

ANALYSIS AND DESIGN OF PILOT-AIDED  
MULTICARRIER SYSTEMS OVER DOUBLY SELECTIVE  
CHANNELS WITH A LOCAL SUBCARRIER  
PROCESSING CONSTRAINT

DISSERTATION

Presented in Partial Fulfillment of the Requirements for  
the Degree Doctor of Philosophy in the  
Graduate School of The Ohio State University

By

Sibasish Das, B. Tech., M.S.

\* \* \* \* \*

The Ohio State University

2008

Dissertation Committee:

Dr. Philip Schniter, Adviser

Dr. Hesham El Gamal

Dr. C. Emre Koksal

Approved by

---

Adviser

Department of Electrical  
and Computer Engineering

## ABSTRACT

In the near future, the need for portable multi-antenna wireless devices supporting high data-rates in a harsh mobile environment is anticipated. Typically, in this scenario, the underlying wireless channels are time- and frequency selective, hence doubly selective (DS). Multicarrier modulation (MCM) schemes, wherein data is split into sub-streams and transmitted on a set of subcarriers in parallel, is frequently used on wireless communication systems. DS channels generate both inter-symbol interference (ISI) as well as inter-carrier interference (ICI) for MCM transmissions, thereby making the design of MCM receivers a challenging task. The challenge is compounded by strict computational power constraints as a result of the demand for light, portable, low-power devices, and by the lack of channel state information (CSI) in practical wireless receivers. Local subcarrier processing (LSP) using CSI acquired via pilot-aided channel estimation is arguably the most popular MCM receiver design technique. The LSP constraint allows the use of only a small subset of local observations and *dominant* ICI coefficients to retrieve information transmitted on each subcarrier. The characterization of performance limits imposed by the LSP constraint on generic MCM schemes serves as the primary motivation for this dissertation.

We characterize a lower-bound on the ergodic LSP-constrained achievable-rate of channel-estimation based generic multi-antenna MCM receivers operating over DS

channels. In doing so, we assume the use of pilot-aided MCM transmission and independent and identically distributed (i.i.d.) Gaussian codebooks. We use this lower-bound to evaluate the high-SNR spectral efficiency in two distinct regimes of high signal-to-noise ratio (SNR). In the regime where practical wireless systems are likely to operate, we find that multiple receive antennas are both necessary and sufficient to achieve spectral efficiency when i.i.d. Gaussian codebooks are used. There is another regime of even higher SNRs (unlikely for practical wireless systems to be operating in) where residual ICI and ISI components result in an achievable-rate ceiling. We also characterize a trade-off between LSP-constrained performance and complexity.

This LSP-constrained achievable-rate metric is a versatile tool for analysis and design of MCM systems over DS channels. As an example, we use this metric to compare the performance of various MCM schemes under a reception complexity constraint. As yet another example, we design novel LSP-constrained achievable-rate maximizing beamforming and combining vectors for multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems operating on DS channels.

Channel estimation errors are a major hindrance to the performance of MCM reception over DS channels. We show that channel re-estimation coupled with successive-decoding can be applied to reduce the effect of channel estimation errors on achievable-rate performance over DS channels. Specifically, we show that the channel re-estimation based approach is also spectrally efficient and provides a performance boost over “one-shot” channel estimation approaches.

To *Maa*, *Baba* and *Jhinuk*...

## ACKNOWLEDGMENTS

I have spent five wonderful years here at The Ohio State University. I shall always remember these years as some of the most interesting of my life.

First and foremost, I would like to express my gratitude towards my advisor Professor Philip Schniter for every thing that he has done for me. I have learnt more from him than any other teacher.

I would like to thank *Baba, Maa* and *Jhinuk* for being constant sources of support and inspiration. Thanks also to Mistu Masi, Mesho, Arko and Soumi for making me feel closer to home.

I would like to thank our sponsors, Motorola, Inc and the National Science Foundation, for providing me with this opportunity. This material is based upon work supported by the National Science Foundation under Grant No. 0237037. I would like to thank the faculty here at The Ohio State University, and in particular, the faculty of the Department of Electrical and Computer Engineering for the academic training I received here. I would like to thank Jeri McMichael for her help throughout with all non-academic issues.

