mathbf{X}_{.j.} = \mathbf{C}D_j(\mathbf{B})\mathbf{A}^T, j = 1, \dots, J, \quad (2.13b)$$

$$\mathbf{X}_{..k} = \mathbf{A}D_k(\mathbf{C})\mathbf{B}^T, k = 1, \dots, K, \quad (2.13c)$$

em que $D_l(\mathbf{R})$ constrói uma matriz diagonal a partir da l -ésima linha de \mathbf{R} . Uma outra forma de representar \mathcal{X} é através da notação produto modo- n (**Definição 8**), que para um tensor de ordem 3 pode ser expressa como [16]:

$$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}, \quad (2.14)$$

em que $\mathcal{I}_{3,R}$ é o tensor identidade de tamanho $R \times R \times R$.

As formas *unfolding* (**Definição 7**) nada mais são que a concatenação horizontal dos *slices* em (2.13). Podemos expressar as formar *unfolding* modo-1, modo-2 e modo-3 de um tensor \mathcal{X} como

$$\begin{aligned} [\mathcal{X}]_{(1)} &= [\mathbf{X}_{..1} \cdots \mathbf{X}_{..K}] = \mathbf{A} [D_1(\mathbf{C})\mathbf{B}^T \cdots D_K(\mathbf{C})\mathbf{B}^T] = \\ &= \mathbf{A}(\mathbf{C} \diamond \mathbf{B})^T \in \mathbb{C}^{I \times JK}, \end{aligned} \quad (2.15)$$

$$\begin{aligned} [\mathcal{X}]_{(2)} &= [\mathbf{X}_{.1.}^T \cdots \mathbf{X}_{.K.}^T] = \mathbf{B} [D_1(\mathbf{C})\mathbf{A}^T \cdots D_K(\mathbf{C})\mathbf{A}^T] = \\ &= \mathbf{B}(\mathbf{C} \diamond \mathbf{A})^T \in \mathbb{C}^{J \times IK}, \end{aligned} \quad (2.16)$$

$$\begin{aligned} [\mathcal{X}]_{(3)} &= [\mathbf{X}_{.1.} \cdots \mathbf{X}_{.J.}] = \mathbf{C} [D_1(\mathbf{B})\mathbf{A}^T \cdots D_J(\mathbf{B})\mathbf{A}^T] = \\ &= \mathbf{C}(\mathbf{B} \diamond \mathbf{A})^T \in \mathbb{C}^{K \times IJ}. \end{aligned} \quad (2.17)$$

Através das equações (2.15-2.17) podemos construir um algoritmo para estimação das matrizes fatores \mathbf{A} , \mathbf{B} e \mathbf{C} . Este algoritmo é conhecido como **PARAFAC-ALS** e está resumido na Tabela 2.1

Um resultado devido a Kruskal [23] estabelece que as matrizes \mathbf{A} , \mathbf{B} e \mathbf{C} são únicas e podem ser estimadas a menos de um fator de escala e permutação caso a seguinte inequação seja satisfeita:

$$k_A + k_B + k_C \geq R + 2, \quad (2.18)$$

em que k_A , k_B e k_C são os k -ranks de \mathbf{A} , \mathbf{B} e \mathbf{C} respectivamente e R é o *rank* do tensor.

Tabela 2.1: Algoritmo PARAFAC-ALS

Inicialização: Atribua $i = 0$; Inicialize aleatoriamente $\hat{A}^{(i=0)}$, $\hat{B}^{(i=0)}$ e $\hat{C}^{(i=0)}$

- 1) Atribua $i = i + 1$.
- 2) Estime $\hat{A}^{(i)}$ usando $\hat{B}^{(i-1)}$ e $\hat{C}^{(i-1)}$ através da equação:

$$\hat{A}^{(i)} = [\mathcal{X}]_{(1)} [(\hat{C}^{(i-1)} \diamond \hat{B}^{(i-1)})^T]^\dagger$$
- 3) Estime $\hat{B}^{(i)}$ usando $\hat{C}^{(i-1)}$ e $\hat{A}^{(i)}$ através da equação:

$$\hat{B}^{(i)} = [\mathcal{X}]_{(2)} [(\hat{C}^{(i-1)} \diamond \hat{A}^{(i)})^T]^\dagger$$
- 4) Estime $\hat{C}^{(i)}$ usando $\hat{B}^{(i)}$ e $\hat{A}^{(i)}$ através da equação:

$$\hat{C}^{(i)} = [\mathcal{X}]_{(3)} [(\hat{B}^{(i)} \diamond \hat{A}^{(i)})^T]^\dagger$$
- 5) Repita os passos 1 a 4 até a convergência.

2.1.2.2 Decomposição Tucker

A decomposição *Tucker* foi introduzida por L. R. Tucker em 1963 [27], mas um artigo de 1966 [15] geralmente é o mais citado na literatura. A decomposição *Tucker* decompõe um tensor em um tensor núcleo multiplicado por uma matriz ao longo de cada modo. Assim, para um tensor de ordem 3 $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, temos

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} \mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r, \quad (2.19)$$

em que $\mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{B} \in \mathbb{C}^{J \times Q}$ e $\mathbf{C} \in \mathbb{C}^{K \times R}$ são as matrizes fatores. Observe que em (2.19) temos um fator de ponderação g_{pqr} para cada componente *rank-1* de \mathcal{X} . Este fator é o (p, q, r) -ésimo elemento do tensor núcleo \mathcal{G} .

Podemos expressar o (i, j, k) -ésimo elemento de \mathcal{X} da seguinte forma

$$x_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_{ip} b_{jq} c_{kr}. \quad (2.20)$$

Observe que se P, Q, R são menores que I, J, K podemos pensar no tensor \mathcal{G} como uma forma comprimida do tensor \mathcal{X} original. A Figura 2.3 ilustra a decomposição *Tucker*.

A decomposição **PARAFAC** pode ser vista como um caso especial da decomposição *Tucker* em que o tensor núcleo é o tensor identidade e $P = Q = R$.

As formas *unfolding* do tensor \mathcal{X} são dadas por

$$[\mathcal{X}]_{(1)} = \mathbf{A}[\mathcal{G}]_{(1)}(\mathbf{C} \otimes \mathbf{B})^T \in \mathbb{C}^{I \times JK}, \quad (2.21)$$

$$[\mathcal{X}]_{(2)} = \mathbf{B}[\mathcal{G}]_{(2)}(\mathbf{C} \otimes \mathbf{A})^T \in \mathbb{C}^{J \times IK}, \quad (2.22)$$

$$[\mathcal{X}]_{(3)} = \mathbf{C}[\mathcal{G}]_{(3)}(\mathbf{B} \otimes \mathbf{A})^T \in \mathbb{C}^{K \times IJ}. \quad (2.23)$$

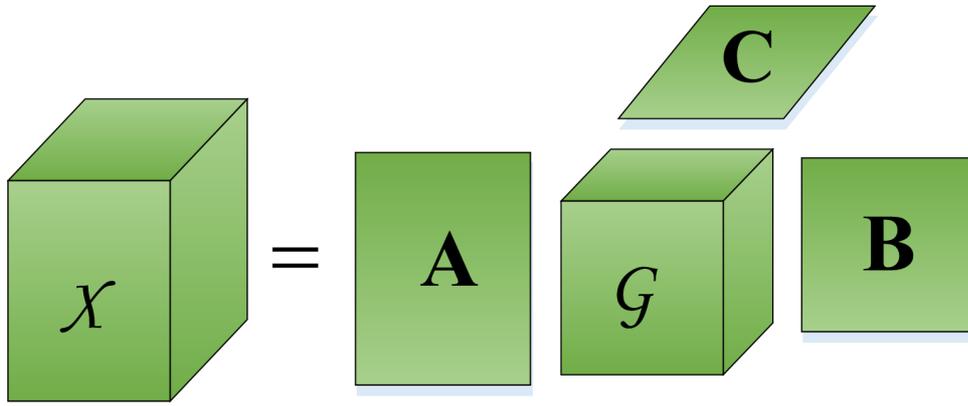


Figura 2.3: Decomposição *Tucker* mostrando as matrizes fatores e o tensor núcleo.

A decomposição *Tucker* como em (2.18) também é chamada decomposição *Tucker3*. Alguns casos especiais surgem quando uma das matrizes fatores é igual a identidade. Por exemplo, quando $C = I$ temos a decomposição *Tucker2* com a seguinte representação em produto modo- n

$$\mathcal{X} = \mathcal{G} \times_1 A \times_2 B. \quad (2.24)$$

As formas *unfolding* de um tensor na forma (2.23) são dadas pelas equações (2.21-2.23) substituindo C por I .

2.1.2.3 Decomposição PARATUCK2

Muitos modelos tensoriais foram desenvolvidos como casos especiais dos modelos *PARAFAC* e *Tucker* [16]. O modelo *INDSCAL* (do inglês, Individual Differences in Scaling) [22] é um caso especial do modelo *PARAFAC* para tensores que são simétricos em dois modos. Outro derivado do modelo *PARAFAC* é o modelo *CANDELINC* (do inglês, *Canonical Decomposition with Linear Constraints*) [28] que pode ser visto como um modelo *PARAFAC* com restrições nas matrizes fatores. Em um trabalho de Harshman [29] foi introduzido o modelo *DEDICOM* (do inglês, *Decomposition Into Directional Components*). Este modelo busca basicamente descrever relações assimétricas entre fatores associados à diferentes matrizes da decomposição. O modelo *PARATUCK2* foi introduzido por Harshman e Lundy [30] e é uma generalização do modelo *DEDICOM*, e pode ser considerado um modelo híbrido entre os modelos *PARAFAC* e *Tucker2*. Dado um tensor de ordem 3 $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ sua decomposição *PARATUCK2* é definida na forma escalar por [30]

$$x_{i,j,k} = \sum_{r=1}^R \sum_{m=1}^M a_{i,r} b_{j,m} g_{r,m} c_{k,r}^{(A)} c_{k,m}^{(B)}, \quad (2.25)$$

em que $x_{i,j,k}$ é o (i, j, k) -ésimo elemento de \mathcal{X} , $A \in \mathbb{C}^{I \times R}$, $B \in \mathbb{C}^{J \times M}$, $C^{(A)} \in \mathbb{C}^{K \times R}$,

$C^{(B)} \in \mathbb{C}^{K \times M}$ e $G \in \mathbb{C}^{R \times M}$. As matrizes A e B estão associadas com a primeira e segunda dimensões de \mathcal{X} . As matrizes $C^{(A)}$ e $C^{(B)}$ definem o padrão de interação entre as R colunas de A e as M colunas de B ao longo da terceira dimensão. A matriz G é chamada matriz núcleo, cujo $g_{r,m}$ elemento define o peso da interação entre a r -ésima coluna de A e a m -ésima coluna de B .

O k -ésimo *slice* frontal de um tensor PARATUCK2 é expresso por [16]

$$\mathbf{X}_{..k} = \mathbf{A} D_k(\mathbf{C}^{(A)}) \mathbf{G} D_k(\mathbf{C}^{(B)}) \mathbf{B}^T, k = 1, \dots, K, \quad (2.26)$$

em que $D_K(\mathbf{C}^{(A)})$ e $D_K(\mathbf{C}^{(B)})$ cria matrizes diagonais a partir das k -ésima linhas de $\mathbf{C}^{(A)}$ e $\mathbf{C}^{(B)}$ respectivamente. A representação PARATUCK2 é mostrada na Figura 2.4.

A seguir, iremos descrever um algoritmo ALS para a estimação das matrizes fatores A , G e B (PARATUCK2-ALS), supondo que as matrizes $\mathbf{C}^{(A)}$ e $\mathbf{C}^{(B)}$ são conhecidas. Pode-se provar que com estas restrições (matrizes $\mathbf{C}^{(A)}$ e $\mathbf{C}^{(B)}$ conhecidas) as matrizes A , G e B podem ser estimadas a menos de um fator de escala [31].

Introduzamos

$$\mathbf{x}_k \doteq \text{vec}(\mathbf{X}_{..k}) \in \mathbb{C}^{IJ \times 1}, k = 1, \dots, K, \quad (2.27)$$

$$\mathbf{X} \doteq [\mathbf{x}_1, \dots, \mathbf{x}_K] \in \mathbb{C}^{IJ \times K}, \quad (2.28)$$

ou seja, cada coluna de \mathbf{X} é a versão vetorizada do k -ésimo *slice* frontal de \mathcal{X} . Considere a seguinte propriedade do produto de Kronecker: $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$. Aplicando esta propriedade em (2.26), \mathbf{x}_k pode ser expresso da seguinte forma

$$\mathbf{x}_k = (\mathbf{B} \otimes \mathbf{A})\text{vec}(D_k(\mathbf{C}^{(A)})\mathbf{G}D_k(\mathbf{C}^{(B)})). \quad (2.29)$$

Aplicando mais uma vez esta propriedade em (2.28), ficamos com

$$\mathbf{x}_k = (\mathbf{B} \otimes \mathbf{A})(D_k(\mathbf{C}^{(B)}) \otimes D_k(\mathbf{C}^{(A)}))\text{vec}(\mathbf{G}). \quad (2.30)$$

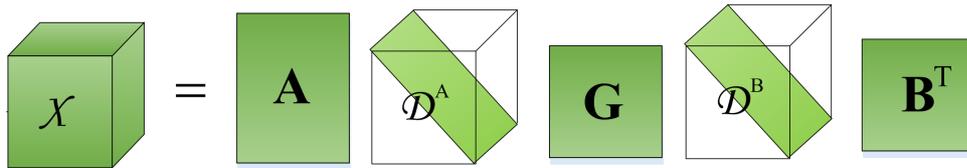


Figura 2.4: Decomposição PARATUCK2 de um tensor \mathcal{X} . Em que os k -ésimos *slices* frontais de \mathcal{D}^A e \mathcal{D}^B são $D_k(\mathbf{C}^{(A)})$ e $D_k(\mathbf{C}^{(B)})$ respectivamente.

Como $\text{diag}(\mathbf{a})\mathbf{b} = \text{diag}(\mathbf{b})\mathbf{a}$, podemos reescrever (2.30) na forma

$$\mathbf{x}_k = (\mathbf{B} \otimes \mathbf{A})\text{diag}(\text{vec}(\mathbf{G})) \left(\mathbf{C}_{(k,\cdot)}^{(\mathbf{B})^T} \otimes \mathbf{C}_{(k,\cdot)}^{(\mathbf{A})^T} \right). \quad (2.31)$$

Desta maneira podemos escrever \mathbf{X} da seguinte forma

$$\mathbf{X} = (\mathbf{B} \otimes \mathbf{A})\text{diag}(\text{vec}(\mathbf{G})) \left(\mathbf{C}^{(\mathbf{B})^T} \diamond \mathbf{C}^{(\mathbf{A})^T} \right). \quad (2.32)$$

Finalmente vetorizando a expressão (2.32) e aplicando a propriedade $\text{vec}(\text{A} \text{diag}(\mathbf{x}) \mathbf{B}) = (\mathbf{B}^T \diamond \mathbf{A})\mathbf{x}$, chegamos a

$$\text{vec}(\mathbf{X}) = \left((\mathbf{C}^{(\mathbf{B})^T} \diamond \mathbf{C}^{(\mathbf{A})^T})^T \diamond (\mathbf{B} \otimes \mathbf{A}) \right) \text{vec}(\mathbf{G}). \quad (2.33)$$

Esta equação será usada para estimar \mathbf{G} . Para estimar \mathbf{A} e \mathbf{B} , definamos

$$\mathbf{Y} \doteq \begin{bmatrix} \mathbf{X}_{\cdot,1} \\ \vdots \\ \mathbf{X}_{\cdot,K} \end{bmatrix} \in \mathbb{C}^{IK \times J}, \quad \mathbf{W} \doteq \begin{bmatrix} \mathbf{X}_{\cdot,1}^T \\ \vdots \\ \mathbf{X}_{\cdot,K}^T \end{bmatrix} \in \mathbb{C}^{JK \times I}, \quad (2.34)$$

e

$$\mathbf{F}_k = D_k(\mathbf{C}^{(\mathbf{A})}) \mathbf{G} D_k(\mathbf{C}^{(\mathbf{B})}), k = 1, \dots, K. \quad (2.35)$$

De (2.26) segue que

$$\begin{aligned} \mathbf{Y} &= \begin{bmatrix} \mathbf{A}\mathbf{F}_1 \\ \vdots \\ \mathbf{A}\mathbf{F}_K \end{bmatrix} \mathbf{B}^T = \begin{bmatrix} \mathbf{A} & & \\ & \ddots & \\ & & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_K \end{bmatrix} \mathbf{B}^T \\ &= (\mathbf{I} \otimes \mathbf{A}) \mathbf{F}_y \mathbf{B}^T, \end{aligned} \quad (2.36)$$

$$\begin{aligned} \mathbf{W} &= \begin{bmatrix} \mathbf{B}\mathbf{F}_1^T \\ \vdots \\ \mathbf{B}\mathbf{F}_K^T \end{bmatrix} \mathbf{A}^T = \begin{bmatrix} \mathbf{B} & & \\ & \ddots & \\ & & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{F}_1^T \\ \vdots \\ \mathbf{F}_K^T \end{bmatrix} \mathbf{A}^T \\ &= (\mathbf{I} \otimes \mathbf{B}) \mathbf{F}_w \mathbf{A}^T, \end{aligned} \quad (2.37)$$

em que

$$\mathbf{F}_y = \begin{bmatrix} \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_K \end{bmatrix} \in \mathbb{C}^{RK \times M}, \quad \mathbf{F}_w = \begin{bmatrix} \mathbf{F}_1^T \\ \vdots \\ \mathbf{F}_K^T \end{bmatrix} \in \mathbb{C}^{MK \times R}. \quad (2.38)$$

Tabela 2.2: Algoritmo PARATUCK2-ALS

Inicialização: Atribua $i = 0$; Inicialize aleatoriamente $\hat{A}^{(i=0)}$, $\hat{B}^{(i=0)}$ e $\hat{G}^{(i=0)}$ e construa as matrizes $F_y^{(i=0)}$ e $F_w^{(i=0)}$ a partir das equações (2.34);

1) Atribua $i = i + 1$.

2) Estime $\hat{B}^{(i)}$ através da equação:

$$\hat{B}^{(i)} = ((I \otimes \hat{A}^{(i-1)}) F_y^{(i-1)})^\dagger Y$$

3) Estime $\hat{A}^{(i)}$ através da equação:

$$\hat{A}^{(i)} = ((I \otimes \hat{B}^{(i-1)}) F_w^{(i-1)})^\dagger W$$

4) Estime $g \doteq \text{vec}(G)$ através da equação:

$$g = \left((C^{(B)T} \diamond C^{(A)T})^T \diamond (B^{(i-1)} \otimes A^{(i-1)}) \right)^\dagger \text{vec}(X)$$

5) Repita os passos 1 a 4 até a convergência.

Através das equações (2.33, 2.36 e 2.37) podemos construir um algoritmo para estimação das matrizes fatores A , B e G . Este algoritmo é conhecido como PARATUCK2-ALS e está resumido na Tabela 2.2 em que Y e W são as formas *unfolding* de \mathcal{X} conforme as equações (2.34).

2.2 Comunicações Cooperativas

Como já foi discutido na Seção 1.1 a introdução de um ou mais nós retransmissores (*relays*) à um sistema de comunicação tem como efeito a obtenção, de forma virtual, de alguns benefícios de um sistema MIMO sem necessariamente utilizarmos nós com várias antenas [2], [3]. Dentre esses benefícios destacam-se [32]:

► Diminuição da perda de percurso

Como podemos observar pela Figura 2.5 o enlace direto fonte-destino é maior do que o canal fonte-relay, desta forma utilizando técnicas de retransmissão adequadas podemos minimizar os efeitos de perda de percurso em sistemas cooperativos. Uma das consequências disto é que a rede aumenta o sua área de cobertura.

► Ganho de diversidade

A transmissão de cópias dos sinais pelos nós *relays* ao longo do processo de comunicação possibilita uma melhora no desempenho da probabilidade de erro, pois várias cópias do mesmo sinal serão recebidas, e, em virtude disso, as chances de erro de decodificação serão menores. Verificamos o ganho de diversidade pela mudança na inclinação do gráfico taxa de erro de bit \times SNR.

► Ganho de multiplexação

O ganho de multiplexação é um ganho devido ao aumento do número de caminhos independentes dentre os quais a informação pode percorrer. Esses enlaces possibilitam um aumento na taxa de transmissão.

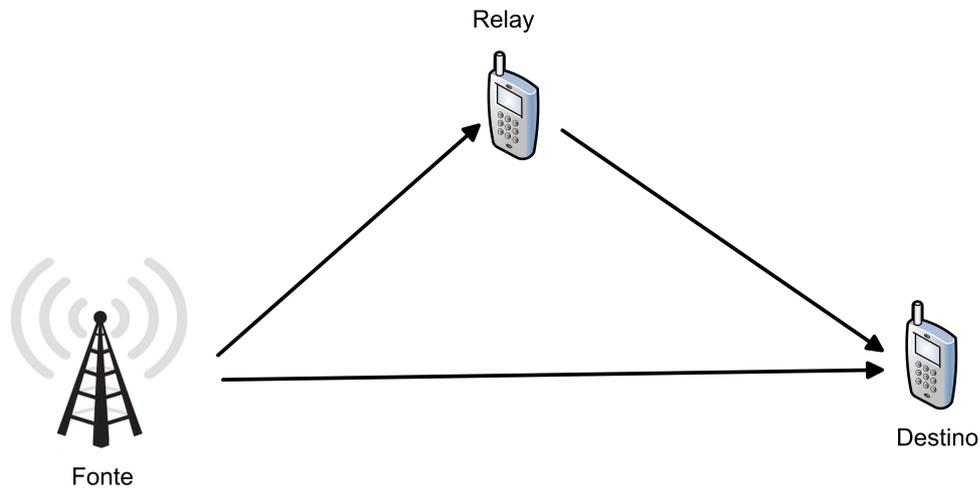


Figura 2.5: Sistema cooperativo com um nó retransmissor (*relay*).