Thanks to Arul, Arun (Pachai), Praveen, Aditi, Rahul, Subhojit, Anand and Sugumar for all the fun we have had together. Thanks to all my room-mates for the exciting days I spent with them. Thanks to all the IPS Laboratory folks for making school seem so much fun, and to all the other friends for making my time away from

school seem even more fun. Thanks to Indrajit and Arun (Kumar) for being much more than just friends.

Finally, I would like to thank this university for all that it has given me to cheer about. I will be a Buckeye faithful forever! *Go Bucks!*

## VITA

Born	.....	May 9, 1980, Jamshedpur, India
B. Tech. in E & ECE	.....	1998 to 2002 I.I.T. Kharagpur, India
Graduate Student	.....	2002 to present, The Ohio State University

## Publications

- **IEEE Trans. Info. Theory:** S.Das and P. Schniter, “On Local Subcarrier Processing for Pilot-Aided SIMO Multicarrier Systems over Doubly Selective Channels”, *to be submitted*.
- **IEEE Trans. Sig. Proc.:** P. Schniter and S. Das, “Max-SINR ISI/ICI-Shaping Multi-Carrier Communication Over the Doubly Dispersive Channel,” *vol. 55, no. 12, pp. 5782-5795, Dec. 2007*.
- **Allerton 2007:** S.Das and P. Schniter, “Non-coherent Communication over the Doubly Selective Channel Via Successive Decoding and Channel Re-estimation”, *Proc. Allerton Conf. on Communication, Control, and Computing, (Monticello, IL), Oct. 2007*.
- **Asilomar 2006:** S.Das and P. Schniter, “Beamforming and Combining Strategies for MIMO-OFDM over Doubly Selective Channels”, *in Proc. Asilomar Conf. on Signals, Systems, and Computers, (Pacific Grove, CA), Nov. 2006..*
- **Asilomar 2006:** S.Das and P. Schniter, “Design of Multi-carrier Modulation Based on a Complexity Constrained Achievable-Rate Metric”, *in Proc. Asilomar Conf. on Signals, Systems, and Computers, (Pacific Grove, CA), Nov. 2006..*

- **Asilomar 2004:** S. Das and P. Schniter, “A New Pulse Shaped Frequency Division Multiplexing Technique for Doubly Dispersive Channels,” *in Proc. Asilomar Conf. on Signals, Systems, and Computers, (Pacific Grove, CA), Nov. 2004.*

## LIST OF FIGURES

Figure	Page	
3.1	Transmission epoch and frame structure. Pilot MCM symbols and data MCM symbols are denoted by $\mathbf{P}$ and $\mathbf{D}$ , respectively. . . . .	37
3.2	Estimation for the $k^{th}$ subcarrier under the LSP constraint only uses coefficients within the dashed rectangle, i.e., the matrix $\hat{\mathbf{H}}_k(i)$ . After the column vector $\hat{\mathbf{h}}_{k,k}(i)$ (dotted rectangle) is removed, the remaining columns within $\hat{\mathbf{H}}_k(i)$ form the dominant ICI matrix $\hat{\hat{\mathbf{H}}}_k(i)$ . . . . .	38
4.1	Quasi-banded channel matrix. . . . .	55
4.2	The modulation pulse $\{a_n\}$ and demodulation pulse $\{b_n\}$ used. . . . .	56
4.3	The effect of channel estimation error $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ on the observations for pilot spacing $N_p = 4$ , $N_p = 8$ and $N_p = 16$ . . . . .	57
4.4	The interference variance $\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)})/N$ , labeled <b>Eff. Noise</b> , and its components: the variance of the additive noise $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}(i)})/N$ , labeled <b>Noise</b> ; the residual ICI $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{rici}}(i)})/N$ , labeled <b>RICI</b> ; the ISI $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}}(i)})/N$ , labeled <b>ISI</b> ; and the effect of channel estimation error $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ , labeled <b>CEE</b> . . . . .	58
4.5	Performance of LSP and GSP on the surrogate MCM system for (a) $N_r = 1$ , (b) $N_r = 2$ and (c) $N_r = 3$ . . . . .	59
4.6	The rate-complexity trade-offs with $N_r = 2$ at two SNRs. . . . .	60
4.7	Lower-bounds on LSP-constrained achievable rate of CP-OFDM, jointly optimized max-SINR MCM, and Gaussian pulsed MCM scheme for a SIMO-MCM system with $N_r = 2$ receive antennas and $N = 128$ subcarriers for $D_h = D_h = 1$ . . . . .	61