2.2.1 Classificação dos Nós Retransmissores (*Relays*)

Os *relays* podem ser subdivididos quanto acesso ao meio em:

► *Half-Duplex*

Relays deste tipo são caracterizados por não enviarem e receberem os símbolos ao mesmo tempo. Esse tipo de técnica evita a interferência entre o que está sendo recebido e o que está sendo enviado (este tipo de interferência é chamada de interferência de *loop-back*). Por outro lado, o tempo de retransmissão aumenta.

► *Full-Duplex*

Relays deste tipo são caracterizados por enviarem e receberem os símbolos ao mesmo tempo. Esse tipo de técnica diminui a latência da rede, mas introduz a interferência de *loop-back*, logo, alguma técnica deverá ser usada para minimizar os efeitos desta interferência.

As técnicas de retransmissão podem obedecer aos seguintes protocolos de cooperação:

► Amplifica-e-encaminha (**AF**, do inglês, *Amplify-and-Forward*)

Relays que utilizam a técnica de retransmissão **AF**, amplificam o sinal recebido em cada antena por um fator de amplificação e encaminham o sinal resultante para o restante da rede. Podemos modelar o sinal recebido em *relay* multi-antena como um vetor, em que o sinal recebido em cada antena corresponde a um elemento do vetor. Assim, o processo de amplificação dos sinais nesse tipo de *relay* equivale a pré-multiplicarmos o vetor de sinal recebido no *relay* por uma matriz diagonal G em que cada elemento da

diagonal principal de G corresponde a um fator de amplificação para uma antena. Este tipo de técnica apresenta baixa complexidade, mas tem a desvantagem de propagar eventuais erros que o canal de transmissão ou ruído causam nos símbolos recebidos.

- Decodifica-e-encaminha (DF, do inglês, *Decode-and-Forward*)

Relays DF, detectam os sinais recebidos, decodificam os símbolos e codifica-os novamente antes de reenviar para o restante da rede. Este tipo de técnica é mais complexa que a técnica AF, mas tem a vantagem de não propagar os erros causados pelo canal. Por outro lado, a eficiência do protocolo DF está diretamente ligada a qualidade do enlace, ou seja, quanto pior o estado do canal mais erros de decodificação estarão presentes e, desta forma, estes erros de decodificação serão propagados para o restante dos nós da rede.

2.3 Trabalhos Relacionados

2.3.1 Método Baseado em SVD para o Cenário *One-Way Two-Hop*

2.3.1.1 Modelo do Sistema

Considere o sistema MIMO cooperativo mostrado na Figura 2.6. Ele consiste em um nó fonte com N antenas transmissoras, um nó destino com M antenas receptoras e um *relay AF* com R antenas. Ele é chamado *one-way* pois os dados trafegam somente na direção fonte-destino e também *two-hop* pelo fato de ser uma rede cooperativa de dois saltos. O processo de comunicação se divide em duas fases: na primeira fase a fonte transmite um vetor de sinal $\mathbf{x} \in \mathbb{C}^{N \times 1}$ para o *relay*. Assumindo desvanecimento plano, o sinal recebido no *relay* em forma discreta é dado por

$$\mathbf{r} = \mathbf{H}_{sr}\mathbf{x} + \mathbf{v}_r \in \mathbb{C}^{R \times 1}, \quad (2.39)$$

em que \mathbf{H}_{sr} é a matriz de canal MIMO fonte-*relay* e \mathbf{v}_r é o vetor de ruído no *relay* (aqui e em toda esta dissertação consideramos que o ruído aditivo é branco e Gaussiano (AWGN, do inglês, *Additive White Gaussian Noise*)).

Na segunda fase de transmissão, o *relay* amplifica o sinal 2.39 multiplicando-o por uma matriz diagonal $\mathbf{G} \in \mathbb{C}^{R \times R}$ que contem os fatores de amplificação para cada uma das antenas, ou seja, o i -ésimo termo da diagonal principal de \mathbf{G} modela o fator de amplificação da i -ésima antena do *relay*. O sinal resultante é enviado para nó destino. O sinal recebido no destino é dado

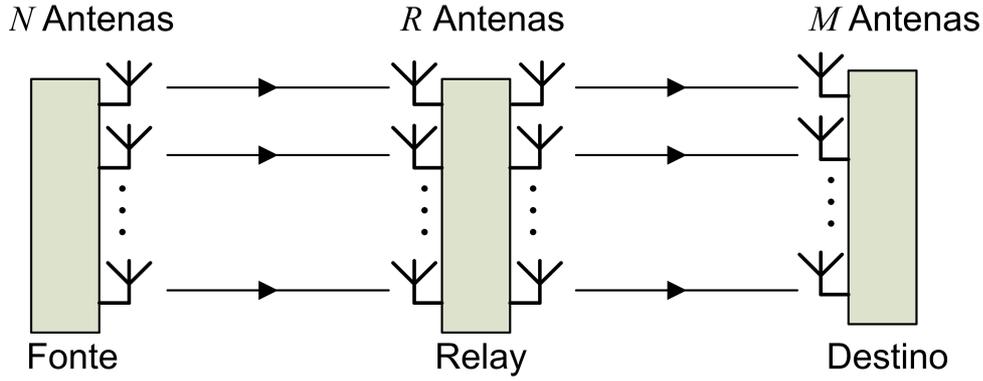


Figura 2.6: Cenário *One-way* MIMO *Two-Hop* com N antenas transmissoras, M antenas receptoras e *relay* AF com R antenas

por

$$\mathbf{y} = \mathbf{H}_{rd}\mathbf{G}\mathbf{H}_{sr}\mathbf{x} + \mathbf{H}_{rd}\mathbf{G}\mathbf{v}_r + \mathbf{v}_d \in \mathbb{C}^{M \times 1}, \quad (2.40)$$

em que \mathbf{H}_{rd} é a matriz de canal **MIMO** *relay*-destino e \mathbf{v}_d é o vetor de ruído **AWGN** no destino. O termo $\mathbf{H}_{rd}\mathbf{G}\mathbf{H}_{sr}\mathbf{x}$ é a componente de sinal do modelo e $\mathbf{H}_{rd}\mathbf{G}\mathbf{v}_r + \mathbf{v}_d$ é a componente de ruído. Afim de simplificar as equações iremos aqui desconsiderar a componente de ruído.

2.3.1.2 Algoritmo de Estimação de Canal

A seguir descreveremos o método introduzido por Lioliou e Viberg [9] para a estimação das matrizes de canal \mathbf{H}_{rd} e \mathbf{H}_{sr} . Para viabilizar a estimação de canal, assumimos que uma matriz $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_L]$, contendo sequências de treinamento conhecidas, é enviada em R blocos de tempo sucessivos e que os canais variam muito lentamente com o tempo, de sorte que podemos considerá-los constantes durante a duração temporal das sequências de treinamento. Assim, podemos construir a matriz $\mathbf{Y}^{(i)} = [\mathbf{y}_1, \dots, \mathbf{y}_L]$ de sinais recebidos no destino durante o i -ésimo bloco de tempo, a qual admite a seguinte expressão:

$$\mathbf{Y}^{(i)} = \mathbf{H}_{rd}\mathbf{G}^{(i)}\mathbf{H}_{sr}\mathbf{X} \in \mathbb{C}^{M \times L}, \quad i = 1, \dots, R. \quad (2.41)$$

Aqui assumimos que as matrizes de amplificação $\mathbf{G}^{(i)}$ são todas distintas e conhecidas pelo nó destino. Note ainda que este método exige que o número de blocos de tempo, i.e., o número de vezes que enviamos a matriz de treinamento \mathbf{X} , deve ser igual ao número de *relays* R . Assumindo $L \geq N$, podemos fazer uma estimativa do canal efetivo do i -ésimo bloco de tempo, $\mathbf{H}^{(i)} = \mathbf{H}_{rd}\mathbf{G}^{(i)}\mathbf{H}_{sr}$,

multiplicando (2.41) por X^\dagger , assim obtemos

$$Y^{(i)} X^\dagger = H^{(i)} = H_{\text{rd}} G^{(i)} H_{\text{sr}} \in \mathbb{C}^{M \times N}, i = 1, \dots, R. \quad (2.42)$$

A equação (2.42) sugere a seguinte estimativa de mínimos quadrados (LS, do inglês *Least Squares*) para H_{sr} e H_{rd}

$$\arg \min_{H_{\text{sr}}, H_{\text{rd}}} \sum_{i=1}^R \left\| H^{(i)} - H_{\text{rd}} G^{(i)} H_{\text{sr}} \right\|_F^2 \quad (2.43)$$

Aplicando o operador $\text{vec}(\cdot)$ em cada um dos termos de (2.43) podemos reescrevê-la como

$$\arg \min_{H_{\text{sr}}, H_{\text{rd}}} \sum_{i=1}^R \left\| \text{vec}(H^{(i)}) - (H_{\text{sr}}^T \diamond H_{\text{rd}}) \text{vecd}(G^{(i)}) \right\|^2, \quad (2.44)$$

em que aplicamos a propriedade $\text{vec}(\text{Adiag}(x)B) = (B^T \diamond A)x$. A equação (2.44) também pode ser expressa como

$$\arg \min_{H_{\text{sr}}, H_{\text{rd}}} \left\| H - \Theta G \right\|_F^2, \quad (2.45)$$

em que $H = [\text{vec}(H^{(1)}), \dots, \text{vec}(H^{(R)})]$, $\Theta = (H_{\text{sr}}^T \diamond H_{\text{rd}})$ e $G = [\text{vecd}(G^{(1)}), \dots, \text{vecd}(G^{(R)})]$. Podemos obter uma estimativa de Θ no sentido dos mínimos quadrados multiplicando (2.45) à direita pela pseudo-inversa de G , assim ficamos com

$$\arg \min_{H_{\text{sr}}, H_{\text{rd}}} \left\| \Theta - (H_{\text{sr}}^T \diamond H_{\text{rd}}) \right\|_F^2. \quad (2.46)$$

Como cada a i -ésima coluna de Θ consiste no produto de Kronecker entre a i -ésima linha de H_{sr} e a i -ésima coluna de H_{rd} , ou seja

$$\psi_i = H_{\text{sr}(i, \cdot)}^T \otimes H_{\text{rd}(\cdot, i)} \in \mathbb{C}^{MN \times 1}, i = 1, \dots, R, \quad (2.47)$$

podemos minimizar (2.46) coluna a coluna. Assim (2.46) pode ser reformulada como

$$\arg \min_{H_{\text{sr}(i, \cdot)}, H_{\text{rd}(\cdot, i)}} \left\| \psi_i - H_{\text{sr}(i, \cdot)}^T \otimes H_{\text{rd}(\cdot, i)} \right\|^2, i = 1, \dots, R \quad (2.48)$$

Tabela 2.3: Algoritmo *Least Squares Khatri-Rao Factorization*

Considere a matriz $C \in \mathbb{C}^{MN \times P}$ que é o resultado do produto de Khatri-Rao entre a matriz $A \in \mathbb{C}^{M \times P}$ e a matriz $B \in \mathbb{C}^{N \times P}$, i.e., $C = A \diamond B$. Atribua $p = 1$.

- 1) Seja c_p , a_p e b_p a p -ésima coluna das matrizes C , A e B , respectivamente. Sabemos que que $c_p = a_p \otimes b_p$.
- 2) Faça o *reshape* do vetor c_p na matriz C_p . Podemos ver que $C_p = b_p a_p^T$, i.e., C_p possui *rank* 1.
- 3) Expresse C_p em termos de sua **SVD**: $C_p = U_p \Sigma_p V_p^H$. Sabemos que a melhor aproximação de *rank* 1 de C_p pode ser obtida truncando a **SVD**, i.e., $\hat{a}_p = \sqrt{\sigma_1} v^*$ e $\hat{b}_p = \sqrt{\sigma_1} u$, em que u_1 e v_1 representam a primeira coluna de U e V , respectivamente, e σ_1 é o maior valor singular.
- 4) Se $p \leq P$, atribua $p = p + 1$ e retorne ao passo 1.

Rearranjando o vetor ψ_i em uma matriz $M \times N$ Ψ_i , podemos reformular o problema (2.48) da seguinte forma

$$\arg \min_{\mathbf{H}_{\text{sr}(i,\cdot)}, \mathbf{H}_{\text{rd}(\cdot,i)}} \left\| \Psi_i - \mathbf{H}_{\text{rd}(\cdot,i)} \mathbf{H}_{\text{sr}(i,\cdot)}^T \right\|_F^2, \quad i = 1, \dots, R \quad (2.49)$$

Claramente Ψ_i possui *rank* 1. Seja a sua **SVD**

$$\Psi_i = U \Sigma V^H = \sum_k \sigma_k \mathbf{u}_k \mathbf{v}_k^H, \quad (2.50)$$

então uma estimativa para $\mathbf{H}_{\text{rd}(\cdot,i)}$ e $\mathbf{H}_{\text{sr}(i,\cdot)}^T$ é dada por

$$\mathbf{H}_{\text{sr}(i,\cdot)}^T = \sqrt{\sigma_1} \mathbf{v}_1^H \quad (2.51)$$

$$\mathbf{H}_{\text{rd}(\cdot,i)} = \sqrt{\sigma_1} \mathbf{u}_1 \quad (2.52)$$

em que σ_1 é o maior valor singular de Ψ_i . O processo de estimação das matrizes de canal \mathbf{H}_{sr}^T e \mathbf{H}_{rd} consiste na fatoração do produto de Khatri-Rao ($\mathbf{H}_{\text{sr}}^T \diamond \mathbf{H}_{\text{rd}}$) em (2.46) e o algoritmo aqui empregado para essa estimação é chamado *Least Squares Khatri-Rao Factorization*. A Tabela 2.3 resume este algoritmo.

2.3.2 Método Baseado em PARAFAC para o Cenário *One-Way Two-Hop*

Ainda em relação ao cenário mostrado na Figura 2.6, um método de estimação de canal introduzido por Rong et al. [33] usa o modelo tensorial **PARAFAC** para a estimação das matrizes de canal \mathbf{H}_{sr} e \mathbf{H}_{rd} . A equação (2.42) pode ser vista como o i -ésimo *slice* frontal de um tensor **PARAFAC** (equação 2.13c) (visto que $\mathbf{G}^{(i)}$ é diagonal). Assim podemos construir um tensor

$$\mathcal{Y} = [\tilde{\mathbf{Y}}^{(1)} \sqcup_3 \tilde{\mathbf{Y}}^{(1)} \dots \sqcup_3 \tilde{\mathbf{Y}}^{(K)}] \in \mathbb{C}^{M \times N \times K} \quad (2.53)$$

em que o operador \sqcup_3 concatena as matrizes $\tilde{\mathbf{Y}}^{(i)}$ ao longo do terceiro modo de modo a formar o tensor \mathcal{Y} . Podemos expressar \mathcal{Y} em termos de suas matrizes fatores:

$$\mathcal{Y} = \mathcal{I} \times_1 \mathbf{H}_{\text{rd}} \times_2 \mathbf{H}_{\text{sr}}^T \times_3 \mathbf{G} \in \mathbb{C}^{M \times N \times K} \quad (2.54)$$

O método de estimação das matrizes \mathbf{H}_{sr} e \mathbf{H}_{rd} consiste em aplicar uma versão simplificada do algoritmo [PARAFAC-ALS](#) (Tabela 2.1): as matrizes fatores A e B equivalem às matrizes \mathbf{H}_{rd} \mathbf{H}_{sr}^T e a matriz C equivale à matriz $\mathbf{G} \in \mathbb{C}^{K \times R}$ em que a k -ésima linha de \mathbf{G} corresponde à diagonal de $\mathbf{G}^{(i)}$. Como assumimos que as matrizes $\mathbf{G}^{(i)}$ são conhecidas no receptor, o passo de estimação da matriz fator do modo-3 (matriz C) é omitido. Atente ainda para o fato de que este método não restringe que o número de blocos de tempo seja igual ao número de *relays*, i.e, em (2.53) K não necessariamente é igual a R . Ver [33] para mais detalhes.

Receptores Propostos

Ao longo deste capítulo, serão descritos os receptores propostos para o cenário *three-hop* em estudo. Este será o cenário básico que servirá de base para o desenvolvimento de outros protocolos de transmissão e, conseqüentemente, outros algoritmos de estimação. Como veremos através das simulações computacionais, em comparação com o cenário de 2 saltos, o cenário de 3 saltos apresenta um ganho bastante elevado tanto em termos de BER quanto em termos de NMSE. Isso se deve ao fato dos algoritmos propostos explorarem de forma rica a diversidade cooperativa que existe ao se combinar os vários enlaces que ligam o transmissor ao receptor.

O modelo consiste de um nó transmissor com N antenas, um nó receptor com M antenas e entre esses nós, e igualmente espaçados, 2 relays AF *half-duplex* com R_1 e R_2 antenas cada um, cujo papel é fazer a retransmissão dos sinais. Neste modelo, nos relays, o número de antenas usadas na recepção é igual ao número de antenas utilizadas na transmissão (esta consideração será relaxada em um segundo momento). Neste cenário o enlace transmissor-receptor é desconsiderado, de tal sorte que os sinais enviados pelo transmissor atingem no máximo o segundo relay, isto é exposto na Figura 3.1. Observe que, a partir do segundo relay, todos os sinais conseguem atingir o nó receptor.

Assumimos também que os canais permanecem constantes durante todas as etapas de transmissão e que transmissores e receptores estão sincronizados a nível de símbolo. Em cada etapa da transmissão somente um nó da rede faz o envio dos sinais (essa consideração não é válida para o caso *two-way*, detalhado no Apêndice A), dessa forma, temos uma rede que opera em múltiplo acesso por divisão no tempo (TDMA, do inglês *Time-Division Multiple Access*). É assumido também que todos os sinais enviados ocupam a mesma faixa de frequência.

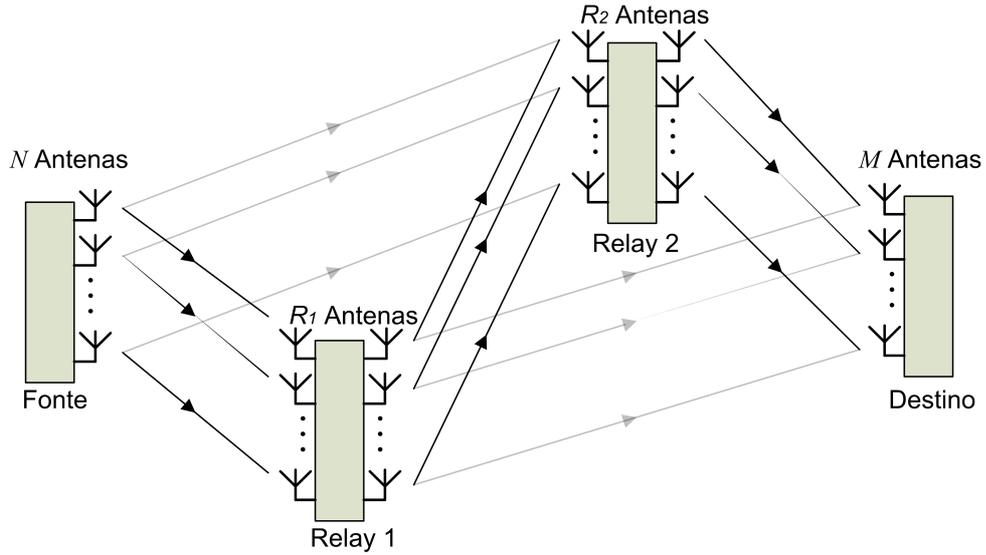


Figura 3.1: Modelo básico da rede cooperativa. N antenas transmissoras M antenas receptoras e dois *relays* AF.

3.1 Modelo em 4 Fases *One-Way*

Neste modelo, os dados trafegam do nó fonte para o nó destino com a ajuda de dois grupos de *relays* repetidores (ver Figura 3.1) de forma unidirecional (*one-way*) obedecendo o protocolo de transmissão que será descrito a seguir.

3.1.1 Descrição do Protocolo de Transmissão

O protocolo de transmissão é dividido em quatro fases. Na primeira fase de transmissão, o nó fonte transmite um vetor de sinal $\mathbf{x} \in \mathbb{C}^{N \times 1}$ e este sinal é recebido pelo primeiro e segundo *relays*. Estes sinais são dados em banda base discreta respectivamente por:

$$\mathbf{y}_{r_1}(1) = \mathbf{H}_{sr_1} \mathbf{x} + \mathbf{v}_{r_1}(1) \in \mathbb{C}^{R_1 \times 1}, \quad (3.1)$$

$$\mathbf{y}_{r_2}(1) = \mathbf{H}_{sr_2} \mathbf{x} + \mathbf{v}_{r_2}(1) \in \mathbb{C}^{R_2 \times 1}, \quad (3.2)$$

em que $\mathbf{H}_{sr_1} \in \mathbb{C}^{R_1 \times N}$ é a matriz de canal do enlace que liga o nó fonte ao primeiro grupo de *relays*, $\mathbf{H}_{sr_2} \in \mathbb{C}^{R_2 \times N}$ é a matriz de canal do enlace que liga o nó fonte ao segundo grupo de *relays*, e $\mathbf{v}_{r_1}(1)$ e $\mathbf{v}_{r_2}(1)$ são os vetores de ruído aditivo Gaussiano na fase 1 nos *relays* 1 e 2, respectivamente.

Na segunda fase de transmissão, o segundo *relay* amplifica o sinal recebido (3.2) multiplicando este sinal por uma matriz de amplificação diagonal \mathbf{G} de dimensões $R_2 \times R_2$ e transmite o sinal resultante para o destino. O sinal

recebido no destino é dado por:

$$\mathbf{y}_d(2) = \mathbf{H}_{r_2d} \mathbf{G} \mathbf{H}_{sr_2} \mathbf{x} + \mathbf{H}_{r_2d} \mathbf{G} \mathbf{v}_{r_2}(1) + \mathbf{v}_d(2) \in \mathbb{C}^{M \times 1}, \quad (3.3)$$

em que $\mathbf{H}_{r_2d} \in \mathbb{C}^{M \times R_2}$ é a matriz de canal ligando o segundo *relay* ao destino e $\mathbf{v}_d(2)$ é o vetor de ruído **AWGN** no destino. Na terceira fase de transmissão, o primeiro *relay* amplifica o sinal recebido (3.1) multiplicando-o por uma matriz diagonal \mathbf{F} de dimensões $R_1 \times R_1$ e transmite o sinal resultante para o destino e para o segundo *relay*. Os sinais recebido no segundo *relay* e no destino são dados, respectivamente por:

$$\mathbf{y}_{r_2}(3) = \mathbf{H}_{r_1r_2} \mathbf{F} \mathbf{H}_{sr_1} \mathbf{x} + \mathbf{H}_{r_1r_2} \mathbf{F} \mathbf{v}_{r_1}(1) + \mathbf{v}_{r_2}(3) \in \mathbb{C}^{R_2 \times 1}, \quad (3.4)$$

$$\mathbf{y}_d(3) = \mathbf{H}_{r_1d} \mathbf{F} \mathbf{H}_{sr_1} \mathbf{x} + \mathbf{H}_{r_1d} \mathbf{F} \mathbf{v}_{r_1}(1) + \mathbf{v}_d(3) \in \mathbb{C}^{M \times 1}, \quad (3.5)$$

em que $\mathbf{H}_{r_1d} \in \mathbb{C}^{M \times R_1}$ e $\mathbf{H}_{r_1r_2} \in \mathbb{C}^{R_2 \times R_1}$ são as matrizes de canal que modelam os enlaces que ligam o primeiro grupo de *relays* ao destino e ao segundo grupo de *relays*, respectivamente.

Finalmente, na quarta fase os *relays* do segundo grupo amplificam o sinal recebido (3.4) usando a mesma matriz diagonal \mathbf{G} e transmitem o sinal resultante para o nó destino. Nessa fase, o sinal recebido no destino é dado por:

$$\mathbf{y}_d(4) = \mathbf{H}_{r_2d} \mathbf{G} \mathbf{H}_{r_1r_2} \mathbf{F} \mathbf{H}_{sr_1} \mathbf{x} + \mathbf{H}_{r_2d} \mathbf{G} \mathbf{H}_{r_1r_2} \mathbf{F} \mathbf{v}_{r_1}(1) + \mathbf{H}_{r_2d} \mathbf{G} \mathbf{v}_{r_2}(3) + \mathbf{v}_d(4) \in \mathbb{C}^{M \times 1}. \quad (3.6)$$

Observe que a taxa de transmissão de dados para este protocolo é a metade da taxa de transmissão do cenário em duas fases (Seção 2.3), pois o protocolo de transmissão aqui descrito possui o dobro de retransmissões.

3.1.2 Formulação Tensorial

Como em [11] e [33] dividimos a fase de treinamento em K blocos de tempo. Para cada um desses blocos de tempo a mesma matriz de treinamento $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_L] \in \mathbb{C}^{N \times L}$ ortogonal (i.e., $\mathbf{X} \mathbf{X}^H = \mathbf{I}_N$) é transmitida pela fonte e retransmitida pelos *relays* até o destino obedecendo o protocolo de transmissão acima descrito. Sejam $\mathbf{W} \in \mathbb{C}^{K \times R_1}$ e $\mathbf{T} \in \mathbb{C}^{K \times R_2}$ matrizes cujas linhas contêm os fatores de amplificação utilizados no primeiro e segundo *relays*, respectivamente, durante os K blocos de tempo (o design dessas matrizes será descrito a seguir), podemos reescrever os modelos de dados das

equações (3.3), (3.5) e (3.6) como segue:

$$\begin{aligned} \mathbf{Y}_{d,k}(2) &= \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{H}_{sr_2} \mathbf{X} \\ &\quad + \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{V}_{r_2}(1) + \mathbf{V}_d(2) \in \mathbb{C}^{M \times L}, \end{aligned} \quad k = 1, \dots, K, \quad (3.7)$$

$$\begin{aligned} \mathbf{Y}_{d,k}(3) &= \mathbf{H}_{r_1d} D_k(\mathbf{W}) \mathbf{H}_{sr_1} \mathbf{X} \\ &\quad + \mathbf{H}_{r_1d} D_k(\mathbf{W}) \mathbf{V}_{r_1}(1) + \mathbf{V}_d(3) \in \mathbb{C}^{M \times L}, \end{aligned} \quad k = 1, \dots, K, \quad (3.8)$$

$$\begin{aligned} \mathbf{Y}_{d,k}(4) &= \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{H}_{r_1r_2} D_k(\mathbf{W}) \mathbf{H}_{sr_1} \mathbf{X} \\ &\quad + \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{H}_{r_1r_2} D_k(\mathbf{W}) \mathbf{V}_{r_1}(1) + \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{V}_{r_2}(3) \\ &\quad + \mathbf{V}_d(4) \in \mathbb{C}^{M \times L}, \end{aligned} \quad k = 1, \dots, K. \quad (3.9)$$

No nó destino, pós-multiplicando ambos os lados da equação (3.7) por \mathbf{X}^H obtemos:

$$\tilde{\mathbf{P}}_k = \mathbf{P}_k + \mathbf{V}_k^{(1)} \in \mathbb{C}^{M \times N}, \quad k = 1, \dots, K, \quad (3.10)$$

onde

$$\mathbf{P}_k = \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{H}_{sr_2}, \quad (3.11)$$

é a componente de sinal de nosso modelo e

$$\mathbf{V}_k^{(1)} = \mathbf{H}_{r_2d} D_k(\mathbf{T}) \mathbf{V}_{r_2}(1) \mathbf{X}^H + \mathbf{V}_d(2) \mathbf{X}^H, \quad (3.12)$$

é a componente de ruído.

Nosso modelo pode ser escrito em uma forma mais compacta usando a notação tensorial. Para isso, podemos introduzir

$$\mathcal{P} = [\mathbf{P}_1 \sqcup_3 \mathbf{P}_2 \dots \sqcup_3 \mathbf{P}_K] \in \mathbb{C}^{M \times N \times K}, \quad (3.13)$$

que pode ser expresso em termos de sua representação em produto modo- n

$$\mathcal{P} = \mathcal{I}_{3,R_2} \times_1 \mathbf{H}_{r_2d} \times_2 \mathbf{H}_{sr_2}^T \times_3 \mathbf{T}, \quad (3.14)$$

em que \mathcal{I}_{3,R_2} é o tensor identidade de tamanho $R_2 \times R_2 \times R_2$. Comparando (3.14) com (2.14) podemos ver que \mathcal{P} é um tensor que é modelado pela

decomposição **PARAFAC** [16].

Os *unfoldings* modo-1 e modo-2 de (3.14) são dados respectivamente por:

$$\begin{aligned} [\mathcal{P}]_{(1)} &= \mathbf{H}_{r_2d}[\mathcal{I}_{3,R_2}]_{(1)}(\mathbf{T} \otimes \mathbf{H}_{sr_2}^T)^T \\ &= \mathbf{H}_{r_2d}(\mathbf{T} \diamond \mathbf{H}_{sr_2}^T)^T \in \mathbb{C}^{N \times MK}, \end{aligned} \quad (3.15)$$

$$\begin{aligned} [\mathcal{P}]_{(2)} &= \mathbf{H}_{sr_2}^T[\mathcal{I}_{3,R_2}]_{(2)}(\mathbf{T} \otimes \mathbf{H}_{r_2d})^T \\ &= \mathbf{H}_{sr_2}^T(\mathbf{T} \diamond \mathbf{H}_{r_2d})^T \in \mathbb{C}^{M \times NK}, \end{aligned} \quad (3.16)$$

em que \diamond denota o produto de Khatri-Rao (**Definição 10**). Multiplicando ambos os lados de (3.8) à direita por \mathbf{X}^H , obtemos:

$$\tilde{\mathbf{Q}}_k = \mathbf{Q}_k + \mathbf{V}_k^{(2)} \in \mathbb{C}^{M \times N}, \quad k = 1, \dots, K, \quad (3.17)$$

em que

$$\mathbf{Q}_k = \mathbf{H}_{r_1d}D_k(\mathbf{W})\mathbf{H}_{sr_1}, \quad (3.18)$$

é a componente de sinal e

$$\mathbf{V}_k^{(2)} = \mathbf{H}_{r_1d}D_k(\mathbf{W})\mathbf{V}_{r_1}(1)\mathbf{X}^H + \mathbf{V}_d(3)\mathbf{X}^H, \quad (3.19)$$

é a componente de ruído. Similarmente a (3.13), podemos construir o tensor

$$\mathcal{Q} = [\mathbf{Q}_1 \sqcup_3 \mathbf{Q}_2 \dots \sqcup_3 \mathbf{Q}_K] \in \mathbb{C}^{M \times N \times K}, \quad (3.20)$$

o qual admite a seguinte decomposição **PARAFAC**:

$$\mathcal{Q} = \mathcal{I}_{3,R_1} \times_1 \mathbf{H}_{r_1d} \times_2 \mathbf{H}_{sr_1}^T \times_3 \mathbf{W}, \quad (3.21)$$

em que \mathcal{I}_{3,R_1} é o tensor identidade de tamanho $R_1 \times R_1 \times R_1$. Os *unfoldings* modo-1 e modo-2 de (3.20) são dados respectivamente por:

$$\begin{aligned} [\mathcal{Q}]_{(1)} &= \mathbf{H}_{r_1d}[\mathcal{I}_{3,R_1}]_{(1)}(\mathbf{W} \otimes \mathbf{H}_{sr_1}^T)^T \\ &= \mathbf{H}_{r_1d}(\mathbf{W} \diamond \mathbf{H}_{sr_1}^T)^T \in \mathbb{C}^{M \times NK}, \end{aligned} \quad (3.22)$$

$$\begin{aligned} [\mathcal{Q}]_{(2)} &= \mathbf{H}_{sr_1}^T[\mathcal{I}_{3,R_1}]_{(2)}(\mathbf{W} \otimes \mathbf{H}_{r_1d})^T \\ &= \mathbf{H}_{sr_1}^T(\mathbf{W} \diamond \mathbf{H}_{r_1d})^T \in \mathbb{C}^{M \times NK}. \end{aligned} \quad (3.23)$$

Os tensores \mathcal{P} e \mathcal{Q} são mostrados na Figura 3.2.

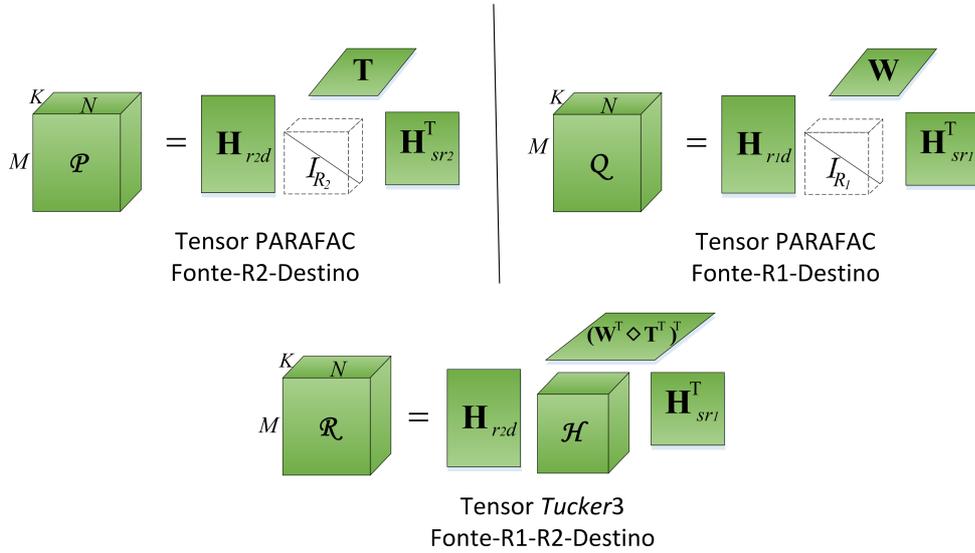


Figura 3.2: Tensores que modelam os dados recebidos dos enlaces Fonte-Relay 2-Destino, Fonte-Relay 1-Destino e Fonte-Relay 1-Relay 2-Destino.

Neste ponto, é importante mencionar que estimativas dos pares de matrizes $\{\mathbf{H}_{r2d}, \mathbf{H}_{sr2}\}$ e $\{\mathbf{H}_{r1d}, \mathbf{H}_{sr1}\}$ poderiam ser obtidas dos tensores PARAFAC (3.14) e (3.21), respectivamente, por meio de dois algoritmos PARAFAC-ALS independentes, como em [33]. Entretanto, explorando a diversidade cooperativa, podemos combinar cada um desses tensores com outro tensor do enlace Fonte-Relay 1-Relay 2-Destino, como será explicado a seguir. Tal abordagem resulta em uma estrutura tensorial mais rica que pode ser explorada para resolver este problema de estimação de canal. Isso geralmente traduz-se em uma condição de identificabilidade mais relaxada e em um aumento de desempenho.

Multiplicando ambos os lados de (3.9) à direita por \mathbf{X}^H , obtemos:

$$\tilde{\mathbf{R}}_k = \mathbf{R}_k + \mathbf{V}_k^{(3)} \in \mathbb{C}^{M \times N}, k = 1, \dots, K, \quad (3.24)$$

em que $\mathbf{R}_k = \mathbf{H}_{r2d} D_k(\mathcal{T}) \mathbf{H}_{r1r2} D_k(\mathcal{W}) \mathbf{H}_{sr1}$ é a componente de sinal e $\mathbf{V}_k^{(3)} = \mathbf{H}_{r2d} D_k(\mathcal{T}) \mathbf{H}_{r1r2} D_k(\mathcal{W}) \mathbf{V}_{r1}(1) \mathbf{X}^H + \mathbf{H}_{r2d} D_k(\mathcal{T}) \mathbf{V}_{r2}(3) \mathbf{X}^H + \mathbf{V}_d(4) \mathbf{X}^H$ é a componente de ruído. Observe que a componente de sinal \mathbf{R}_k pode ser vista como o k -ésimo *slice* frontal de um tensor PARATUCK2 (equação 2.26).

Introduzamos $\mathbf{r}_k \triangleq \text{vec}(\mathbf{R}_k) \in \mathbb{C}^{MN \times 1}$. Aplicando a propriedade $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, obtemos

$$\mathbf{r}_k = (\mathbf{H}_{sr1}^T \otimes \mathbf{H}_{r2d}) \text{diag}(\text{vec}(\mathbf{H}_{r1r2})) (\mathbf{W}_{(k,:)}^T \otimes \mathbf{T}_{(k,:)}^T). \quad (3.25)$$

Definindo $\mathbf{R} \doteq [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_K]^T \in \mathbb{C}^{K \times MN}$, vem

$$\mathbf{R} = (\mathbf{W}^T \diamond \mathbf{T}^T)^T \text{diag}(\text{vec}(\mathbf{H}_{r1r2})) (\mathbf{H}_{sr1}^T \otimes \mathbf{H}_{r2d})^T \in \mathbb{C}^{K \times MN}. \quad (3.26)$$

Comparando (3.26) com (3.15) e (3.23), \mathcal{R} pode ser visto como o *unfolding* modo-3 de um tensor \mathcal{H} , que admite a seguinte decomposição *Tucker3* [15]:

$$\mathcal{R} = \mathcal{H} \times_1 \mathbf{H}_{r_{2d}} \times_2 \mathbf{H}_{sr_1}^T \times_3 (\mathbf{W}^T \diamond \mathbf{T}^T)^T \in \mathbb{C}^{M \times N \times K}, \quad (3.27)$$

em que $\mathcal{H} \in \mathbb{C}^{R_2 \times R_1 \times R_1 R_2}$ é um tensor cujo *unfolding* modo-3 é dado por $[\mathcal{H}]_{(3)} = \text{Diag}(\text{vec}(\mathbf{H}_{r_{1r_2}}))$. O tensor \mathcal{R} é mostrado na Figura 3.2. Nós estamos interessados nos *unfoldings* modo-1 e modo-2 deste tensor, elas são dadas respectivamente por:

$$[\mathcal{R}]_{(1)} = \mathbf{H}_{r_{2d}} [\mathcal{H}]_{(1)} ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \mathbf{H}_{sr_1}^T)^T \in \mathbb{C}^{M \times NK}, \quad (3.28)$$

$$[\mathcal{R}]_{(2)} = \mathbf{H}_{sr_1}^T [\mathcal{H}]_{(2)} ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \mathbf{H}_{r_{2d}})^T \in \mathbb{C}^{N \times MK}. \quad (3.29)$$

Neste trabalho, escolhemos \mathbf{W} e \mathbf{T} tais que $\mathbf{U} = (\mathbf{W}^T \diamond \mathbf{T}^T)^T$ é uma matriz de transformada de Fourier (**DFT**, do inglês *Discrete Fourier Transform*) truncada de dimensões $K \times R_1 R_2$, que, se $K \geq R_1 R_2$ é *rank* coluna completo por definição. Deste design, as matrizes de amplificação \mathbf{W} e \mathbf{T} são obtidas por meio de um algoritmo de fatoração Least Squares Khatri-Rao Factorization (ver Tabela 2.3 para mais detalhes). O requisito de *rank* da componente $(\mathbf{W}^T \diamond \mathbf{T}^T)^T$ é necessário para garantir a estimação de $\mathbf{H}_{r_{1r_2}}$, como veremos a seguir.

3.1.3 Algoritmo de Estimação de Canal

Definamos $\mathbf{r} \doteq \text{vec}(\mathbf{R}^T) \in \mathbb{C}^{MNK \times 1}$. De (3.26) temos:

$$\mathbf{r} = ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \diamond (\mathbf{H}_{sr_1}^T \otimes \mathbf{H}_{r_{2d}})) \text{vec}(\mathbf{H}_{r_{1r_2}}), \quad (3.30)$$

em que foi aplicada mais uma vez a propriedade $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$. Definindo $\mathbf{h}_{r_{1r_2}} \doteq \text{vec}(\mathbf{H}_{r_{1r_2}}) \in \mathbb{C}^{R_1 R_2 \times 1}$, (3.29) conduz ao seguinte problema de mínimos quadrados:

$$\hat{\mathbf{h}}_{r_{1r_2}} = \arg \min_{\mathbf{h}_{r_{1r_2}}} \left\| \tilde{\mathbf{r}} - ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \diamond (\hat{\mathbf{H}}_{sr_1}^T \otimes \hat{\mathbf{H}}_{r_{2d}})) \mathbf{h}_{r_{1r_2}} \right\|^2,$$

em que $\tilde{\mathbf{r}}$ é uma versão corrompida por ruído de (3.30). A solução é dada por:

$$\hat{\mathbf{h}}_{r_{1r_2}} = \left[(\mathbf{W}^T \diamond \mathbf{T}^T)^T \diamond (\hat{\mathbf{H}}_{sr_1}^T \otimes \hat{\mathbf{H}}_{r_{2d}}) \right]^\dagger \tilde{\mathbf{r}}. \quad (3.31)$$

Para que a estimação de $\mathbf{h}_{r_{1r_2}}$ seja possível a pseudo-inversa em (3.31) deve existir e, para isso, a componente $(\mathbf{W}^T \diamond \mathbf{T}^T)^T$ deve ser *rank* coluna completo. Combinando (3.15) e (3.28) uma estimação LS para $\mathbf{H}_{r_{2d}}$ pode ser obtida

resolvendo o seguinte problema:

$$\arg \min_{\mathbf{H}_{r2d}^T} \left\| \begin{bmatrix} [\tilde{\mathcal{P}}]_{(1)}^T \\ [\tilde{\mathcal{R}}]_{(1)}^T \end{bmatrix} - \begin{bmatrix} (\mathbf{T} \diamond \hat{\mathbf{H}}_{sr2}^T) \\ ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{sr1}^T) [\mathcal{H}]_{(1)}^T \end{bmatrix} \mathbf{H}_{r2d}^T \right\|_F^2,$$

cujas solução é dada por:

$$\hat{\mathbf{H}}_{r2d}^T = \begin{bmatrix} (\mathbf{T} \diamond \hat{\mathbf{H}}_{sr2}^T) \\ ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{sr1}^T) [\mathcal{H}]_{(1)}^T \end{bmatrix}^\dagger \begin{bmatrix} [\tilde{\mathcal{P}}]_{(1)}^T \\ [\tilde{\mathcal{R}}]_{(1)}^T \end{bmatrix}, \quad (3.32)$$

em que $\tilde{\mathcal{P}}$ e $\tilde{\mathcal{R}}$ são versões corrompidas por ruído de \mathcal{P} e \mathcal{R} , respectivamente. Com um procedimento similar, podemos combinar (3.16) e (3.29) para formular o seguinte problema para estimar \mathbf{H}_{sr1} :

$$\arg \min_{\mathbf{H}_{sr1}} \left\| \begin{bmatrix} [\tilde{\mathcal{Q}}]_{(2)}^T \\ [\tilde{\mathcal{R}}]_{(2)}^T \end{bmatrix} - \begin{bmatrix} (\mathbf{W} \diamond \hat{\mathbf{H}}_{r1d}) \\ ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{r2d}) [\mathcal{H}]_{(2)}^T \end{bmatrix} \mathbf{H}_{sr1} \right\|_F^2,$$

cujas solução é dada por:

$$\hat{\mathbf{H}}_{sr1} = \begin{bmatrix} (\mathbf{W} \diamond \hat{\mathbf{H}}_{r1d}) \\ ((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{r2d}) [\mathcal{H}]_{(2)}^T \end{bmatrix}^\dagger \begin{bmatrix} [\tilde{\mathcal{Q}}]_{(2)}^T \\ [\tilde{\mathcal{R}}]_{(2)}^T \end{bmatrix}, \quad (3.33)$$

em que $\tilde{\mathcal{Q}}$ é uma versão com ruído de \mathcal{Q} . Adicionalmente, de (3.16) formulamos o seguinte problema LS para estimar \mathbf{H}_{sr2} :

$$\begin{aligned} \hat{\mathbf{H}}_{sr2}^T &= \arg \min_{\mathbf{H}_{sr2}^T} \left\| [\tilde{\mathcal{P}}]_{(2)} - \mathbf{H}_{sr2}^T (\mathbf{T} \diamond \hat{\mathbf{H}}_{r2d})^T \right\|_F^2 \\ &= [\tilde{\mathcal{P}}]_{(2)} \left[(\mathbf{T} \diamond \hat{\mathbf{H}}_{r2d})^T \right]^\dagger. \end{aligned} \quad (3.34)$$

Similarmente, de (3.22) obtemos a seguinte estimativa para \mathbf{H}_{r1d}

$$\begin{aligned} \hat{\mathbf{H}}_{r1d} &= \arg \min_{\mathbf{H}_{r1d}} \left\| [\tilde{\mathcal{Q}}]_{(1)} - \mathbf{H}_{r1d} (\mathbf{W} \diamond \hat{\mathbf{H}}_{sr1}^T)^T \right\|_F^2 \\ &= [\tilde{\mathcal{Q}}]_{(1)} \left[(\mathbf{W} \diamond \hat{\mathbf{H}}_{sr1}^T)^T \right]^\dagger. \end{aligned} \quad (3.35)$$

A Tabela 3.1 resume o algoritmo proposto, aqui referido como Combined ALS 1 (Comb-ALS-1).