5.1	Structure of MIMO subcarrier coupling matrix $\hat{\mathbf{H}}(i)$ for the surrogate MIMO-OFDM system. Rectangle (in dotted lines) indicates the channel coefficients used for LSP and BV design for the $k^{\text{th}}$ subcarrier. . .	72
5.2	Lower-bound $R^{\text{lb}}$ on LSP-constrained achievable rate versus SNR for (a) $N_t = 4$ , (b) $N_t = 6$ and (c) $N_t = 8$ transmit antennas. . . . .	73
5.3	Lower-bound $R^{\text{lb}}$ on LSP-constrained achievable rate on a system in TDD mode with predicted CSI at the transmitter for (a) $F_{\text{Dop}}T_s = 0.008$ , and (b) $F_{\text{Dop}}T_s = 0.016$ . . . . .	74
6.1	Power allocation parameter $\alpha_{p,*}$ at various SNRs for a $N_s = 128$ -substream spectrally efficient transmission over a CE-BEM doubly selective fading channel with $N_h = 8$ taps of ISSI and $D \in \{1, 2\}$ . . . .	88
B.1	Structure of the dominant ICI matrix $\hat{\mathbf{H}}_k(i)$ when $N_r = 1$ and $D_h \leq D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ . . . . .	103
B.2	Structure of the dominant ICI matrix $\hat{\mathbf{H}}_k(i)$ when $N_r = 1$ and $D_r < D_h$ .	103

# TABLE OF CONTENTS

	<b>Page</b>
Abstract . . . . .	ii
Dedication . . . . .	iv
Acknowledgments . . . . .	v
Vita . . . . .	vii
List of Figures . . . . .	ix
Chapters:	
1. Introduction . . . . .	1
1.1 Wireless Channels . . . . .	1
1.1.1 Doubly Selective Wireless Channels . . . . .	2
1.2 Multicarrier Modulation . . . . .	3
1.3 Motivation . . . . .	4
1.4 Our Contribution . . . . .	6
1.4.1 LSP-Constrained Achievable-Rate . . . . .	6
1.4.2 Choosing MCM System Parameters . . . . .	7
1.4.3 Beamforming for MIMO-OFDM . . . . .	8
1.4.4 Successive Decoding with Channel Re-Estimation . . . . .	9
1.5 Organization . . . . .	11
1.6 Notation Used . . . . .	12
2. System Model . . . . .	13
2.1 Multicarrier Modulation . . . . .	13
2.2 Doubly-Selective Channel Models . . . . .	14

2.2.1	WSSUS Rayleigh-Faded Model . . . . .	14
2.2.2	CE-BEM Model . . . . .	15
2.3	Multicarrier Demodulation . . . . .	17
2.4	CP-OFDM: An Example MCM System . . . . .	20
2.5	Incorporating Multiple Antennas . . . . .	21
2.5.1	Single-Input Multiple-Output (SIMO) Systems . . . . .	21
2.5.2	Beamforming MIMO Systems . . . . .	23
3.	Transmission and Reception Strategy . . . . .	26
3.1	Code and Pilot Structure . . . . .	26
3.2	Pilot-Aided Reception . . . . .	28
3.3	Channel Estimation . . . . .	28
3.4	Local Subcarrier Processing . . . . .	33
4.	LSP-Constrained Performance Analysis . . . . .	39
4.1	Analysis Via Surrogate System . . . . .	39
4.2	SNR Regimes . . . . .	40
4.3	Lower Bound on LSP-Constrained Achievable-Rate . . . . .	42
4.4	Spectral Efficiency . . . . .	45
4.5	Numerical Examples . . . . .	49
4.6	Rate-Complexity Trade-Off . . . . .	52
4.7	Comparing MCM Schemes under A Complexity Constraint . . . . .	53
5.	Beamforming for MIMO-OFDM . . . . .	62
5.1	LSP on Beamforming MIMO-OFDM System . . . . .	62
5.2	Combiner and Beamforming Vector Designs . . . . .	65
5.2.1	Combiner Design . . . . .	65
5.2.2	Max-SNR Beamforming for DS Channels . . . . .	66
5.2.3	Approximate Max-AR Beamforming and Combining . . . . .	67
5.3	Numerical Results and Discussion . . . . .	69
6.	Incorporating Channel Re-Estimation . . . . .	75
6.1	System Model . . . . .	75
6.1.1	Transmission Model . . . . .	75
6.1.2	Channel Model . . . . .	77
6.2	Non-Coherent Pilot-and-Data-Aided Communication . . . . .	79
6.2.1	Description of Scheme . . . . .	79
6.2.2	Achievable-Rate Analysis . . . . .	81
6.2.3	Asymptotic Achievable-Rate Analysis . . . . .	82

6.3	Illustrative Example . . . . .	84
7.	Conclusions . . . . .	89
7.1	Summary of Our Work . . . . .	89
7.2	Future Research Directions . . . . .	91
Appendices:		
A.	Derivation of Interference Statistics . . . . .	93
B.	LSP-Constrained Achievable Rate Analysis . . . . .	96
B.1	Surrogate System . . . . .	96
B.2	Proof of Lemma 1 . . . . .	97
B.3	Proof for Theorem 1 . . . . .	99
B.4	Proof for Theorem 2 . . . . .	100
B.5	Proof for Theorem 3 . . . . .	102
B.6	Proof of Theorem 4 . . . . .	105
B.7	Proof for Theorem 5 . . . . .	107
C.	Analysis for Channel Re-estimating Reception . . . . .	109
C.1	Proof for Lemma 2 . . . . .	109
C.2	Proof for Theorem 1 . . . . .	111
C.3	Proof for Proposition 1 . . . . .	114
C.4	Proof for Proposition 2 . . . . .	115
	Bibliography . . . . .	117

# CHAPTER 1

## INTRODUCTION

### 1.1 Wireless Channels

Wireless links are a pivotal component of existing and future communication infrastructure. Communicating over wireless links poses two main challenges.

On one hand, reflections off physical objects create multiple paths between transmitter and receiver. Consequently, the receiver observes the sum of several delayed and phase-shifted echoes of the transmitted signal. This phenomenon is commonly called *multipath fading* [1]. When the delay spread of the paths is comparable to the transmitted symbol duration, energy from one symbol spills over into observations for neighboring symbols. This effect is called inter-symbol interference (ISI). Large delay-spreads imply that the channel's frequency response changes over the bandwidth of the transmitted signal. Thus, ISI causing channels are often called “frequency selective” channels.

On the other hand, the transmitter, the receiver and the reflectors may individually be in motion in a wireless environment. In a mobile setting, characteristics (i.e., delays and phase-shifts) of each echo observed at the receiver may change appreciably over a few symbol durations. In this scenario, the channel is called “time selective”.

If a pure tone is transmitted over a time selective channel, the receiver observes a band of frequencies. Furthermore, if symbols are transmitted on closely spaced carrier tones over a time selective channel, symbols on one tone contaminate observations for neighboring tones at the receiver. This effect is called inter-carrier interference (ICI) [2–5].

### 1.1.1 Doubly Selective Wireless Channels

A wireless channel that is both time- and frequency selective is called doubly selective (DS). Next generation systems, expected to support high data-rate services in mobile environments, have to be designed for the DS channel. The principal challenge in communicating over the DS channel is combating/exploiting ICI and ISI simultaneously. In this dissertation, we focus on communication over DS wireless channels.

Popularly, DS channels are modeled as noisy linear time-varying systems. In such a noisy linear time-varying system, the received sequence can be written as a convolution between the transmitted sequence and the time-varying channel impulse response that has been corrupted by additive Gaussian noise. A number of DS channel models have been proposed. For instance, Bello proposed the wide-sense stationary uncorrelated scattering (WSSUS) channel model in [6]. An alternative approach is to represent the DS wireless channel in terms of a truncated Fourier basis as seen in [7–9]. The resulting model is popularly called the complex-exponential basis expansion model (CE-BEM) for a DS channel.

The maximum Doppler frequency measures the rate of time-variations of a DS channel, where a higher maximum Doppler frequency implies rapid time-variations.

Mobility being the cause of time-variations in a wireless channel, the maximum Doppler frequency is a function of the velocities of the communicating devices and reflectors. The maximum Doppler frequency is also a function of the transmission carrier frequency.

The use of multicarrier modulation (MCM) for communication over DS channels is popular, and serves as the primary focus of this dissertation.

## 1.2 Multicarrier Modulation

In MCM, the information stream is split into several substreams and modulated on a set of subcarriers in parallel. These modulated subcarriers are transmitted after being shaped by a possibly non-rectangular modulation pulse. At the receiver, the observations are demodulated using another possibly non-rectangular demodulation pulse.