Identificabilidade: Podemos notar que \mathbf{H}_{r2d} e \mathbf{H}_{sr1} são identificáveis de (3.32) e (3.33) se $\mathbf{T} \diamond \mathbf{H}_{sr2}^T$ e $\mathbf{W} \diamond \mathbf{H}_{r1d}$ são *rank* coluna completo. Desde que

Tabela 3.1: Algoritmo Comb-ALS-1

<p><i>Inicialização:</i> Atribua $i = 0$; Inicialize aleatoriamente $\hat{\mathbf{H}}_{\text{sr}_1}^{(i=0)}$, $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i=0)}$, $\hat{\mathbf{H}}_{\text{sr}_2}^{(i=0)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i=0)}$;</p> <ol style="list-style-type: none"> 1) Atribua $i = i + 1$. 2) Estime $\hat{\mathbf{H}}_{\text{r}_1\text{r}_2}^{(i)}$ usando $\hat{\mathbf{H}}_{\text{sr}_1}^{(i-1)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i-1)}$ na equação (3.31); 3) Estime $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i)}$ e $\hat{\mathbf{H}}_{\text{sr}_1}^{(i)}$ usando as equações (3.32) e (3.33); 4) Estime $\hat{\mathbf{H}}_{\text{sr}_2}^{(i)}$ e $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i)}$ usando as equações (3.34) e (3.35); 5) Repita os passos 1 a 4 até a convergência.
--

$\{\hat{\mathbf{H}}_{\text{sr}_2}^T, \hat{\mathbf{H}}_{\text{r}_1\text{d}}\}$ possuem entradas independentes e identicamente distribuídas (do inglês, independent identically distributed (i.i.d.)), essa condição de *rank* é assegurada se $R_1 \leq MK$ e $R_2 \leq NK$. Portanto, a identificabilidade de $\mathbf{H}_{\text{r}_2\text{d}}$ e \mathbf{H}_{sr_1} é garantida se estas condições forem satisfeitas. As equações (3.34) e (3.35), que correspondem as estimativas de \mathbf{H}_{sr_2} e $\mathbf{H}_{\text{r}_1\text{d}}$, também requerem $R_1 \leq MK$ e $R_2 \leq NK$, respectivamente. Adicionalmente, note que a estimação de $\mathbf{h}_{\text{r}_1\text{r}_2}$ de (3.31) requer $R_1R_2 \leq MNK$. Combinando essas condições e atentando para o fato do design de \mathbf{W} e \mathbf{T} requerer $K \geq R_1R_2$ chegamos ao seguinte limitante inferior para o número de blocos de tempo: $K \geq \max(\frac{R_1}{M}, \frac{R_2}{N}, \frac{R_1R_2}{MN}, R_1R_2) = \max(R_1R_2)$.

3.1.4 Modelo em 3 Fases *One-Way*

Este algoritmo é similar ao algoritmo anterior, com exceção de que temos agora uma fase a menos (foi retirada a fase 2, i.e. equação 3.3), como explicaremos a seguir. Este caso é de especial relevância quando o enlace fonte-*relay* 2 apresenta um desvanecimento profundo e o sistema deixa de utilizá-lo. Além disso, com esse caso em 3 fases podemos analisar o impacto que a diminuição de uma fase no protocolo tem em termos de eficiência de estimação de canal e complexidade computacional do sistema.

3.1.4.1 Protocolo de Transmissão

O processo de comunicação é dividido em três fases. Na primeira fase, a fonte envia um vetor de sinal $\mathbf{x} \in \mathbb{C}^{N \times 1}$ e estes dados são recebidos pelo primeiro e segundo *relays*. Em notação banda base discreta o sinal recebido pelo primeiro *relay* é dado por:

$$\mathbf{y}_{\text{r}_1}(1) = \mathbf{H}_{\text{sr}_1} \mathbf{x} + \mathbf{v}_{\text{r}_1}(1) \in \mathbb{C}^{R_1 \times 1}. \quad (3.36)$$

Na segunda fase do protocolo de transmissão, o primeiro *relay* amplifica o sinal (3.36) por meio de uma matriz diagonal de tamanho $R_1 \times R_1$ \mathbf{F} e retransmitem este sinal amplificado para o segundo *relay* e para o destino. Os sinais recebidos no segundo *relay* e no destino são dados respectivamente

por:

$$\mathbf{y}_{r_2}(2) = \mathbf{H}_{r_1r_2}\mathbf{F}\mathbf{H}_{sr_1}\mathbf{x} + \mathbf{H}_{r_1r_2}\mathbf{F}\mathbf{v}_{r_1}(1) + \mathbf{v}_{r_2}(2) \in \mathbb{C}^{R_2 \times 1}, \quad (3.37)$$

$$\mathbf{y}_d(2) = \mathbf{H}_{r_1d}\mathbf{F}\mathbf{H}_{sr_1}\mathbf{x} + \mathbf{H}_{r_1d}\mathbf{F}\mathbf{v}_{r_1}(1) + \mathbf{v}_d(2) \in \mathbb{C}^{M \times 1}. \quad (3.38)$$

Finalmente, na terceira fase de transmissão, os *relays* do segundo grupo amplificam o sinal (3.37) usando uma matriz diagonal de tamanho $R_2 \times R_2$ \mathbf{G} e retransmitem este sinal amplificado para o nó destino. O sinal recebido no destino é dado por:

$$\begin{aligned} \mathbf{y}_d(3) = \mathbf{H}_{r_2d}\mathbf{G}\mathbf{H}_{r_1r_2}\mathbf{F}\mathbf{H}_{sr_1}\mathbf{x} + \mathbf{H}_{r_2d}\mathbf{G}\mathbf{H}_{r_1r_2}\mathbf{F}\mathbf{v}_{r_1}(1) \\ + \mathbf{H}_{r_2d}\mathbf{G}\mathbf{v}_{r_2}(2) + \mathbf{v}_d(3) \in \mathbb{C}^{M \times 1}. \end{aligned} \quad (3.39)$$

A taxa de transmissão deste protocolo é maior do que a taxa de transmissão do protocolo da subseção anterior por um fator de quatro para três, visto que este protocolo tem uma fase a menos. Já em comparação com o protocolo em duas fases (Seção 2.3), a taxa de transmissão é menor por um fator de dois terços.

3.1.4.2 Formulação Tensorial

Da mesma forma que na subseção 3.1.2, podemos dividir a fase de treinamento em K blocos de tempo e para cada um deles a mesma matriz de treinamento $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_L] \in \mathbb{C}^{N \times L}$ ortogonal é transmitida obedecendo o protocolo de transmissão descrito acima. Assim, podemos reescrever os modelos de dados das equações (3.38), (3.39) como segue:

$$\begin{aligned} \mathbf{Y}_{d,k}(2) = \mathbf{H}_{r_1d}D_k(\mathbf{W})\mathbf{H}_{sr_1}\mathbf{X} \\ + \mathbf{H}_{r_1d}D_k(\mathbf{W})\mathbf{V}_{r_1}(1) + \mathbf{V}_d(2) \in \mathbb{C}^{M \times L}, \\ k = 1, \dots, K \end{aligned} \quad (3.40)$$

$$\begin{aligned} \mathbf{Y}_{d,k}(3) = \mathbf{H}_{r_2d}D_k(\mathbf{T})\mathbf{H}_{r_1r_2}D_k(\mathbf{W})\mathbf{H}_{sr_1}\mathbf{X} \\ + \mathbf{H}_{r_2d}D_k(\mathbf{T})\mathbf{H}_{r_1r_2}D_k(\mathbf{W})\mathbf{V}_{r_1}(1) + \mathbf{H}_{r_2d}D_k(\mathbf{T})\mathbf{V}_{r_2}(2) \\ + \mathbf{V}_d(3) \in \mathbb{C}^{M \times L}, \\ k = 1, \dots, K \end{aligned} \quad (3.41)$$

Tabela 3.2: Algoritmo Comb-ALS-1.1

<p><i>Inicialização:</i> Atribua $i = 0$; Inicialize aleatoriamente $\hat{\mathbf{H}}_{\text{sr}_1}^{(i=0)}$, $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i=0)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i=0)}$;</p> <ol style="list-style-type: none"> 1) Atribua $i = i + 1$. 2) Estime $\hat{\mathbf{H}}_{\text{r}_1\text{r}_2}^{(i)}$ usando $\hat{\mathbf{H}}_{\text{sr}_1}^{(i-1)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i-1)}$ na equação (3.31); 3) Estime $\hat{\mathbf{H}}_{\text{sr}_1}^{(i)}$ e $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i)}$ usando as equações (3.33) e (3.35); 4) Estime $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i)}$ usando a equação (3.44); 5) Repita os passos 1 a 4 até a convergência.
--

em que $\mathbf{W} \in \mathbb{C}^{K \times R_1}$ e $\mathbf{T} \in \mathbb{C}^{K \times R_2}$ são matrizes cujas linhas contêm os fatores de amplificação utilizados nos *relays* do primeiro e segundo grupo, durante os K blocos de tempo.

Multiplicando ambos os lados de (3.40) e (3.41) à direita por \mathbf{X}^H , obtemos:

$$\tilde{\mathbf{Q}}_k = \mathbf{Q}_k + \mathbf{V}_k^{(2)} \in \mathbb{C}^{M \times N}, \quad k = 1, \dots, K, \quad (3.42)$$

e

$$\tilde{\mathbf{R}}_k = \mathbf{R}_k + \mathbf{V}_k^{(3)} \in \mathbb{C}^{M \times N}, \quad k = 1, \dots, K, \quad (3.43)$$

em que \mathbf{Q}_k e \mathbf{R}_k são as componente de sinal e $\mathbf{V}_k^{(2)}$ e $\mathbf{V}_k^{(3)}$ são as componentes de ruído definidas em (3.17) e (3.24). Assim, podemos construir de forma similar a (3.20) e (3.27) os tensores \mathcal{Q} e \mathcal{R} .

3.1.4.3 Algoritmo de Estimação de Canal

O algoritmo de estimação de canal para este modelo é similar ao anterior, como a diferença de que o enlace da fonte para o segundo grupo de *relays* não está disponível no destino devido ao protocolo aqui utilizado, ou seja, no destino o tensor \mathcal{P} (equação 3.14) não está disponível, e portanto, não irá fazer parte do algoritmo de estimação de canal. Assim, as matrizes de canal envolvidas no processo de comunicação são: $\mathbf{H}_{\text{r}_1\text{r}_2}$, \mathbf{H}_{sr_1} , $\mathbf{H}_{\text{r}_1\text{d}}$ e $\mathbf{H}_{\text{r}_2\text{d}}$. A solução LS para a estimação de $\hat{\mathbf{h}}_{\text{r}_1\text{r}_2}$ é dado pela equação (3.31), assim como estimativas LS para $\hat{\mathbf{H}}_{\text{sr}_1}$ e $\hat{\mathbf{H}}_{\text{r}_1\text{d}}$ são dadas pelas equações (3.33) e (3.35). Para fazer a estimação de $\hat{\mathbf{H}}_{\text{r}_2\text{d}}$ podemos resolver o seguinte problema LS:

$$\arg \min_{\mathbf{H}_{\text{r}_2\text{d}}^T} \left\| \left[\tilde{\mathcal{R}} \right]_{(1)}^T - \left((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{\text{sr}_1}^T \right) [\mathcal{H}]_{(1)}^T \mathbf{H}_{\text{r}_2\text{d}}^T \right\|_F^2,$$

cuja solução é dada por:

$$\hat{\mathbf{H}}_{\text{r}_2\text{d}}^T = \left[\left((\mathbf{W}^T \diamond \mathbf{T}^T)^T \otimes \hat{\mathbf{H}}_{\text{sr}_1}^T \right) [\mathcal{H}]_{(1)}^T \right]^\dagger \left[\tilde{\mathcal{R}} \right]_{(1)}^T. \quad (3.44)$$

A Tabela 3.2 resume o algoritmo proposto, aqui referido como Combined ALS 1.1 (Comb-ALS-1.1).

Identificabilidade: Podemos notar que as entradas de $[\mathcal{H}]_{(1)}$ são i.i.d portanto $[\mathcal{H}]_{(1)}^T$ em (3.44) é rank coluna completo, além disso o produto de Kronecker nesta expressão também é rank coluna completo pois \mathbf{H}_{sr_1} possui entradas i.i.d, portanto a identificabilidade de \mathbf{H}_{r_2d} é garantida desde que $R_2 \leq NK$. \mathbf{H}_{sr_1} é identificável de (3.33) desde que $R_1 \leq MK$. Essa condição também garante a identificabilidade \mathbf{H}_{r_1d} (equação 3.35). Finalmente, a identificabilidade de $\mathbf{h}_{r_1r_2}$ é garantida se $R_1R_2 \leq MNK$ (equação 3.31). Combinando essas condições e levando em consideração que $K \geq R_1R_2$, chegamos ao seguinte limitante inferior para o número de blocos de tempo: $K \geq \max(\frac{R_1}{M}, \frac{R_2}{N}, \frac{R_1R_2}{MN}, R_1R_2) = \max(R_1R_2)$.

3.1.5 Simulações e Resultados

Nesta seção, os desempenhos dos algoritmos Comb-ALS-1 e Comb-ALS-1.1 serão avaliados através de simulações computacionais. O algoritmo Comb-ALS-1 será comparado com um estimador sequencial que consiste em primeiramente aplicar o método baseado em SVD (seção 2.3) para estimar $\{\mathbf{H}_{r_1d}, \mathbf{H}_{sr_1}\}$ e $\{\mathbf{H}_{r_2d}, \mathbf{H}_{sr_2}\}$ através das equações (3.18) e (3.15), respectivamente, seguido de um passo LS para estimar $\mathbf{H}_{r_1r_2}$ através da equação (3.31). Este método será aqui referido como SVD-LS e está resumido na Tabela 3.3. O algoritmo Comb-ALS-1.1 será comparado com outro algoritmo sequencial que consiste em aplicar o algoritmo PARAFAC-ALS em (3.14) para estimar $\{\mathbf{H}_{r_1d}, \mathbf{H}_{sr_1}\}$ seguido de um PARATUCK2-ALS no conjunto de equações (3.24), que pode ser visto como *slices* frontais de um tensor que é modelado através do modelo PARATUCK2, para estimar \mathbf{H}_{sr_1} , \mathbf{H}_{r_2d} e $\mathbf{H}_{r_1r_2}$. Este algoritmo será referido como PARAFAC-PARATUCK2. Note que este métodos sequenciais (SVD-LS e PARAFAC-PARATUCK2) não combinam os enlaces dos dois grupos de *relays* no destino para estimar conjuntamente \mathbf{H}_{sr_1} , \mathbf{H}_{r_2d} e $\mathbf{H}_{r_1r_2}$, em contraste com os algoritmos Comb-ALS-1 e Comb-ALS-1.1.

Nestas simulações, assumimos um sistema em que a fonte e o destino têm $N = 4$ e $M = 4$ antenas, respectivamente. O processo de comunicação é assistido por 2 grupos de *relays*, possuindo 4 antenas cada um (i.e., $R_1 = R_2 = 4$). O expoente de perda de percurso utilizado nas simulações é igual a 3.

Outra consideração que fazemos aqui é que, em cada etapa de transmissão, ruído aditivo Gaussiano é adicionado ao sinal recebido com uma potência fixa de 20dB, calculada da seguinte forma:

$$\text{Pot} = 10 \times \log \left\| \mathbf{V} \right\|_F^2. \quad (3.45)$$

Em que \mathbf{V} é a matriz com as componentes de ruído aditivo Gaussiano.

Para o método SVD-LS, \mathbf{W} e \mathbf{T} são DFTs normalizadas tais que $\mathbf{W}^T \mathbf{W}^* =$

Tabela 3.3: Algoritmo SVD-LS

O Algoritmo consiste em 3 fases:

- 1) Utilizando o enlace Fonte-Relay 1-Destino, estime as matrizes \hat{H}_{sr_1} e \hat{H}_{r_1d} utilizando o Algoritmo baseado em SVD descrito na Seção 2.3;
- 2) Utilizando o enlace Fonte-Relay 2-Destino, estime as matrizes \hat{H}_{sr_2} e \hat{H}_{r_2d} utilizando este mesmo algoritmo baseado em SVD;
- 3) Finalmente, estime $\hat{H}_{r_1r_2}$ através da equação (3.31) utilizando as estimativas de $\{\hat{H}_{sr_1} \hat{H}_{r_1d}\}$ e $\{\hat{H}_{sr_2} \hat{H}_{r_2d}\}$;

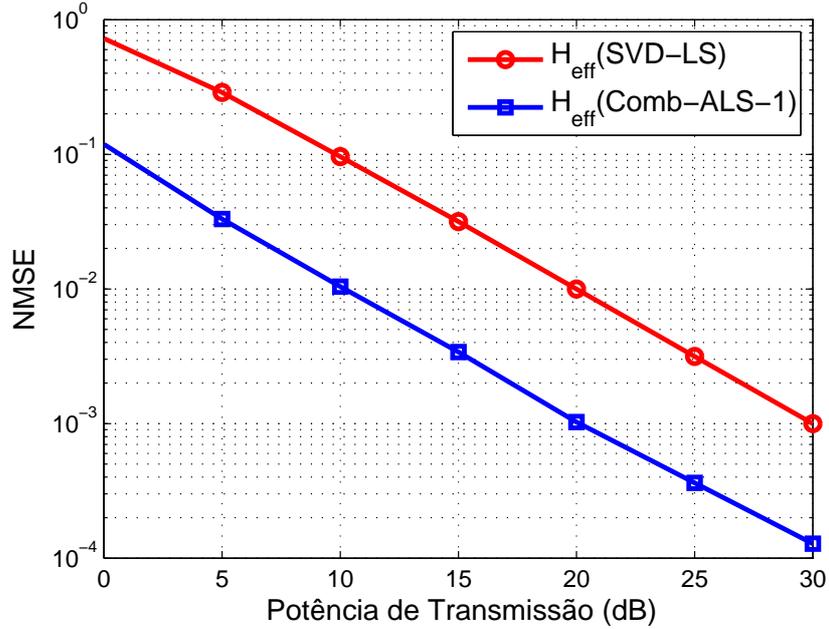


Figura 3.3: NMSE vs. Potência de Transmissão (Comparação entre o algoritmo Comb-ALS-1 e SVD-LS). Para o canal H_{eff} em (3.47)

I_{R_1} e $T^T T^* = I_{R_2}$. Consideramos que o sistema irá transmitir em $K = 4$ blocos de tempo .

A precisão da estimação de canal é avaliada em termos da NMSE do canal efetivo H_{eff} em (3.47) e (3.48) para os algoritmos Comb-ALS-1 e Comb-ALS-1.1, respectivamente. Serão feitas 5000 rodadas de Monte Carlo independentes em que, em cada rodada, uma realização diferente para todas as matrizes de canal e matrizes de ruído é feita, ao final uma média dos resultados é realizada.