Cyclic-prefixed orthogonal frequency division multiplexing (CP-OFDM), [6,7], is perhaps the most popular MCM scheme. On purely frequency-selective channels, CP-OFDM creates parallel non-interfering channels – one on each subcarrier, thereby facilitating simple reception strategies. Unfortunately, on the DS channel, CP-OFDMs subcarriers interfere. Moreover, the rectangular CP-OFDM pulses imply that the side-lobes of the subcarrier spectra decay slowly with increasing frequency distance. The resulting ICI pattern necessitates the use of expensive equalization techniques at the receiver.

Among the most popular MCM schemes for communication over the DS channel are those based on time-frequency shifts of prototype pulses, e.g., [10–15]. Matheus and Kammeyer [16,17] proposed the use of dilated Gaussian prototype pulses with the

aim of suppressing ICI and/or ISI. Kozek and Molisch [12] designed non-rectangular pulses for robustness against ISI/ICI using knowledge of the channel cross-ambiguity function. Strohmer and Beaver [15] used a similar approach, but used pulses generated by a Gabor basis instead. Vahlin and Holte [18] designed pulses for an offset quadrature amplitude modulation (OQAM) OFDM system to reduce interference to adjacent frequency bands. Slimane [19] designed non-rectangular pulsed OFDM to reduce peak-to-average power ratio. Schniter [20] proposed a receiver for CP-OFDM using ISI-suppressing demodulation pulses. Hunziker and Dahlhaus [21] proposed a scheme using a Gaussian modulation pulse that minimized an out-of-target ISI/ICI metric for a purely frequency-selective and purely time-selective channels. In our previous work [22,23], we designed MCM pulses for the WSSUS DS channel to shape ISI/ICI into a desired target pattern.

### 1.3 Motivation

In DS channels, implementation of the optimal MCM demodulator is made difficult by the complicated structure of the ISI/ICI response. For example, the number of ISI/ICI coefficients grows with the square of the number of subcarriers, which itself may be quite large (e.g.,  $> 100$ ) in practical systems. The number of ISI/ICI coefficients increases further when multiple antennas are used. For this reason, practical receiver designs have been proposed that use only the *dominant* ISI/ICI coefficients. Examples include the iterative likelihood maximization in [21], the two-level block DFE structure proposed in [24], efficient sequence detection algorithms in [25], iterative minimum mean squared error (MMSE) and iterative maximum likelihood (ML) equalizers in turbo configuration [22,23], as well as the receiver designs

in [16, 20, 26–30]. In the typical case of narrowband subcarriers and band-limited Doppler spreading, the dominant ISI/ICI coefficients are those that correspond to interference from *neighboring* subcarriers; both ISI and ICI from far-away subcarriers is much weaker. This claim is justified, e.g., by the ISI/ICI analysis of CP-OFDM systems in [2, 4, 28, 31–33] and Weyl-Heisenberg systems (e.g., those using time/frequency shifts of prototype pulses) in [5, 15, 21]. For these reasons, we are interested in characterizing the performance of MCM receivers that consider interference from only neighboring subcarriers, i.e., that operate under a *local subcarrier processing* (LSP) constraint. While previous studies focused on particular schemes and particular LSP-constrained MCM systems (e.g., [8, 21, 24, 28, 34]), we are interested in *general* performance limits imposed by the LSP constraint. Hence, we undertake an information-theoretic approach to the study of LSP constrained MCM reception.

Most practical wireless systems are non-coherent, i.e., channel state information (CSI) is not available a priori at the receiver. A receiver operating over a DS channel has to update CSI regularly due to time-variations in the channel. These CSI updates can require significant computational burden and form a critical component of receiver design. A popular non-coherent receiver design strategy is to use channel estimation for pilot-aided transmission (PAT) [35]. In PAT, the transmitted sequence consists of “pilot” symbols that are known a priori at the receiver. These pilots are used to acquire CSI at the receiver via channel estimation algorithms. PAT schemes are popular with the DS channel, (e.g., see [23, 36–39]). Information-theoretic analysis of PAT schemes can be found in [36, 40–45] and specifically for the (CE-BEM) DS channel in [46–49]. Realizing the importance of acquiring CSI in a practical wireless

receiver for DS channels, we consider MCM PAT schemes with channel-estimation based reception for our investigations.

We characterize the LSP-constrained achievable-rate of multi-antenna MCM reception for PAT over non-coherent WSSUS Rayleigh-fading DS channels. In our analysis, we assume the use of i.i.d. Gaussian codebooks only. This LSP-constrained achievable rate is a versatile tool in the design and analysis of practical MCM systems over DS channels. We demonstrate this versatility by presenting a few applications of the LSP-constrained achievable-rate metric in Section 1.4.

## 1.4 Our Contribution

### 1.4.1 LSP-Constrained Achievable-Rate

We find that the statistical properties of interference observed when decoding under the LSP-constraint makes achievable-rate analysis difficult. As an alternative, a surrogate system model that approximates the MCM system is derived. A lower-bound the LSP-constrained achievable-rate of MCM reception using this surrogate system is then derived [50]. We analyze the derived lower-bound at high signal-to-noise-ratio (SNR), since this is the practical operating regime for most MCM systems. It is convenient to split our analysis of the LSP-constrained achievable-rate in two regimes, defined by the relative levels of residual ICI, ISI, channel estimation error and additive noise. We define *Regime 1* as the range of SNRs for which the levels of additive noise are much larger than those of residual ICI (RICI) and ISI, and *Regime 2* as the range of SNRs for which the levels of RICI and ISI are much larger than that of additive noise.

In Regime 1, it is seen that the use of multiple receive antennas is *necessary and sufficient* for “spectrally efficient” LSP-constrained MCM reception when i.i.d Gaussian codebooks are used. Here, “spectrally efficient” refers to the pre-log factor of the LSP-constrained achievable-rate expression being equal to the prelog-factor for the achievable-rate expression for global subcarrier processing (GSP) reception, a strategy employing all observations and all ICI coefficients. Note that the use of a single receive antenna is *insufficient* with LSP, whereas it is sufficient with optimal reception [51]. Furthermore, this lower-bound on the LSP-constrained achievable rate may be tightened by incorporating more observations and ICI coefficients in LSP at the expense of increased implementation complexity.

Finally, in Regime 2, we observe that the presence of non-vanishing ISI / RICI causes an achievable-rate ceiling for LSP. However, we demonstrate that judicious selection of MCM system parameters can ensure that Regime 2 lies beyond the operating SNR range of practical wireless systems.

### 1.4.2 Choosing MCM System Parameters

The choice of modulation and demodulation pulses is a critical component of MCM system design. For a given set of channel spreading characteristics, the number of dominant ICI (and ISI) coefficients is a function of the MCM pulse shape as well as the time-frequency spacing between the MCM pulses. The time-frequency spacing affects the system’s spectral efficiency and hence must be chosen with care. The pulse shape determines the time-frequency dispersion of the pulses, which affects the number of significant ISI/ICI coefficients. It also determines the correlation between these coefficients, which affects the diversity that can be exploited by the decoder. In

short, MCM schemes should be designed with decoding complexity, spectral efficiency, as well as diversity exploitation in mind. We are, however, not aware of any MCM design strategies that attempt to address all three of these aspects simultaneously. For example, the classical MCM schemes that minimize ISI/ICI subject to (bi)orthogonal pulse constraints (e.g., [10–12, 15]) admit good ISI/ICI suppression only at relatively low spectral efficiencies (as a consequence of the Balian-Low Theorem [52]) and never consider coefficient correlation. As another example, modern MCM schemes can attain much higher spectral efficiencies while maintaining only a small number of non-negligible ISI/ICI coefficients (e.g., [20, 22, 30, 53]) but do so without explicit concern for ISI/ICI coefficient correlation. As yet another example, so-called “maximum diversity” pulses [8, 54] have been proposed at the price of low spectral efficiency and high implementation complexity.

The lower-bound on LSP-constrained achievable-rate allows us to incorporate all three design metrics, i.e., implementation complexity, spectral efficiency, as well as coefficient correlation. In [55], we compare the lower-bounds on LSP-constrained achievable rates for a set of MCM schemes. By proper choice of LSP parameters, we ensure identical reception complexity during this comparison.

### 1.4.3 Beamforming for MIMO-OFDM

The last decade has seen the emergence of multiple-input multiple-output (MIMO) systems that provide large spatial diversity gains [56–58]. MIMO systems promise a diversity advantage proportional to the product of the number of transmit and receive antennas. Space-time codes [56, 57] and, more recently, lattice codes [59] have been shown to take advantage of spatial diversity. When channel state information (CSI)

is available at the transmitter, one can also combine traditional coding techniques with beamforming and antenna combining to leverage spatial diversity (see [58, 60] and references therein). Spatial diversity can also be harnessed in a MIMO-OFDM system via beamforming and combining (see, e.g., [61–63]).