Assim:

$$\text{NMSE}(\hat{H}_{\text{eff}}) = \frac{1}{5000} \sum_{i=1}^{5000} \frac{\|\hat{H}_{\text{eff}}(i) - H_{\text{eff}}\|_F^2}{\|H_{\text{eff}}(i)\|_F^2}. \quad (3.46)$$

A Figura 3.3 mostra a NMSE do canal efetivo em função da potência de transmissão para o algoritmo Comb-ALS-1 e SVD-LS, em que a potência de transmissão foi calculada de forma análoga à (3.45). Note que o método

proposto (Comb-ALS-1) possui um desempenho superior ao método **SVD-LS**: para um valor de **NMSE** fixa, o ganho em termos de potência é próximo de 10dB. Adicionalmente, a Figura 3.4 mostra o desempenho de **BER** destes algoritmos para o receptor Zero Forcing (**ZF**) em função da potência de transmissão dos dados (depois da fase de treinamento), em que fixamos a potência da fase de treinamento em 20dB (aqui e em toda as simulações de **BER** desta dissertação, foi utilizada a modulação 4-**QAM**). Essa figura também mostra os resultados obtidos com o receptor **ZF** assumindo o caso de perfeito conhecimento do canal (perfect **CSI**). O receptor **ZF** é baseado na combinação de todos os enlaces envolvidos no processo de transmissão da fonte para o destino. Essa solução **ZF** pode ser escrita como:

$$\hat{\mathbf{x}}_{\text{ZF}}(t) = \underbrace{\begin{bmatrix} \hat{\mathbf{H}}_{r_2d} \mathbf{G} \hat{\mathbf{H}}_{sr_2} \\ \hat{\mathbf{H}}_{r_1d} \mathbf{F} \hat{\mathbf{H}}_{sr_1} \\ \hat{\mathbf{H}}_{r_2d} \mathbf{G} \hat{\mathbf{H}}_{r_1r_2} \mathbf{F} \hat{\mathbf{H}}_{sr_1} \\ \hat{\mathbf{H}}_{\text{eff}} \end{bmatrix}}^{\dagger} \begin{bmatrix} \mathbf{y}_d(2) \\ \mathbf{y}_d(3) \\ \mathbf{y}_d(4) \end{bmatrix}. \quad (3.47)$$

A Figura 3.5 mostra os resultados de simulação de **NMSE** e compara o algoritmo Comb-ALS-1.1 com o algoritmo sequencial PARAFAC-PARATUCK2. Note que o método proposto (Comb-ALS-1.1) consegue atingir melhores estimativas comparado com o método PARAFAC-PARATUCK2. Para altas potências de transmissão, notamos que as estimativas de **NMSE** do método

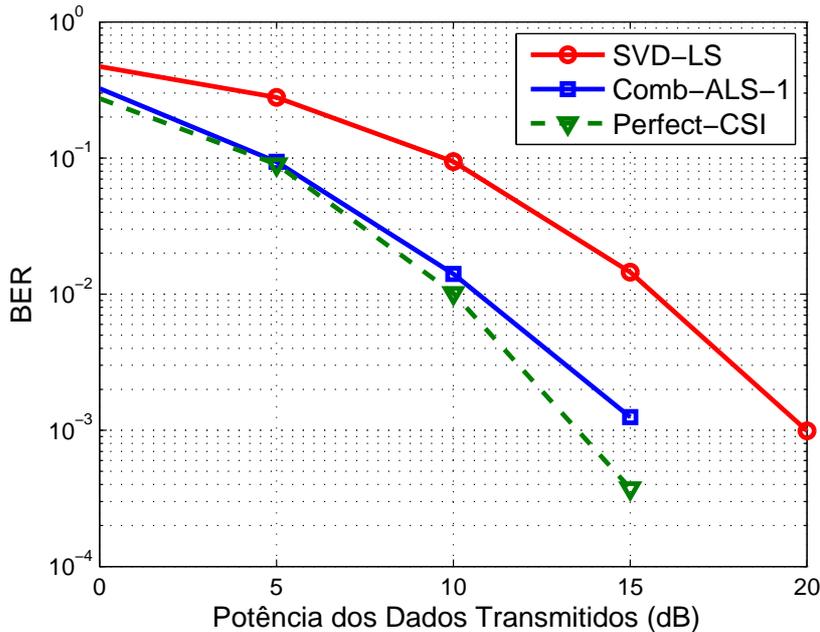


Figura 3.4: BER vs. Potência de transmissão dos dados. Potência de transmissão na fase de treinamento = 20dB. $N = M = R_1 = R_2 = 4$)

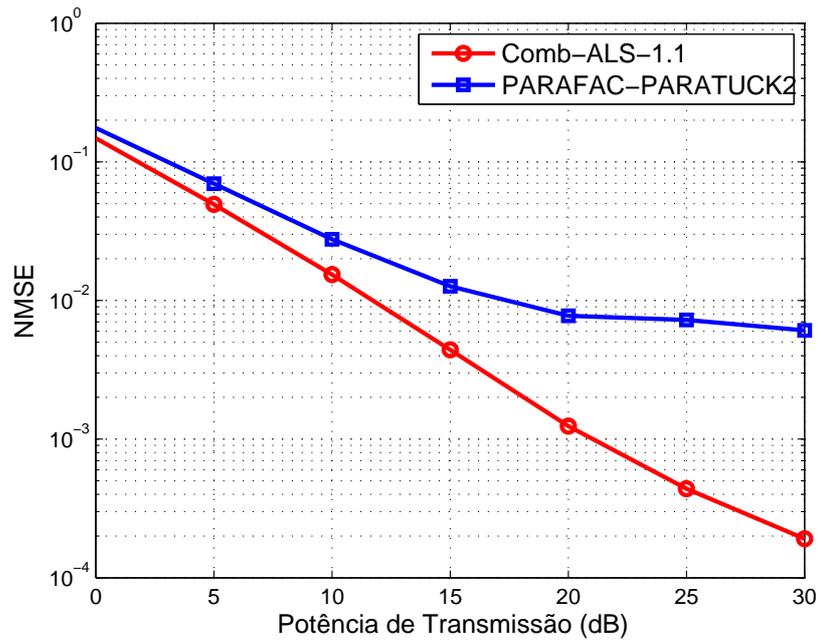


Figura 3.5: NMSE vs. Potência de transmissão (Comparação entre o algoritmo Comb-ALS-1.1 e PARAFAC-PARATUCK2). Para o canal H_{eff} em (3.48).

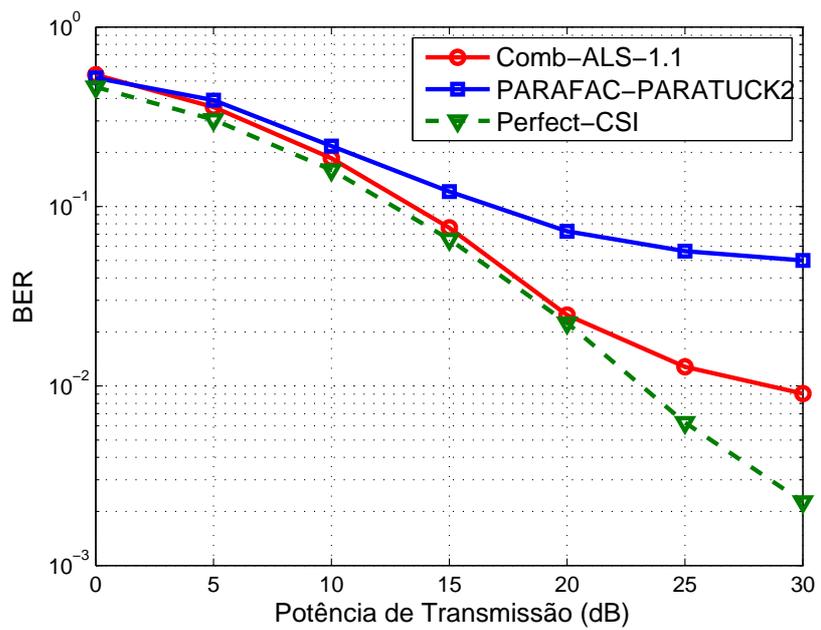


Figura 3.6: BER vs. SNR. Potência de transmissão = 20dB.

PARAFAC-PARATUCK2 não diminuem e permanecem em torno de 10^{-2} enquanto que as do método Comb-ALS-1.1 possuem um decaimento linear.

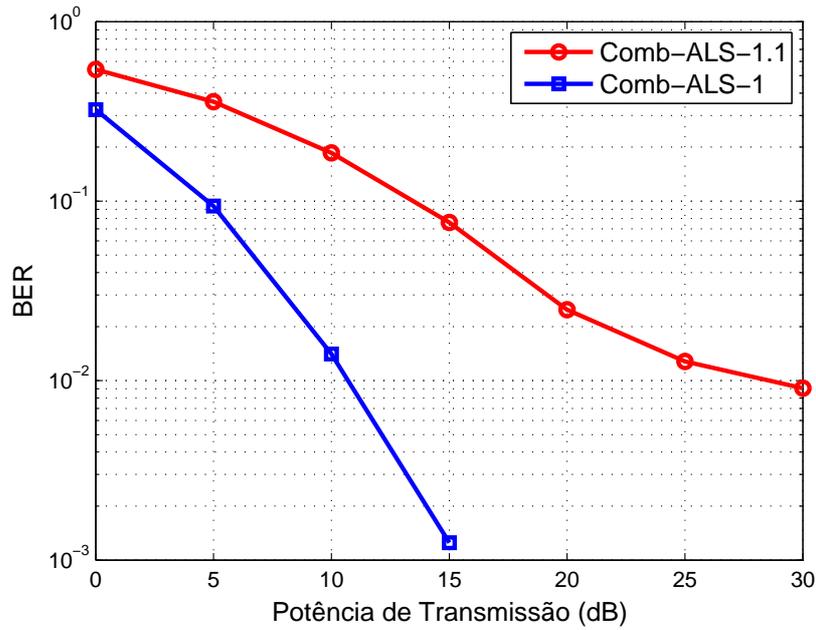


Figura 3.7: BER vs. Potência de transmissão dos dados (Comparação entre os algoritmos Comb-ALS-1 e Comb-ALS-1.1). Potência de transmissão da fase de treinamento = 20dB.

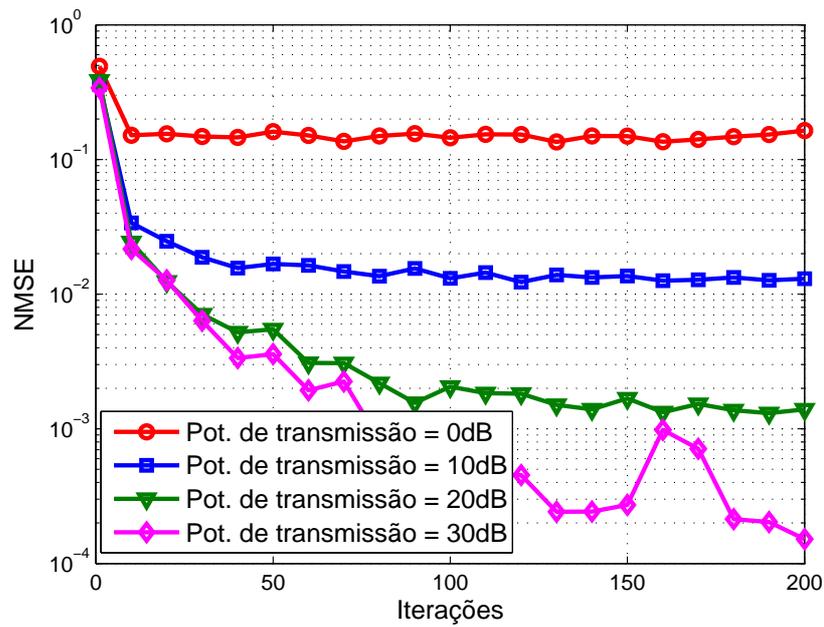


Figura 3.8: NMSE vs. Iterações (Comb-ALS-1).

$$\hat{\mathbf{x}}_{ZF}(t) = \underbrace{\begin{bmatrix} \hat{H}_{r_1d} F \hat{H}_{sr_1} \\ \hat{H}_{r_2d} G \hat{H}_{r_1r_2} F \hat{H}_{sr_1} \end{bmatrix}}_{\hat{H}_{eff}}^{\dagger} \begin{bmatrix} \mathbf{y}_d(2) \\ \mathbf{y}_d(3) \end{bmatrix}. \quad (3.48)$$

A Figura 3.6 compara os algoritmos Comb-ALS-1.1 e

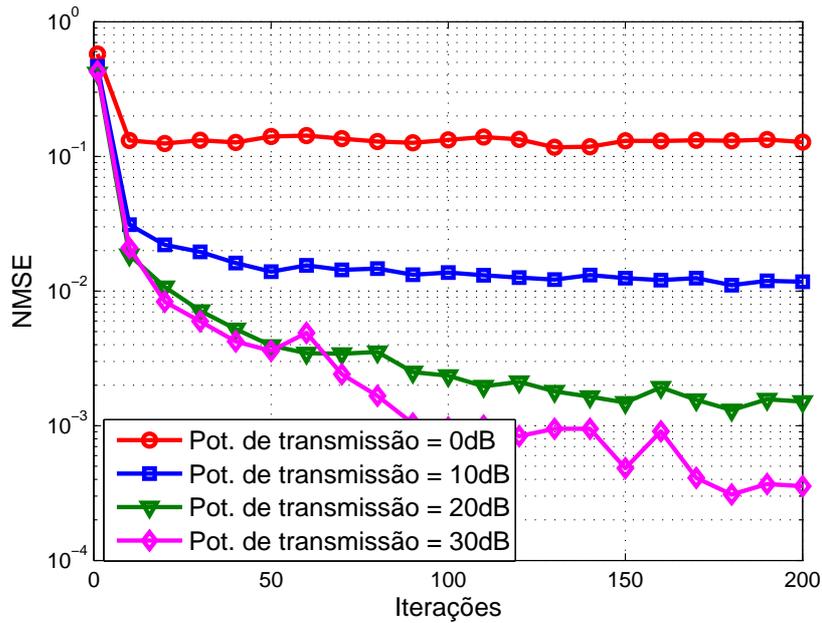


Figura 3.9: NMSE vs. Iterações (Comb-ALS-1.1).

PARAFAC-PARATUCK2 e termos de BER utilizando o receptor ZF dado em (3.48). Como podemos ver, o algoritmo proposto apresenta melhor desempenho do que o algoritmo sequencial PARAFAC-PARATUCK2. Perceba ainda que, em comparação com o algoritmo Comb-ALS-1, o algoritmo Comb-ALS-1.1 apresenta desempenho inferior em termos de NMSE e BER, como podemos ver na Figura 3.7. Isso se deve ao fato de que o algoritmo Comb-ALS-1.1 não faz uso do enlace Fonte-Relay 2-Destino.

As Figuras (3.8) e (3.9) mostra as estimativas de NMSE dos algoritmos Comb-ALS-1 e Comb-ALS-1.1, respectivamente, em função do número de iterações para vários valores de potência de transmissão. Como podemos ver, eles apresentam desempenho semelhante e, para valores de potências de transmissão médias e altas, estes algoritmos convergem em torno de 100 a 150 iterações. Vale ressaltar que o algoritmo Comb-ALS-1 faz a estimação de 5 matrizes de canal ao passo que o algoritmo Comb-ALS-1.1 faz a estimativa de apenas 4.

3.2 Extensão para Matrizes de Amplificação Não-diagonais

O modelo da seção anterior pode ser generalizado para operar com matrizes de amplificação não-diagonais. Desta forma, no processo de amplificação dos sinais nos relays podemos trabalhar com matrizes de amplificação “cheias” (não necessariamente diagonais). Esta estrutura de amplificação é importante, por exemplo, quando se quer otimizar as matrizes de amplificação de relays que operam com uma estrutura de beamforming [34]. Além

disso, podemos trabalhar com matrizes de amplificação não-quadradas. O significado físico deste tipo de matriz de amplificação (não-quadrada) é que os *relays* podem operar com diferentes números de antenas receptoras e transmissoras, assim, por exemplo, no *relay 1*, ao invés de trabalharmos com um conjunto de *relays* com R_1 antenas teremos agora um *relay* com $R_1^{(in)}$ antenas receptoras e $R_1^{(out)}$ antenas transmissoras e a matriz de amplificação do *relay* terá dimensão $R_1^{(out)} \times R_1^{(in)}$. O que será feito nesta seção é uma extensão do algoritmo Comb-ALS-1 para este tipo de esquema de amplificação.

Com essa estrutura de amplificação as equações (3.7)-(3.9) se tornam:

$$\begin{aligned} Y_{d, k(2)} &= \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{H}_{sr_2} \mathbf{X} \\ &+ \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{V}_{r_2}(1) + \mathbf{V}_d(2) \in \mathbb{C}^{M \times L}, \\ &k = 1, \dots, K, \end{aligned} \quad (3.49)$$

$$\begin{aligned} Y_{d, k(3)} &= \mathbf{H}_{r_1d} \mathbf{W}^{(k)} \mathbf{H}_{sr_1} \mathbf{X} \\ &+ \mathbf{H}_{r_1d} \mathbf{W}^{(k)} \mathbf{V}_{r_1}(1) + \mathbf{V}_d(3) \in \mathbb{C}^{M \times L}, \\ &k = 1, \dots, K, \end{aligned} \quad (3.50)$$

$$\begin{aligned} Y_{d, k(4)} &= \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{H}_{r_1r_2} \mathbf{W}^{(k)} \mathbf{H}_{sr_1} \mathbf{X} \\ &+ \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{H}_{r_1r_2} \mathbf{W}^{(k)} \mathbf{V}_{r_1}(1) + \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{V}_{r_2}(2) \\ &+ \mathbf{V}_d(4) \in \mathbb{C}^{M \times L}, \\ &k = 1, \dots, K. \end{aligned} \quad (3.51)$$

Nesta dissertação as matrizes de amplificação $\mathbf{T}^{(k)} \in \mathbb{C}^{R_2^{(out)} \times R_2^{(in)}}$ e $\mathbf{W}^{(k)} \in \mathbb{C}^{R_1^{(out)} \times R_1^{(in)}}$ serão modeladas como o k -ésimo *slice* frontal de tensores que possuem as seguintes decomposições **PARAFAC**:

$$\mathcal{T} = \mathcal{I}_{3, r_T} \times_1 \mathbf{T}_1 \times_2 \mathbf{T}_2 \times_3 \mathbf{T}_3 \in \mathbb{C}^{R_{out}^{(2)} \times R_{in}^{(2)} \times K}, \quad (3.52)$$

$$\mathcal{W} = \mathcal{I}_{3, r_W} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \in \mathbb{C}^{R_{out}^{(1)} \times R_{in}^{(1)} \times K}. \quad (3.53)$$

Onde r_T e r_W são os *ranks* de \mathcal{T} e \mathcal{W} , respectivamente. O *rank* de \mathcal{T} é escolhido como sendo $\max(R_{out}^{(2)}, R_{in}^{(2)})$ e o *rank* de \mathcal{W} é escolhido como sendo $\max(R_{out}^{(1)}, R_{in}^{(1)})$. A Figura 3.10 mostra esses tensores de amplificação. As matrizes $\mathbf{W}_2 \in \mathbb{C}^{R_{in}^{(1)} \times r_W}$ e $\mathbf{W}_1 \in \mathbb{C}^{R_{out}^{(1)} \times r_W}$ ($\mathbf{T}_2 \in \mathbb{C}^{R_{in}^{(2)} \times r_T}$ e $\mathbf{T}_1 \in \mathbb{C}^{R_{out}^{(2)} \times r_T}$) funcionam

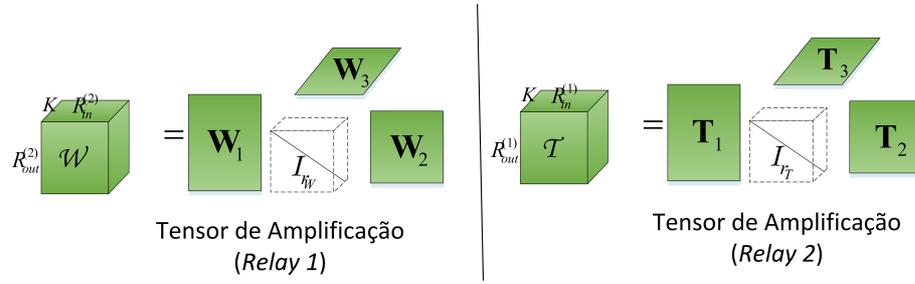


Figura 3.10: Tensores de amplificação do Relay 1 (\mathcal{W}) e do Relay 2 (\mathcal{T}).

como matrizes de seleção das antenas nos *relays*. Nesta dissertação elas são construídas a partir de matrizes identidade: as matrizes \mathbf{W}_1 e \mathbf{W}_2 a partir do truncamento de uma matriz identidade de dimensões $\max(R_{out}^{(2)}, R_{in}^{(2)}) \times \max(R_{out}^{(2)}, R_{in}^{(2)})$ e as matrizes \mathbf{T}_1 e \mathbf{T}_2 a partir do truncamento de uma matriz identidade de dimensões $\max(R_{out}^{(1)}, R_{in}^{(1)}) \times \max(R_{out}^{(1)}, R_{in}^{(1)})$. As matrizes modo-3 \mathbf{W}_3 e \mathbf{T}_3 contêm em suas linhas os fatores de amplificação ao longo da dimensão K . Multiplicando (3.49) e (3.50) à direita por \mathbf{X}^H temos (estamos desprezando as componentes de ruído):

$$\mathbf{P}_k = \mathbf{H}_{r2d} \mathbf{T}^{(k)} \mathbf{H}_{sr2}, \quad (3.54)$$

$$\mathbf{Q}_k = \mathbf{H}_{r1d} \mathbf{T}^{(k)} \mathbf{H}_{sr1}. \quad (3.55)$$

Podemos imaginar (3.54) e (3.55) como sendo *slices* frontais de tensores que possuem as seguintes representações Tucker2:

$$\mathcal{P} = \mathcal{T} \times_1 \mathbf{H}_{r2d} \times_2 \mathbf{H}_{sr2}^T, \quad (3.56)$$

$$\mathcal{Q} = \mathcal{W} \times_1 \mathbf{H}_{r1d} \times_2 \mathbf{H}_{sr1}^T. \quad (3.57)$$

Substituindo (3.52) e (3.53) em (3.56) e (3.57), respectivamente, temos:

$$\mathcal{P} = \mathcal{I}_{3,r_T} \times_1 (\mathbf{H}_{r2d} \mathbf{T}_1) \times_2 (\mathbf{H}_{sr2}^T \mathbf{T}_2) \times_3 \mathbf{T}_3, \quad (3.58)$$

$$\mathcal{Q} = \mathcal{I}_{3,r_W} \times_1 (\mathbf{H}_{r1d} \mathbf{W}_1) \times_2 (\mathbf{H}_{sr1}^T \mathbf{W}_2) \times_3 \mathbf{W}_3. \quad (3.59)$$

Aqui empregamos a propriedade do produto de duas matrizes consecutivas ao longo do mesmo modo (**Definição 8.**).

Como em (3.15-3.16) e (3.22-3.23) temos as seguintes formas matriciadas modo-1 e modo-2 de \mathcal{P} :

$$[\mathcal{P}]_{(1)} = \mathbf{H}_{r_2d} \mathbf{T}_1 (\mathbf{T}_3 \diamond (\mathbf{H}_{sr_2}^T \mathbf{T}_2))^T \in \mathbb{C}^{M \times NK}, \quad (3.60)$$

$$[\mathcal{P}]_{(2)} = \mathbf{H}_{sr_2}^T \mathbf{T}_2 (\mathbf{T}_3 \diamond (\mathbf{H}_{r_2d} \mathbf{T}_1))^T \in \mathbb{C}^{N \times MK}, \quad (3.61)$$

e \mathcal{Q} :

$$[\mathcal{Q}]_{(1)} = \mathbf{H}_{r_1d} \mathbf{W}_1 (\mathbf{W}_3 \diamond (\mathbf{H}_{sr_1}^T \mathbf{W}_2))^T \in \mathbb{C}^{M \times NK}, \quad (3.62)$$

$$[\mathcal{Q}]_{(2)} = \mathbf{H}_{sr_1}^T \mathbf{W}_2 (\mathbf{W}_3 \diamond (\mathbf{H}_{r_1d} \mathbf{W}_1))^T \in \mathbb{C}^{N \times MK}. \quad (3.63)$$

Multiplicando ambos os lados de (3.51) à direita por \mathbf{X}^H , temos (mais uma vez desprezando as componentes de ruído):

$$\mathbf{R}_k = \mathbf{H}_{r_2d} \mathbf{T}^{(k)} \mathbf{H}_{r_1r_2} \mathbf{W}^{(k)} \mathbf{H}_{sr_1}. \quad (3.64)$$

Expressando $\mathbf{T}^{(k)}$ e $\mathbf{W}^{(k)}$ através de suas matrizes fatores (equação (2.13c)), podemos reescrever (3.64) como:

$$\mathbf{R}_k = \mathbf{H}_{r_2d} \mathbf{T}_1 D_k(\mathbf{T}_3) \mathbf{T}_2^T \mathbf{H}_{r_1r_2} \mathbf{W}_1 D_k(\mathbf{W}_3) \mathbf{W}_2^T \mathbf{H}_{sr_1}. \quad (3.65)$$

Esta equação pode ser interpretada como o k -ésimo *slice* frontal de um tensor que possui a seguinte representação *Tucker2*:

$$\mathcal{R} = \mathcal{S} \times_1 \mathbf{H}_{r_2d} \times_2 \mathbf{H}_{sr_1}^T \in \mathbb{C}^{M \times N \times K}, \quad (3.66)$$

em que

$$\begin{aligned} \mathcal{S} = & [\mathbf{T}_1 D_1(\mathbf{T}_3) \mathbf{S} D_1(\mathbf{W}_3) \mathbf{W}_2^T \sqcup_3 \mathbf{T}_1 D_2(\mathbf{T}_3) \mathbf{S} D_2(\mathbf{W}_3) \mathbf{W}_2^T \\ & \dots \sqcup_3 \mathbf{T}_1 D_K(\mathbf{T}_3) \mathbf{S} D_K(\mathbf{W}_3) \mathbf{W}_2^T] \in \mathbb{C}^{R_2 \times R_1 \times K}, \end{aligned} \quad (3.67)$$

e

$$\mathcal{S} = \mathbf{T}_2^T \mathbf{H}_{r_1r_2} \mathbf{W}_1. \quad (3.68)$$

As formas matriciadas modo-1 e modo-2 de (3.66) são dadas,

respectivamente, por:

$$[\mathcal{R}]_{(1)} = \mathbf{H}_{r_2d}[\mathcal{S}]_{(1)}(\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T)^T \in \mathbb{C}^{N \times MK}, \quad (3.69)$$

$$[\mathcal{R}]_{(2)} = \mathbf{H}_{sr_1}^T[\mathcal{S}]_{(2)}(\mathbf{I}_K \otimes \mathbf{H}_{r_2d})^T \in \mathbb{C}^{N \times MK}. \quad (3.70)$$

Como em (3.25), podemos aplicar o operador $\text{vec}(\cdot)$ em (3.65) e definir (aplicando as mesmas operações que foram utilizadas para se chegar em (3.25)):

$$\mathbf{r}_k = ((\mathbf{W}_2^T \mathbf{H}_{sr_1}^T) \otimes (\mathbf{H}_{r_2d} \mathbf{T}_1)) \text{diag}(\text{vec}(\mathcal{S})) (\mathbf{W}_{3(k,:)}^T \otimes \mathbf{T}_{3(k,:)}^T). \quad (3.71)$$

Definindo $\mathbf{R} \doteq [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_K] \in \mathbb{C}^{K \times MN}$, vem

$$\mathbf{R} = ((\mathbf{W}_2^T \mathbf{H}_{sr_1}^T) \otimes (\mathbf{H}_{r_2d} \mathbf{T}_1)) \text{diag}(\text{vec}(\mathcal{S})) (\mathbf{W}_3^T \diamond \mathbf{T}_3^T) \in \mathbb{C}^{K \times MN}. \quad (3.72)$$

Finalmente, aplicando o operador $\text{vec}(\cdot)$ em (3.72), temos:

$$\mathbf{r} \doteq \text{vec}(\mathbf{R}) = \mathbf{Z} \mathbf{h}_{r_1 r_2}, \quad (3.73)$$

em que $\mathbf{Z} \doteq ((\mathbf{W}_3^T \diamond \mathbf{T}_3^T)^T \diamond ((\mathbf{H}_{sr_1}^T \mathbf{W}_2) \otimes (\mathbf{H}_{r_2d} \mathbf{T}_1))) (\mathbf{W}_1^T \otimes \mathbf{T}_2^T)$ e $\mathbf{h}_{r_1 r_2} \doteq \text{vec}(\mathbf{H}_{r_1 r_2})$. Aqui, aplicamos de novo a propriedade $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$.

O processo de estimação das matrizes de canal segue de forma semelhante ao da seção anterior (as estimativas são soluções de problemas de mínimos quadrados). Uma estimativa de $\mathbf{H}_{r_1 r_2}$ segue de (3.73):

$$\hat{\mathbf{h}}_{r_1 r_2} = \mathbf{Z}^\dagger \tilde{\mathbf{r}}, \quad (3.74)$$

em que $\tilde{\mathbf{r}}$ é uma versão de (3.73) com as componentes de ruído. De maneira análoga a (3.32) e (3.33) as estimativas de \mathbf{H}_{r_2d} e \mathbf{H}_{sr_1} são das por:

$$\hat{\mathbf{H}}_{r_2d}^T = \begin{bmatrix} (\mathbf{T}_3 \diamond (\mathbf{H}_{sr_2}^T \mathbf{T}_2)) \mathbf{T}_1^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{sr_1}^T) [\mathcal{S}]_{(1)}^T \end{bmatrix}^\dagger \begin{bmatrix} [\tilde{\mathcal{P}}]_{(1)}^T \\ [\tilde{\mathcal{R}}]_{(1)}^T \end{bmatrix}, \quad (3.75)$$

e

$$\hat{\mathbf{H}}_{sr_1} = \begin{bmatrix} (\mathbf{W}_3 \diamond (\mathbf{H}_{r_1d} \mathbf{W}_1)) \mathbf{W}_2^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{r_2d}) [\mathcal{S}]_{(2)}^T \end{bmatrix}^\dagger \begin{bmatrix} [\tilde{\mathcal{Q}}]_{(2)}^T \\ [\tilde{\mathcal{R}}]_{(2)}^T \end{bmatrix}. \quad (3.76)$$

Da mesma forma, estimativas para \mathbf{H}_{sr_2} e \mathbf{H}_{r_1d} são dadas por (de maneira análoga à (3.34) e (3.35)):

Tabela 3.4: Algoritmo Comb-ALS-2

<p><i>Inicialização:</i> Atribua $i = 0$; Inicialize aleatoriamente $\hat{\mathbf{H}}_{\text{sr}_1}^{(i=0)}$, $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i=0)}$, $\hat{\mathbf{H}}_{\text{sr}_2}^{(i=0)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i=0)}$;</p> <ol style="list-style-type: none"> 1) Atribua $i = i + 1$. 2) Estime $\hat{\mathbf{H}}_{\text{r}_1\text{r}_2}^{(i)}$ usando $\hat{\mathbf{H}}_{\text{sr}_1}^{(i-1)}$ e $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i-1)}$ na equação (3.74); 3) Estime $\hat{\mathbf{H}}_{\text{r}_2\text{d}}^{(i)}$ e $\hat{\mathbf{H}}_{\text{sr}_1}^{(i)}$ usando as equações (3.75) e (3.76); 4) Estime $\hat{\mathbf{H}}_{\text{sr}_2}^{(i)}$ e $\hat{\mathbf{H}}_{\text{r}_1\text{d}}^{(i)}$ usando as equações (3.77) e (3.78); 5) Repita os passos 1 a 4 até a convergência.
--

$$\hat{\mathbf{H}}_{\text{sr}_2}^T = [\tilde{\mathcal{P}}]_{(2)} \left[\mathbf{T}_2 (\mathbf{T}_3 \diamond (\mathbf{H}_{\text{r}_2\text{d}} \mathbf{T}_1))^T \right]^\dagger. \quad (3.77)$$

$$\hat{\mathbf{H}}_{\text{r}_1\text{d}} = [\tilde{\mathcal{Q}}]_{(1)} \left[\mathbf{W}_1 (\mathbf{W}_3 \diamond (\mathbf{H}_{\text{sr}_1}^T \mathbf{W}_2))^T \right]^\dagger. \quad (3.78)$$

O algoritmo de estimação de canal para o caso desta subseção é apresentado na Tabela 3.4 (Algoritmo Comb-ALS-2).

Uma extensão para um protocolo de transmissão bidirecional (*two-way*) é detalhada no Apêndice A. Este modelo também utiliza os tensores de amplificação nos *relays* (tensores \mathcal{T} e \mathcal{W}) e explora, de forma semelhante aos algoritmos desta seção, todos os conjuntos de dados recebidos pelo terminal final.

3.2.1 Simulações e Resultados

Como na seção anterior, assumimos um sistema em que a fonte e o destino têm $N = 4$ e $M = 4$ antenas, respectivamente. Também aqui o sistema de comunicação é assistido por 2 grupos de *relays* (Figura 3.1). O algoritmo Comb-ALS-2 é comparado com o estimador sequencial SVD-LS introduzido na seção anterior (nas simulações a seguir o algoritmo Comb-ALS-1 é equivalente ao algoritmo Comb-ALS-2 no caso em que temos $R_{\text{in}}^{(i)} = R_{\text{out}}^{(i)}$, $i = 1, 2$).

O desempenho da estimação dos canais é medida em termos da NMSE do canal efetivo (equação 3.79). Assim como em (3.46), foram realizadas 5000 rodadas de Monte-Carlo, em que, para cada rodada, diferentes matrizes de canal e de ruído foram geradas.

A Figura 3.11 mostra a NMSE do canal efetivo em função da potência de transmissão. Como podemos ver, o método Comb-ALS-2 consegue melhores estimativas de NMSE quando comparado com o estimador SVD-LS. Observe ainda que o algoritmo SVD-LS, não pode operar com seleção de antenas nos *relays*, desta forma, o número de antenas de entrada e saída para este algoritmo devem ser iguais (neste caso estamos usando $R_{\text{(in)}}^{(i)} = R_{\text{(out)}}^{(i)} = 4$, para

$i = 1, 2$). Nesta configuração de antenas, o algoritmo Comb-ALS-2 apresenta um ganho médio em torno de 10dB em comparação ao algoritmo **SVD-LS**.

A Figura 3.12 mostra os resultado de **BER** em função da potência de transmissão dos dados para uma potência de transmissão da fase de treinamento igual a 20dB para a configuração de antenas $R_{in}^{(1)} = R_{out}^{(1)} = R_{in}^{(2)} = R_{out}^{(2)} = 4$. As curvas em vermelho e azul mostram as estimativas de **BER** para os algoritmos **SVD-LS** e Comb-ALS-2, respectivamente, utilizando o receptor **ZF** dado pela equação a seguir:

$$\hat{\mathbf{x}}_{ZF} = \left[\begin{array}{c} \hat{\mathbf{H}}_{r2d} \mathbf{G} \hat{\mathbf{H}}_{sr2} \\ \hat{\mathbf{H}}_{r1d} \mathbf{F} \hat{\mathbf{H}}_{sr1} \\ \underbrace{\hat{\mathbf{H}}_{r2d} \mathbf{G} \hat{\mathbf{H}}_{r1r2} \mathbf{F} \hat{\mathbf{H}}_{sr1}}_{\hat{\mathbf{H}}_{\text{eff}}} \end{array} \right]^{\dagger} \begin{bmatrix} \mathbf{y}_d(2) \\ \mathbf{y}_d(3) \\ \mathbf{y}_d(4) \end{bmatrix}, \quad (3.79)$$

Observe que o algoritmo Comb-ALS-2 apresenta um desempenho superior ao estimador **SVD-LS**, sendo próximo ao caso de perfeito conhecimento do canal (Perfect **CSI**). A fim de demonstrar a importância de se utilizar todos os enlaces para estimar os símbolos, a Figura 3.12 mostra ainda as estimativas de **BER** para o receptor **ZF** da equação (3.79) suprimindo-se o último enlace, ou seja, o receptor **ZF** é dado por:

$$\hat{\mathbf{x}}_{ZF}(t) = \left[\begin{array}{c} \hat{\mathbf{H}}_{r2d} \mathbf{G} \hat{\mathbf{H}}_{sr2} \\ \hat{\mathbf{H}}_{r1d} \mathbf{F} \hat{\mathbf{H}}_{sr1} \\ \underbrace{\hspace{10em}}_{\hat{\mathbf{H}}_{\text{eff}}} \end{array} \right]^{\dagger} \begin{bmatrix} \mathbf{y}_d(t+1) \\ \mathbf{y}_d(t+2) \end{bmatrix}, \quad (3.80)$$

Por utilizar somente os enlaces de 2 saltos ($\hat{\mathbf{H}}_{r2d} \mathbf{G} \hat{\mathbf{H}}_{sr2}$ e $\hat{\mathbf{H}}_{r1d} \mathbf{F} \hat{\mathbf{H}}_{sr1}$) os resultados de **BER** na Figura 3.12 são apresentados com o nome Comb-ALS-2(*two-hop*) e **SVD-LS** (*two-hop*) (curvas em azul claro e magenta). Como podemos ver, existe um ganho considerável quando se usa todos os enlaces disponíveis para se estimar os símbolos.

Finalmente, o método proposto continua sendo válido em outros cenários em que o número de antenas é diferente do número de blocos de tempo (desde que as condições de identificabilidade sejam satisfeitas), ao contrário do método **SVD-LS** que exige $K = R_1 = R_2$.

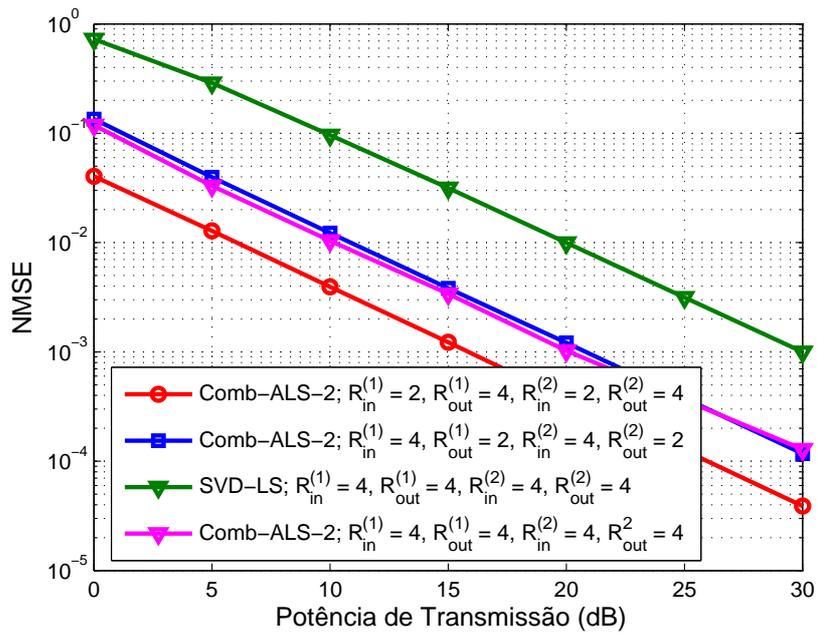


Figura 3.11: NMSE vs. Potência de transmissão para $N = M = K = 4$ (Comparação entre o algoritmo Comb-ALS-2 e SVD-LS).

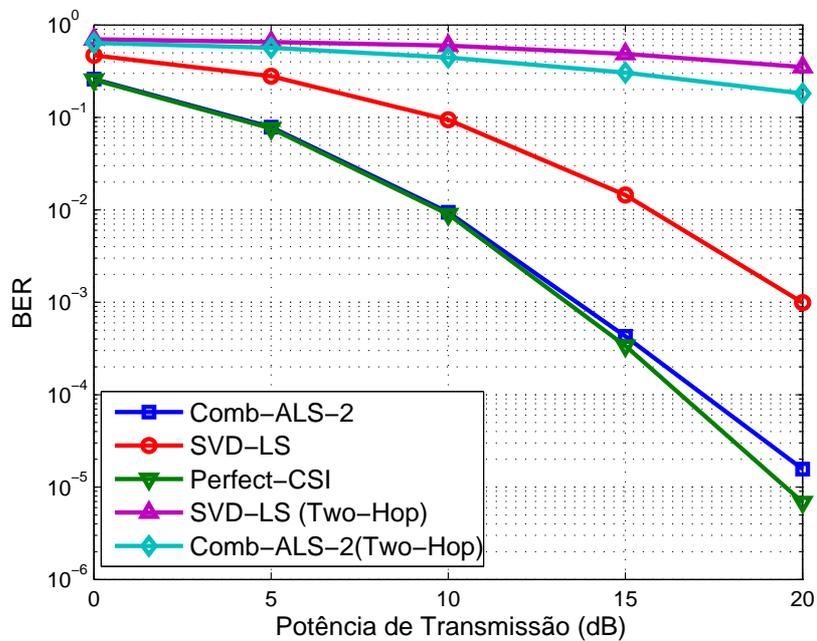


Figura 3.12: Desempenho de BER como o receptor ZF em função da potência de transmissão dos dados (Comparação entre Comb-ALS-2 e SVD-LS para $R_{in}^{(1)} = R_{out}^{(1)} = R_{in}^{(2)} = R_{out}^{(2)} = 4$). Potência de transmissão na fase de treinamento = 20dB.

Conclusões e Perspectivas

Nesta dissertação foram investigados algoritmos de estimação de canal para uma rede **MIMO** cooperativa com 3 saltos. Foram propostos métodos de estimação assistidos por sequências de treinamento. Estes métodos utilizam algoritmos iterativos baseados em **ALS** para estimar as matrizes de canal, explorando todos os sinais dos diversos enlaces de chegada. Os dados recebidos podem ser modelados como tensores que, por sua vez, podem ser combinados de forma a se estimar matrizes de canal que são comuns entre eles.

O primeiro algoritmo de estimação (Comb-ALS-1), utiliza um protocolo de transmissão em 4 fases e pode estimar todas as matrizes de canal. Um algoritmo em 3 fases para o mesmo cenário foi proposto a fim de diminuir o tempo da fase de treinamento (Comb-ALS-1.1). Através dos resultados de simulações, foi verificado que o algoritmo (Comb-ALS-1.1) possui um desempenho inferior ao algoritmo (Comb-ALS-1). Essa diferença de desempenho se deve à diminuição do número de enlaces utilizados na estimação das matrizes de canal do algoritmo (Comb-ALS-1.1).

Em um segundo momento, foi considerado que as matrizes de amplificação no *relays* podem ser não-diagonais. Desta forma, os *relays* podem operar em um esquema de amplificação mais geral. Além disso, as matrizes de amplificação podem funcionar como matrizes de seleção das antenas nos *relays*, desta forma, os *relays* podem operar com diferentes conjuntos de antenas transmissoras e receptoras. Além desses algoritmos para o caso *one-way*, foi proposto um algoritmo para o caso *two-way*, ele está detalhado no Apêndice **A**.

Os métodos propostos fazem uso eficiente da diversidade cooperativa e exploram de forma rica todos os conjuntos de dados no receptor final. Apesar disso, existem alguns pontos que podem ser analisados em trabalhos futuros, dentre eles podemos abordar questões como:

- ▶ Diminuição do número de blocos de tempo afim de minimizar o tempo da fase de treinamento, ou propor uma técnica cega de estimação conjunta de canais e símbolos.
- ▶ Extensão do modelo proposto para o caso geral com N saltos.
- ▶ Uso de *relays* decodifica-e-encaminha (DF, do inglês Decode-and-Forward) em alguns saltos da rede e análise do impacto do seu uso na estimação de símbolos e na latência da rede.
- ▶ Otimização das matrizes de amplificação utilizando métricas como LS ou MMSE.

Receptor para o Modelo *Two-Way*

Neste apêndice, descreveremos o algoritmo para o cenário em que os dados trafegam de forma bidirecional (*Two-way*). Este cenário consiste de dois terminais que enviam seus dados através da rede cooperativa de forma semelhante ao modelo do Capítulo 3. A diferença aqui é que os dados trafegam nas duas direções (ver Figura A.1). Os mesmos parâmetros de perda de percurso e potência de ruído adotados no modelo do Capítulo 3 serão utilizados aqui. Além disso, os *relays* desta rede ainda operam no modo *AF half-duplex*.

O algoritmo de estimação das matrizes de canal também segue o modelo do Capítulo 3, ou seja, as estimativas das matrizes de canal serão alcançadas através da solução de problemas de mínimos quadrados. Faremos a descrição deste algoritmo apenas para um dos terminais, já que, devido a simetria do problema, todas as equações desenvolvidas para o Terminal 1 são análogas às equações do Terminal 2 (ver a descrição do protocolo de transmissão, Seção A.1).

A.1 Descrição do Protocolo de Transmissão

O protocolo de transmissão é dividido em cinco fases. Na primeira, fase os terminais 1 e 2 transmitem ao mesmo tempo, os vetores de sinal $\mathbf{x}_1 \in \mathbb{C}^{N \times 1}$ e $\mathbf{x}_2 \in \mathbb{C}^{M \times 1}$, respectivamente. Estes sinais são recebidos pelo primeiro e segundo *relays* e são dados em banda base discreta respectivamente por:

$$\mathbf{y}_{r_1}(1) = \mathbf{H}_{t_1r_1}\mathbf{x}_1 + \mathbf{H}_{t_2r_1}\mathbf{x}_2 + \mathbf{v}_{r_1}(1) \in \mathbb{C}^{R_1 \times 1}, \quad (\text{A.1})$$

$$\mathbf{y}_{r_2}(1) = \mathbf{H}_{t_2r_2}\mathbf{x}_2 + \mathbf{H}_{t_1r_2}\mathbf{x}_1 + \mathbf{v}_{r_2}(1) \in \mathbb{C}^{R_2 \times 1}, \quad (\text{A.2})$$

em que $\mathbf{H}_{t_i r_j}$ é a matriz que modela o canal entre o terminal i e o *relay* j e $\mathbf{v}_{r_j}(1)$ é o ruído aditivo Gaussiano no grupo de *relays* j . Na segunda fase, o

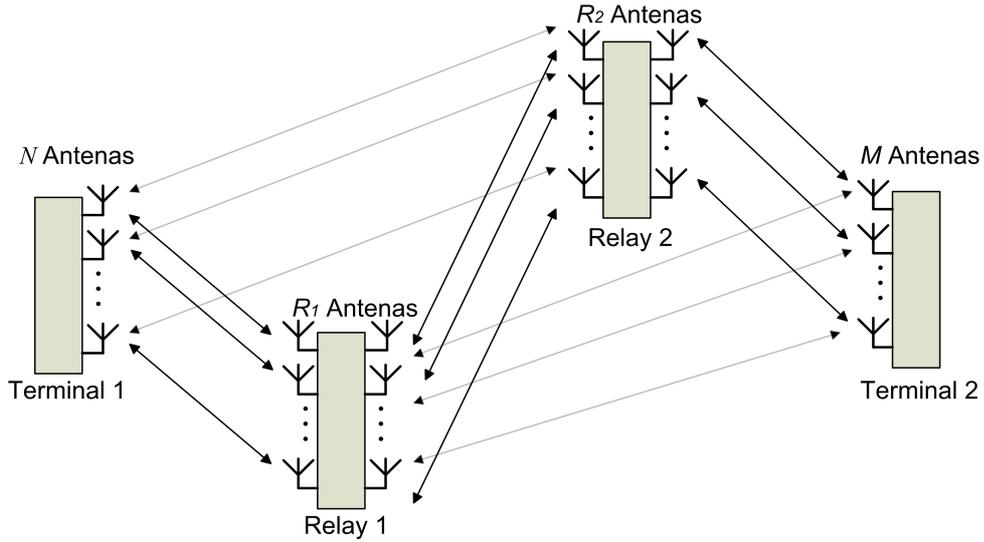


Figura A.1: Modelo básico da rede cooperativa. Terminal 1 com N antenas, Terminal 2 com M antenas e dois relays AF.

relay 1 amplifica o sinal (A.1) multiplicando-o por uma matriz de amplificação W e retransmite este sinal amplificado para os outros nós da rede. Os sinais recebidos pelo terminal 1 e pelo *relay 2* são dados respectivamente por:

$$\mathbf{y}_{t_1}(2) = \mathbf{H}_{t_1r_1}^T \mathbf{W} \mathbf{H}_{t_1r_1} \mathbf{x}_1 + \mathbf{H}_{t_1r_1}^T \mathbf{W} \mathbf{H}_{t_2r_1} \mathbf{x}_2 + \mathbf{H}_{t_1r_1}^T \mathbf{W} \mathbf{v}_{r_1}(1) + \mathbf{v}_{t_1}(2) \in \mathbb{C}^{N \times 1}, \quad (\text{A.3})$$

$$\mathbf{y}_{r_2}(2) = \mathbf{H}_{r_1r_2} \mathbf{W} \mathbf{H}_{t_1r_1} \mathbf{x}_1 + \mathbf{H}_{r_1r_2} \mathbf{W} \mathbf{H}_{t_2r_1} \mathbf{x}_2 + \mathbf{H}_{r_1r_2} \mathbf{W} \mathbf{v}_{r_1}(1) + \mathbf{v}_{r_2}(2) \in \mathbb{C}^{R_2 \times 1}. \quad (\text{A.4})$$

Na fase três, o *relay 2* amplifica o sinal (A.2) multiplicando-o por uma matriz de amplificação T e faz o *broadcast* do sinal resultante. Os sinais recebidos pelo terminal 1 e pelo *relay 1* são dados respectivamente por:

$$\mathbf{y}_{t_1}(3) = \mathbf{H}_{t_1r_2}^T \mathbf{T} \mathbf{H}_{t_2r_2} \mathbf{x}_2 + \mathbf{H}_{t_1r_2}^T \mathbf{T} \mathbf{H}_{t_1r_2} \mathbf{x}_1 + \mathbf{H}_{t_1r_2}^T \mathbf{T} \mathbf{v}_{r_2}(1) + \mathbf{v}_{t_1}(3) \in \mathbb{C}^{N \times 1}, \quad (\text{A.5})$$

$$\mathbf{y}_{r_1}(3) = \mathbf{H}_{r_1r_2}^T \mathbf{T} \mathbf{H}_{t_2r_2} \mathbf{x}_2 + \mathbf{H}_{r_1r_2}^T \mathbf{T} \mathbf{H}_{t_1r_2} \mathbf{x}_1 + \mathbf{H}_{r_1r_2}^T \mathbf{T} \mathbf{v}_{r_2}(1) + \mathbf{v}_{r_1}(3) \in \mathbb{C}^{R_1 \times 1}. \quad (\text{A.6})$$

Na fase quatro, o *relay 1* amplifica o sinal (A.6) multiplicando-o pela mesma matriz de amplificação W e o retransmite. Finalmente, na fase cinco o *relay 2* faz um procedimento análogo multiplicando (A.4) por sua matriz de amplificação T e retransmite o sinal resultante para os outros nós da rede. Os

sinais recebidos pelo terminal 1 nestas duas fases são dados respectivamente por:

$$\begin{aligned} \mathbf{y}_{t_1}(4) = & \mathbf{H}_{t_1 r_1}^T \mathbf{W} \mathbf{H}_{r_1 r_2}^T \mathbf{T} \mathbf{H}_{t_2 r_2} \mathbf{x}_2 + \mathbf{H}_{t_1 r_1}^T \mathbf{W} \mathbf{H}_{r_1 r_2}^T \mathbf{T} \mathbf{H}_{t_1 r_2} \mathbf{x}_1 + \\ & \mathbf{H}_{t_1 r_1}^T \mathbf{W} \mathbf{H}_{r_1 r_2}^T \mathbf{T} \mathbf{v}_{r_2}(1) + \mathbf{H}_{t_1 r_1}^T \mathbf{W} \mathbf{v}_{t_2}(3) + \mathbf{v}_{t_1}(4) \in \mathbb{C}^{N \times 1}, \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \mathbf{y}_{t_1}(5) = & \mathbf{H}_{t_1 r_2}^T \mathbf{T} \mathbf{H}_{r_1 r_2} \mathbf{W} \mathbf{H}_{t_1 r_1} \mathbf{x}_1 + \mathbf{H}_{t_1 r_2}^T \mathbf{T} \mathbf{H}_{r_1 r_2} \mathbf{W} \mathbf{H}_{t_2 r_1} \mathbf{x}_2 + \\ & \mathbf{H}_{t_1 r_2}^T \mathbf{T} \mathbf{H}_{r_1 r_2} \mathbf{W} \mathbf{v}_{r_1}(1) + \mathbf{H}_{t_1 r_2}^T \mathbf{T} \mathbf{v}_{r_2}(2) + \mathbf{v}_{t_1}(5) \in \mathbb{C}^{N \times 1}. \end{aligned} \quad (\text{A.8})$$

A.2 Formulação Tensorial

Da mesma forma que no Capítulo 3, podemos dividir a fase de treinamento em K blocos de tempo e, para cada um desses blocos de tempo, a mesma matriz de treinamento $\mathbf{X}_1 = [\mathbf{x}_{1,1}, \dots, \mathbf{x}_{1,L}] \in \mathbb{C}^{N \times L}$ ortogonal é transmitida pelo terminal 1 e $\mathbf{X}_2 = [\mathbf{x}_{2,1}, \dots, \mathbf{x}_{2,L}] \in \mathbb{C}^{M \times L}$ é transmitida pelo terminal 2 obedecendo o protocolo de transmissão descrito acima. Assim, podemos escrever as equações (A.3), (A.5), (A.7) e (A.8) na forma matricial como segue:

$$\mathbf{Y}_{t_1,k}(2) = \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{H}_{t_1 r_1} \mathbf{X}_1 + \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{H}_{t_2 r_1} \mathbf{X}_2 + \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{V}_{r_1}(1) + \mathbf{V}_{t_1}(2) \in \mathbb{C}^{N \times L}, \quad (\text{A.9})$$

$$\mathbf{Y}_{t_1,k}(3) = \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{t_2 r_2} \mathbf{X}_2 + \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{t_1 r_2} \mathbf{X}_1 + \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{V}_{r_2}(1) + \mathbf{V}_{t_2}(3) \in \mathbb{C}^{N \times L}, \quad (\text{A.10})$$

$$\begin{aligned} \mathbf{Y}_{t_1,k}(4) = & \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{H}_{r_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{t_2 r_2} \mathbf{X}_2 + \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{H}_{r_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{t_1 r_2} \mathbf{X}_1 + \\ & \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{H}_{r_1 r_2}^T \mathbf{T}^{(k)} \mathbf{V}_{r_2}(1) + \mathbf{H}_{t_1 r_1}^T \mathbf{W}^{(k)} \mathbf{V}_{t_2}(3) + \\ & \mathbf{V}_{t_1}(4) \in \mathbb{C}^{N \times L}, \end{aligned} \quad (\text{A.11})$$

$$\begin{aligned} \mathbf{Y}_{t_1,k}(5) = & \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{r_1 r_2} \mathbf{W}^{(k)} \mathbf{H}_{t_1 r_1} \mathbf{X}_1 + \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{r_1 r_2} \mathbf{W}^{(k)} \mathbf{H}_{t_2 r_1} \mathbf{X}_2 + \\ & \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{H}_{r_1 r_2} \mathbf{W}^{(k)} \mathbf{V}_{r_1}(1) + \mathbf{H}_{t_1 r_2}^T \mathbf{T}^{(k)} \mathbf{V}_{r_2}(2) + \\ & \mathbf{V}_{t_1}(5) \in \mathbb{C}^{N \times L}, k = 1, \dots, K, \end{aligned} \quad (\text{A.12})$$

em que as duas primeiras parcelas em (A.9-A.12) são as componentes de sinal e as restantes são as componentes de ruído.

Como na seção anterior, nossas matrizes de amplificação $\mathbf{W}^{(k)}$ e $\mathbf{T}^{(k)}$ correspondem ao k th slice frontal de tensores PARAFAC que são dados pelas (3.52) e (3.53) (aqui estamos considerando que, nos *relays*, o mesmo conjunto de antenas faz a transmissão e a recepção dos sinais, i.e, $R_i^{(in)} = R_i^{(out)} = R_i$, para $i = 1, 2$):

$$\mathbf{W} = \mathcal{I}_{3,r_W} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \in \mathbb{C}^{R_1 \times R_1 \times K}, \quad (\text{A.13})$$

$$\mathcal{T} = \mathcal{I}_{3,r_T} \times_1 \mathbf{T}_1 \times_2 \mathbf{T}_2 \times_3 \mathbf{T}_3 \in \mathbb{C}^{R_2 \times R_2 \times K}. \quad (\text{A.14})$$

Assim como em [11], as matrizes fatores \mathbf{W}_i e \mathbf{T}_i , $i = 1, 2$ e 3 são construídas a partir de matrizes DFT normalizadas.

Introduzindo:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{t_1 r_1} & \mathbf{H}_{t_2 r_1} \\ \mathbf{H}_{t_1 r_2} & \mathbf{H}_{t_2 r_2} \end{bmatrix} \in \mathbb{C}^{(R_1+R_2) \times (N+M)}, \quad (\text{A.15})$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \in \mathbb{C}^{(N+M) \times L}, \quad (\text{A.16})$$

em que a matriz (A.16) possui linhas ortonormais, ou seja, as relações

$$\mathbf{X}_1^* \mathbf{X}^T = \mathbf{X}_1^* \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{I}_N & \mathbf{0}_{N \times M} \end{bmatrix}$$

e

$$\mathbf{X}_2^* \mathbf{X}^T = \mathbf{X}_2^* \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{M \times N} & \mathbf{I}_M \end{bmatrix},$$

são válidas. Podemos construir \mathbf{X} a partir de uma matriz Discrete Fourier Transform (DFT) normalizada (esta matriz é quadrada, portanto, aqui utilizamos $L = M + N$).

Assim como foi feito no Capítulo 3, podemos concatenar as equações (A.9) em um tensor

$$\mathcal{Y}_{t_1}(2) = [\mathbf{Y}_{t_1,1}(2) \sqcup_3 \dots \sqcup_3 \mathbf{Y}_{t_1,K}(2)] \in \mathbb{C}^{N \times L \times K}. \quad (\text{A.17})$$

Inspecionando as equações (A.9), podemos caracterizar (A.17) como um tensor *Tucker2* com a seguinte representação:

$$\mathcal{Y}_{t_1}(2) = \mathcal{W} \times_1 \mathbf{H}_{t_1 r_1}^T \times_2 (\Psi_{R_1} \mathbf{H} \mathbf{X})^T + \mathcal{V}(2), \quad (\text{A.18})$$

em que $\mathcal{V}(2)$ é um tensor que contém todas as parcelas de ruído em (A.9) e Ψ_{R_1} é uma matriz construída tomando as R_1 primeiras linhas de uma matriz

identidade $(R_1 + R_2) \times (R_1 + R_2)$ (esta matriz faz a seleção das R_1 primeiras linhas de H). Com um procedimento similar podemos introduzir:

$$\mathcal{Y}_{t_1}(3) = [\mathbf{Y}_{t_1,1}(3) \sqcup_3 \dots \sqcup_3 \mathbf{Y}_{t_1,K}(3)] \in \mathbb{C}^{N \times L \times K}. \quad (\text{A.19})$$

Que tem possui a seguinte representação *Tucker2*:

$$\mathcal{Y}_{t_1}(3) = \mathcal{T} \times_1 \mathbf{H}_{t_1 r_2}^T \times_2 (\Phi_{R_2} \mathbf{H} \mathbf{X})^T + \mathcal{V}(3), \quad (\text{A.20})$$

em que Φ_{R_2} é uma matriz construída tomando as R_2 últimas linhas de uma matriz identidade $(R_1 + R_2) \times (R_1 + R_2)$ (esta matriz seleciona as R_2 últimas linhas de H).

Como $\mathbf{W}^{(k)}$ e $\mathbf{T}^{(k)}$ são *slices* frontais de um tensor **PARAFAC**, podemos reescrever (A.11)(as componentes de ruído foram omitidas) como:

$$\begin{aligned} \mathbf{Y}_{t_1}^{(k)}(4) = & \mathbf{H}_{t_1 r_1}^T \mathbf{W}_1 D_k(\mathbf{W}_3) \mathbf{W}_2^T \mathbf{H}_{r_1 r_2}^T \mathbf{T}_1 D_k(\mathbf{T}_3) \mathbf{T}_2^T \mathbf{H}_{t_2 r_2} \mathbf{X}_2 + \\ & \mathbf{H}_{t_1 r_1}^T \mathbf{W}_1 D_k(\mathbf{W}_3) \mathbf{W}_2^T \mathbf{H}_{r_1 r_2}^T \mathbf{T}_1 D_k(\mathbf{T}_3) \mathbf{T}_2^T \mathbf{H}_{t_1 r_2} \mathbf{X}_1. \end{aligned} \quad (\text{A.21})$$

Procedendo de uma forma similar a (A.17) e (A.19) podemos, a partir das equações (A.21), construir o tensor

$$\mathcal{Y}_{t_1}(4) = [\mathbf{Y}_{t_1,1}(4) \sqcup_3 \dots \sqcup_3 \mathbf{Y}_{t_1,K}(4)] \in \mathbb{C}^{N \times L \times K}. \quad (\text{A.22})$$

Cuja representação *Tucker2* é:

$$\mathcal{Y}_{t_1}(4) = \mathcal{A} \times_1 \mathbf{H}_{t_1 r_1}^T \times_2 (\Phi_{R_2} \mathbf{H} \mathbf{X})^T + \mathcal{V}(4), \quad (\text{A.23})$$

em que

$$\mathcal{A} = [\mathbf{W}_1 D_1(\mathbf{W}_3) \mathbf{R} D_1(\mathbf{T}_3) \mathbf{T}_2^T \sqcup_3 \dots \sqcup_3 \mathbf{W}_1 D_K(\mathbf{W}_3) \mathbf{R} D_K(\mathbf{T}_3) \mathbf{T}_2^T] \in \mathbb{C}^{R_1 \times R_2 \times K} \quad (\text{A.24})$$

e

$$\mathbf{R} = \mathbf{W}_2^T \mathbf{H}_{r_1 r_2}^T \mathbf{T}_1. \quad (\text{A.25})$$

Finalmente, de forma similar a (A.22), podemos a partir das equações (A.12) construir o seguinte tensor:

$$\mathcal{Y}_{t_1}(5) = [\mathbf{Y}_{t_1,1}(5) \sqcup_3 \dots \sqcup_3 \mathbf{Y}_{t_1,K}(5)] \in \mathbb{C}^{N \times L \times K}. \quad (\text{A.26})$$

Cuja representação *Tucker2* é dada por:

$$\mathcal{Y}_{t_1}(5) = \mathcal{B} \times_1 \mathbf{H}_{t_1 r_2}^T \times_2 (\Psi_{R_1} \mathbf{H} \mathbf{X})^T + \mathcal{V}(5), \quad (\text{A.27})$$

em que

$$\begin{aligned} \mathcal{B} = [& \mathbf{T}_1 D_1(\mathbf{T}_3) S D_1(\mathbf{W}_3) \mathbf{W}_2^T \sqcup_3 \mathbf{T}_1 D_2(\mathbf{T}_3) S D_2(\mathbf{W}_3) \mathbf{W}_2^T \\ & \dots \sqcup_3 \mathbf{T}_1 D_K(\mathbf{T}_3) S D_K(\mathbf{W}_3) \mathbf{W}_2^T] \in \mathbb{C}^{R_2 \times R_1 \times K} \end{aligned} \quad (\text{A.28})$$

e

$$\mathbf{S} = \mathbf{T}_2^T \mathbf{H}_{r_1 r_2} \mathbf{W}_1. \quad (\text{A.29})$$

A.3 Algoritmo de Estimação de Canal

► Estimação de $\mathbf{H}_{t_1 r_2}$

Tomando o *unfolding* modo-2 de (A.23), temos

$$\begin{aligned} [\mathcal{Y}_{t_1}(4)]_{(2)} &= (\mathbf{X}^T \mathbf{H}^T \Phi_{R_2}^T) [\mathcal{A}]_{(2)} (\mathbf{H}_{t_1 r_1}^T \otimes \mathbf{I}_K)^T = \\ &= \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{t_1 r_2}^T \\ \mathbf{H}_{t_2 r_2}^T \end{bmatrix} [\mathcal{A}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T. \end{aligned} \quad (\text{A.30})$$

Pré-multiplicando (A.30) por \mathbf{X}_1^* obtemos

$$\mathbf{X}_1^* [\mathcal{Y}_{t_1}(4)]_{(2)} = \mathbf{H}_{t_1 r_2}^T [\mathcal{A}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T. \quad (\text{A.31})$$

Tomando o *unfolding* modo-1 de (A.27), temos

$$[\mathcal{Y}_{t_1}(5)]_{(1)} = \mathbf{H}_{t_1 r_2}^T [\mathcal{B}]_{(1)} (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T)^T. \quad (\text{A.32})$$

Finalmente, tomando o *unfolding* modo-1 de (A.20)

$$[\mathcal{Y}_{t_1}(3)]_{(1)} = \mathbf{H}_{t_1 r_2}^T [\mathcal{T}]_{(1)} (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T)^T. \quad (\text{A.33})$$

De (A.31-A.33), uma estimativa de $\mathbf{H}_{t_1 r_2}$ pode ser obtida resolvendo o problema de mínimos quadrados

$$\arg \min_{\mathbf{H}_{t_1 r_2}} \left\| \begin{bmatrix} (\mathbf{X}_1^* [\mathcal{Y}_{t_1}(4)]_{(2)})^T \\ ([\mathcal{Y}_{t_1}(5)]_{(1)})^T \\ ([\mathcal{Y}_{t_1}(3)]_{(1)})^T \end{bmatrix} - \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T) [\mathcal{A}]_{(2)}^T \\ (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T) [\mathcal{B}]_{(1)}^T \\ (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T) [\mathcal{T}]_{(1)}^T \end{bmatrix} \mathbf{H}_{t_1 r_2} \right\|_F^2,$$

cuja solução é dada por:

$$\hat{\mathbf{H}}_{t_1 r_2} = \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T) [\mathcal{A}]_{(2)}^T \\ (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T) [\mathcal{B}]_{(1)}^T \\ (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T) [\mathcal{T}]_{(1)}^T \end{bmatrix}^\dagger \begin{bmatrix} (\mathbf{X}_1^* [\mathcal{Y}_{t_1}(4)]_{(2)})^T \\ ([\mathcal{Y}_{t_1}(5)]_{(1)})^T \\ ([\mathcal{Y}_{t_1}(3)]_{(1)})^T \end{bmatrix}. \quad (\text{A.34})$$

► Estimação de $\mathbf{H}_{t_1 r_1}$

Tomando o *unfolding* modo-2 de (A.27), temos

$$\begin{aligned} [\mathcal{Y}_{t_1}(5)]_{(2)} &= (\mathbf{X}^T \mathbf{H}^T \Psi_{R_1}^T) [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T = \\ &= \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{t_1 r_1}^T \\ \mathbf{H}_{t_2 r_1}^T \end{bmatrix} [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T. \end{aligned} \quad (\text{A.35})$$

Pré-multiplicando (A.35) por \mathbf{X}_1^* obtemos

$$\mathbf{X}_1^* [\mathcal{Y}_{t_1}(5)]_{(2)} = \mathbf{H}_{t_1 r_1}^T [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T. \quad (\text{A.36})$$

Tomando o *unfolding* modo-1 de (A.23), temos

$$[\mathcal{Y}_{t_1}(4)]_{(1)} = \mathbf{H}_{t_1 r_1}^T [\mathcal{A}]_{(1)} (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T)^T. \quad (\text{A.37})$$

Finalmente, tomando o *unfolding* modo-1 de (A.18), temos

$$[\mathcal{Y}_{t_1}(2)]_{(1)} = \mathbf{H}_{t_1 r_1}^T [\mathcal{W}]_{(1)} (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T)^T. \quad (\text{A.38})$$

De (A.36)-(A.38) uma estimativa de $\mathbf{H}_{t_1 r_1}$ pode ser obtida resolvendo o seguinte problema

$$\arg \min_{\mathbf{H}_{t_1 r_1}} \left\| \begin{bmatrix} (\mathbf{X}_1^* [\mathcal{Y}_{t_1}(5)]_{(2)})^T \\ ([\mathcal{Y}_{t_1}(4)]_{(1)})^T \\ ([\mathcal{Y}_{t_1}(2)]_{(1)})^T \end{bmatrix} - \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T) [\mathcal{B}]_{(2)}^T \\ (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T) [\mathcal{A}]_{(1)}^T \\ (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T) [\mathcal{W}]_{(1)}^T \end{bmatrix} \mathbf{H}_{t_1 r_1} \right\|_F^2,$$

cuja solução é dada por:

$$\hat{\mathbf{H}}_{t_1 r_1} = \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T) [\mathcal{B}]_{(2)}^T \\ (\mathbf{I}_K \otimes (\Phi_{R_2} \mathbf{H} \mathbf{X})^T) [\mathcal{A}]_{(1)}^T \\ (\mathbf{I}_K \otimes (\Psi_{R_1} \mathbf{H} \mathbf{X})^T) [\mathcal{W}]_{(1)}^T \end{bmatrix}^\dagger \begin{bmatrix} (\mathbf{X}_1^* [\mathcal{Y}_{t_1}(5)]_{(2)})^T \\ ([\mathcal{Y}_{t_1}(4)]_{(1)})^T \\ ([\mathcal{Y}_{t_1}(2)]_{(1)})^T \end{bmatrix}. \quad (\text{A.39})$$

► Estimação de $\mathbf{H}_{t_2 r_1}$

Tomando o *unfolding* modo-2 de (A.27), temos

$$\begin{aligned} [\mathbf{y}_{t_1}(5)]_{(2)} &= (\mathbf{X}^T \mathbf{H}^T \Psi_{R_1}^T) [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T = \\ &= \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{t_1 r_1}^T \\ \mathbf{H}_{t_2 r_1}^T \end{bmatrix} [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T. \end{aligned} \quad (\text{A.40})$$

Pré-multiplicando (A.40) por \mathbf{X}_2^* obtemos

$$\mathbf{X}_2^* [\mathbf{y}_{t_1}(5)]_{(2)} = \mathbf{H}_{t_2 r_1}^T [\mathcal{B}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T)^T. \quad (\text{A.41})$$

Tomando o *unfolding* modo-2 de (A.18), temos

$$\begin{aligned} [\mathbf{y}_{t_1}(2)]_{(2)} &= (\mathbf{X}^T \mathbf{H}^T \Psi_{R_1}^T) [\mathcal{W}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T = \\ &= \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{t_1 r_1}^T \\ \mathbf{H}_{t_2 r_1}^T \end{bmatrix} [\mathcal{W}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T. \end{aligned} \quad (\text{A.42})$$

Pré-multiplicando (A.42) por \mathbf{X}_2^* obtemos

$$\mathbf{X}_2^* [\mathbf{y}_{t_1}(2)]_{(2)} = \mathbf{H}_{t_2 r_1}^T [\mathcal{W}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T. \quad (\text{A.43})$$

A partir de (A.41) e (A.43) uma estimativa de $\mathbf{H}_{t_2 r_1}$ pode ser obtida resolvendo o problema

$$\arg \min_{\mathbf{H}_{t_2 r_1}} \left\| \begin{bmatrix} (\mathbf{X}_2^* [\mathbf{y}_{t_1}(5)]_{(2)})^T \\ (\mathbf{X}_2^* [\mathbf{y}_{t_1}(2)]_{(2)})^T \end{bmatrix} - \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T) [\mathcal{B}]_{(2)}^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T) [\mathcal{W}]_{(2)}^T \end{bmatrix} \mathbf{H}_{t_2 r_1} \right\|_F^2,$$

cuja solução é dada por:

$$\hat{\mathbf{H}}_{t_2 r_1} = \begin{bmatrix} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_2}^T) [\mathcal{B}]_{(2)}^T \\ (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T) [\mathcal{W}]_{(2)}^T \end{bmatrix}^\dagger \begin{bmatrix} (\mathbf{X}_2^* [\mathbf{y}_{t_1}(5)]_{(2)})^T \\ (\mathbf{X}_2^* [\mathbf{y}_{t_1}(2)]_{(2)})^T \end{bmatrix}. \quad (\text{A.44})$$

► Estimação de $\mathbf{H}_{t_2 r_2}$

Tomando o *unfolding* modo-2 de (A.23), temos

$$\begin{aligned} [\mathbf{y}_{t_1}(4)]_{(2)} &= (\mathbf{X}^T \mathbf{H}^T \Phi_{R_2}^T) [\mathcal{A}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T = \\ &= \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{t_1 r_2}^T \\ \mathbf{H}_{t_2 r_2}^T \end{bmatrix} [\mathcal{A}]_{(2)} (\mathbf{I}_K \otimes \mathbf{H}_{t_1 r_1}^T)^T. \end{aligned} \quad (\text{A.45})$$

Pré-multiplicando (A.45) por X_2^* obtemos

$$X_2^*[\mathcal{Y}_{t_1}(4)]_{(2)} = H_{t_2r_2}^T[\mathcal{A}]_{(2)}(I_K \otimes H_{t_1r_1}^T)^T. \quad (\text{A.46})$$

Tomando o *unfolding* modo-2 de (A.20), temos

$$\begin{aligned} [\mathcal{Y}_{t_1}(3)]_{(2)} &= (X^T H^T \Phi_{R_2}^T)[\mathcal{T}]_{(2)}(I_K \otimes H_{t_1r_2}^T)^T = \\ &= \begin{bmatrix} X_1^T & X_2^T \end{bmatrix} \begin{bmatrix} H_{t_1r_2}^T \\ H_{t_2r_2}^T \end{bmatrix} [\mathcal{T}]_{(2)}(I_K \otimes H_{t_1r_2}^T)^T. \end{aligned} \quad (\text{A.47})$$

Pré-multiplicando (A.47) por X_2^* obtemos

$$X_2^*[\mathcal{Y}_{t_1}(3)]_{(2)} = H_{t_2r_2}^T[\mathcal{T}]_{(2)}(I_K \otimes H_{t_1r_2}^T)^T. \quad (\text{A.48})$$

A partir de (A.46) e (A.48) podemos estimar $H_{t_2r_2}$ resolvendo o problema

$$\arg \min_{H_{t_2r_2}} \left\| \begin{bmatrix} (X_2^*[\mathcal{Y}_{t_1}(4)]_{(2)})^T \\ (X_2^*[\mathcal{Y}_{t_1}(3)]_{(2)})^T \end{bmatrix} - \begin{bmatrix} (I_K \otimes H_{t_1r_1}^T)[\mathcal{A}]_{(2)}^T \\ (I_K \otimes H_{t_1r_2}^T)[\mathcal{T}]_{(2)}^T \end{bmatrix} H_{t_2r_2} \right\|_F^2,$$

cuja solução é dada por:

$$\hat{H}_{t_2r_2} = \begin{bmatrix} (I_K \otimes H_{t_1r_1}^T)[\mathcal{A}]_{(2)}^T \\ (I_K \otimes H_{t_1r_2}^T)[\mathcal{T}]_{(2)}^T \end{bmatrix}^\dagger \begin{bmatrix} (X_2^*[\mathcal{Y}_{t_1}(4)]_{(2)})^T \\ (X_2^*[\mathcal{Y}_{t_1}(3)]_{(2)})^T \end{bmatrix}. \quad (\text{A.49})$$

► Estimação de $H_{r_1r_2}$

Para estimar $H_{r_1r_2}$ partimos das equações (A.11) (estamos desprezando as componentes de ruído)

$$Y_{t_1,k}(4) = H_{t_1r_1}^T W^{(k)} H_{r_1r_2}^T T^{(k)} H_{t_2r_2} X_2 + H_{t_1r_1}^T W^{(k)} H_{r_1r_2}^T T^{(k)} H_{t_1r_2} X_1. \quad (\text{A.50})$$

Pós-multiplicando (A.50) por X_2^H obtemos

$$\begin{aligned} P^{(k)} &= H_{t_1r_1}^T W^{(k)} H_{r_1r_2}^T T^{(k)} H_{t_2r_2} = \\ &= H_{t_1r_1}^T W_1 D_k(W_3) W_2^T H_{r_1r_2}^T T_1 D_k(T_3) T_2^T H_{t_2r_2}, \end{aligned} \quad (\text{A.51})$$

em que $P^{(k)} = Y_{t_1,k}(4) X_2^H$ e $W^{(k)}$ e $T^{(k)}$ foram expressos em termos de suas matrizes fatores. Aplicando o operador $\text{vec}(\cdot)$ nas expressões acima e aplicando a propriedade: $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$, obtemos

Tabela A.1: Algoritmo Comb-ALS-Two-Way

Inicialização: Atribua $i = 0$; Inicialize aleatoriamente as matrizes $\{\hat{\mathbf{H}}_{t_1 r_1}^{(i=0)}, \{\hat{\mathbf{H}}_{t_2 r_1}^{(i=0)}\}$ e $\{\hat{\mathbf{H}}_{t_1 r_2}^{(i=0)}, \{\hat{\mathbf{H}}_{t_2 r_2}^{(i=0)}\}$

- 1) Atribua $i = i + 1$.
- 2) Estime $\hat{\mathbf{H}}_{r_1 r_2}^{(i)}$ usando a equação (A.55)
- 3) Estime $\hat{\mathbf{H}}_{t_1 r_2}^{(i)}$ e $\hat{\mathbf{H}}_{t_1 r_1}^{(i)}$ usando as equações (A.34) e (A.39);
- 4) Estime $\hat{\mathbf{H}}_{t_2 r_1}^{(i)}$ e $\hat{\mathbf{H}}_{t_2 r_2}^{(i)}$ usando as equações (A.44) e (A.49);
- 5) Repita os passos 1 a 4 até a convergência.

$$\text{vec}(\mathbf{P}^{(k)}) = \mathbf{p}_k = ((\mathbf{H}_{t_2 r_2}^T \mathbf{T}_2) \otimes (\mathbf{H}_{t_1 r_1}^T \mathbf{W}_1)) \text{diag}(\text{vec}(\mathbf{S})) (\mathbf{T}_{3(k,:)}^T \otimes \mathbf{W}_{3(k,:)}^T), \quad (\text{A.52})$$

em que $\mathbf{S} \doteq \mathbf{W}_2^T \mathbf{H}_{r_1 r_2}^T \mathbf{T}_1$. Definindo

$$\begin{aligned} \mathbf{P} &= [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K] = \\ &= ((\mathbf{H}_{t_2 r_2}^T \mathbf{T}_2) \otimes (\mathbf{H}_{t_1 r_1}^T \mathbf{W}_1)) \text{diag}(\text{vec}(\mathbf{S})) (\mathbf{T}_3^T \otimes \mathbf{W}_3^T), \end{aligned} \quad (\text{A.53})$$

aplicando o operador $\text{vec}(\cdot)$ na expressão acima e mais uma vez aplicando a propriedade $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$ obtemos

$$\text{vec}(\mathbf{P}) = \left((\mathbf{T}_3^T \diamond \mathbf{W}_3^T)^T \diamond ((\mathbf{H}_{t_2 r_2}^T \mathbf{T}_2) \otimes (\mathbf{H}_{t_1 r_1}^T \mathbf{W}_1)) \right) (\mathbf{T}_1^T \otimes \mathbf{W}_2^T) \mathbf{h}_{r_1 r_2}, \quad (\text{A.54})$$

em que $\mathbf{h}_{r_1 r_2} \doteq \text{vec}(\mathbf{H}_{r_1 r_2}^T)$. A equação para a estimação de $\mathbf{H}_{r_1 r_2}$ vem do seguinte problema

$$\arg \min_{\mathbf{H}_{r_1 r_2}} \left\| \text{vec}(\mathbf{P}) - \left((\mathbf{T}_3^T \diamond \mathbf{W}_3^T)^T \diamond ((\mathbf{H}_{t_2 r_2}^T \mathbf{T}_2) \otimes (\mathbf{H}_{t_1 r_1}^T \mathbf{W}_1)) \right) (\mathbf{T}_1^T \otimes \mathbf{W}_2^T) \mathbf{h}_{r_1 r_2} \right\|^2,$$

cujas solução é dada por:

$$\hat{\mathbf{h}}_{r_1 r_2} = \left[\left((\mathbf{T}_3^T \diamond \mathbf{W}_3^T)^T \diamond ((\mathbf{H}_{t_2 r_2}^T \mathbf{T}_2) \otimes (\mathbf{H}_{t_1 r_1}^T \mathbf{W}_1)) \right) (\mathbf{T}_1^T \otimes \mathbf{W}_2^T) \right]^\dagger \left[\text{vec}(\mathbf{P}) \right]. \quad (\text{A.55})$$

A tabela A.1 resume o algoritmo proposto, aqui referido como Combined ALS Two-Way (Comb-ALS-Two-Way).

Referências Bibliográficas

- [1] K. Liu, A. Sadek, W. Su, and A. Kwasinski, *Cooperative communications and networking*. Cambridge University Press, 2008.
- [2] A. Sendonaris, E. Erkip, and B. Aazhang, “User Cooperation Diversity–Part I: System Description,” *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [3] —, “User Cooperation Diversity–Part II: Implementation Aspects and Performance Analysis,” *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.
- [4] M. Dohler and Y. Li, *Cooperative communications: hardware, channel and PHY*. John Wiley & Sons, 2010.
- [5] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, , and G. Fettweis, “Relay-based deployment concepts for wireless and mobile broadband radio,” *IEEE Communications Magazine*, vol. 42, pp. 80–89, Sept. 2004.
- [6] L. Cao, J. Zhang, and N. Kanno, “Multi-user cooperative communications with relay-coding for uplink IMT-advanced 4G systems,” *Proc. IEEE GLOBECOM’09*, pp. 1–6, Nov. 2009.
- [7] Y. Rong, X. Tang, and Y. Hua, “A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems,” *IEEE Transactions on Signal Processing*, vol. 57, pp. 4837–4851, Dec. 2009.
- [8] T. Kong and Y. Hua, “Optimal channel estimation and training design for MIMO relays,” in *IEEE Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, CA, USA, Nov. 2010, pp. 663–667.

- [9] P. Lioliou and M. Viberg, “Least-squares based channel estimation for MIMO relays,” in *2008 International ITG Workshop on Smart Antennas (WSA)*, 2008, pp. 90–95.
- [10] P. Lioliou, M. Viberg, and M. Coldrey, “Efficient channel estimation techniques for amplify and forward relaying systems,” *IEEE Transactions on Wireless Communications*, vol. 60, no. 11, pp. 3150–3155, Oct. 2012.
- [11] F. Roemer and M. Haardt, “Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse,” *IEEE Transactions on Signal Processing*, vol. 58, pp. 5720–5735, Nov. 2010.
- [12] C. A. R. Fernandes, A. L. F. de Almeida, and D. B. Costa, “Unified tensor modeling for blind receivers in multiuser uplink cooperative systems,” *IEEE Signal Processing Letters*, vol. 19, pp. 247–250, May. 2012.
- [13] A. L. F. de Almeida, C. A. R. Fernandes, and D. B. da Costa, “Multiuser Detection for Uplink DS-CDMA Amplify-and-Forward Relaying Systems,” *IEEE Signal Processing Letters*, vol. 20, no. 7, pp. 697–700, 2013.
- [14] R. A. Harshman, “Foundations of the PARAFAC procedure: Model and conditions for an ‘explanatory’ multi-mode factor analysis,” *UCLA Working Papers Phonetics*, vol. 16, pp. 1–84, 1970.
- [15] L. R. Tucker, “Some mathematical notes on three-mode factor analysis,” *Psychometrika*, vol. 31, pp. 279–311, 1966.
- [16] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, Nov. 2009.
- [17] A. L. F. de Almeida, “Tensor Modeling and Signal Processing for Wireless Communication Systems,” Ph.D. dissertation, Université de Nice-Sophia Antipolis - École Doctorale STIC, 2007.
- [18] R. Bro, “Multi-way analysis in the food industry: models, algorithms and applications,” Ph.D. dissertation, Royal Veterinary and Agricultural University, 1998.
- [19] A. Cichocki, D. Mandic, A.-H. Phan, C. Caiafa, G. Zhou, Q. Zhao, and L. D. Lathauwer, “Tensor decompositions for signal processing applications from two-way to multiway component analysis,” *IEEE Signal Processing Magazine*, 2014.

- [20] F. L. Hitchcock, "Multiple invariants and generalized rank of a p-way matrix or tensor," *Journal of Mathematics and Physics*, vol. 7, pp. 39–79, 1927.
- [21] L. R. Tucker, "The extension of factor analysis to three-dimensional matrices," *Contributions to Mathematical Psychology*, H. Gulliksen and N. Frederiksen, Eds. New York: Holt, Rinehart and Winston, pp. 110–127, 1964.
- [22] J. 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, 1970.
- [23] J. Kruskal, "Three way arrays: Rank and uniqueness of trilinear decomposition with applications to arithmetic complexity and statistics," *Lin. Algebra Applicat.*, vol. 18, pp. 95–138, 1977.
- [24] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, pp. 164–189, 1927.
- [25] R. B. Cattell, "Parallel proportional profiles and other principles for determining the choice of factors by rotation," *Psychometrika*, vol. 9, pp. 267–283, 1952.
- [26] —, "The three basic factor-analytic research design – their interrelations and derivatives," *Psych. Bull.*, vol. 49, pp. 499–452, 1952.
- [27] L. R. Tucker, "Implications of factor analysis of three-way matrices for measurement of change," *Problems in Measuring Change*, C. W. Harris, ed., University of Wisconsin Press, 1963.
- [28] J. D. Carroll, S. Pruzansky, and J. B. Kruskal, "CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters," *Psychometrika*, vol. 45, pp. 3–24, 1980.
- [29] R. A. Harshman, "Models for analysis of asymmetrical relationships among n objects or stimuli," *First Joint Meeting of the Psychometric Society and the Society for Mathematical Psychology*, McMaster University, Hamilton, Ontario, 1978.
- [30] R. A. Harshman and M. E. Lundy, "Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP," *Psychometrika*, vol. 61, pp. 133–154, 1996.

- [31] L. R. Ximenes, G. Favier, A. L. F. de Almeida, and Y. C. B. Silva, "PARAFAC-PARATUCK Semi-Blind Receivers for Two-Hop Cooperative MIMO Relay Systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3604–3615, Jul. 2014.
- [32] F. R. V. Guimarães, "Técnicas de seleção de nós em redes cooperativas cognitivas com compartilhamento espectral," Master's thesis, Universidade Federal do Ceará, 2013.
- [33] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay system via parallel factor analysis," *IEEE Transactions Wireless Communications*, vol. 11, pp. 2224–2233, Jun. 2012.
- [34] Y. Rong, "Optimal Joint Source and Relay Beamforming for MIMO Relays with Direct Link," *IEEE Communications Letters*, vol. 14, no. 5, pp. 390–392, May 2010.