One of the central assumptions in traditional MIMO-OFDM beamforming and combining is that *subcarriers do not interfere* and, hence, the SNR for each subcarrier can be maximized independently. However, DS channels cause subcarriers to interfere with each other, rendering traditional MIMO-OFDM beam-forming strategies ineffective. We design beamforming and combining strategies for MIMO-OFDM operating over DS channels aimed at maximizing the LSP-constrained achievable rate [64]. These novel MIMO-OFDM beamforming and combining strategies are effective even in the presence of significant ICI, and also show robustness to channel prediction errors.

#### 1.4.4 Successive Decoding with Channel Re-Estimation

We have used a “one-shot” PAT scheme in our studies, whereby the receiver computes a pilot-aided channel estimate and subsequently uses it for coherent data decoding. Channel estimation error acts as additional noise that degrades decoding performance and thus the rate of reliable communication [40–42]. Though channel estimation can be improved by allocating more transmission resources (e.g., rate and power) to pilots, doing so limits the resources that remain for data transmission.

As an improvement to one-shot estimation/decoding of PAT, several authors have considered iterative (i.e., “turbo”) estimation/decoding strategies, whereby soft decoder outputs are employed to refine channel estimates, which can then be used for

improved decoding, and so on [65–68]. Such systems are generally suboptimal and difficult to analyze.

More recently, the use of block interleaving with successive decoding has been proposed as a more structured approach to joint estimation/decoding of PAT [49,69]. There the idea is to split the information stream into independently coded substreams and decode them successively. While a pilot-aided channel estimate is used to decode the first substream, reliably decoded symbols can be employed to refine the channel estimates used by later decoding stages. For long coding blocks and properly chosen substream rates (e.g., assuming known channel statistics), each substream can be reliably decoded, greatly simplifying the design and analysis of such systems. PAT with successive decoding has been used successfully in time-selective and frequency-selective SISO channels [69] as well as time-selective MIMO channels [49].

In [70], we propose a scheme for non-coherent communication over doubly (i.e, time- and frequency-) selective fading channels that uses successive decoding and channel re-estimation at the receiver. Assuming perfect decoding of each stream, we calculate an achievable-rate lower-bound and use it to infer a set of substream rates which are sufficient to ensure perfect decoding. We also characterize the high-SNR spectral efficiency of the proposed communication strategy. we consider the case of transmissions over a CE-BEM DS channel in further detail. For this CE-BEM DS channel, we design a suitable pilot pattern and through it verify that the pre-log factor of the high-SNR achievable rate expression coincides with that of the ergodic capacity [47,71]. We also propose a pilot/data power allocation strategy that maximizes a lower-bound on the achievable rate.

## 1.5 Organization

In Chapter 2, we describe the generic MCM modulation and demodulation procedure. We present details of the WSSUS and CE-BEM DS channel models. We also describe the process of incorporating multiple receive and transmit antenna into the MCM system model.

In Chapter 3, we describe the construction of the transmission sequence from information bits through coding, interleaving and insertion of pilots. We present details of the one-shot channel estimation procedure using transmitted pilot MCM symbols. Finally, we define and present details of the LSP constraint on MCM reception. We develop a model to analyze MCM reception under the LSP constraint. We also investigate the statistics of the interference observed for LSP-constrained reception.

We derive and characterize the performance of reception in single-input multiple-output (SIMO) MCM systems under an LSP constraint in Chapter 4. In order to do so, we derive a surrogate system model that makes analysis tractable. We then define SNR regimes based on interference characteristics. Using the surrogate model, we derive a lower-bound on the LSP-constrained achievable-rate, and study the spectral efficiency of LSP-constrained MCM reception. We explore a trade-off between LSP constrained performance and LSP complexity, and compare various MCM schemes under a complexity constraint.

In Chapter 5, we consider the MIMO-OFDM system over DS channels. We design novel beamforming and combining strategies aimed at maximizing the LSP-constrained achievable rate.

In Chapter 6, we explore how channel re-estimation improves performance over DS channels. We also present a pilot-data power allocation strategy for communication

systems over CE-BEM DS channels that use channel re-estimation and successive decoding at the receiver.

Finally, conclusions are presented in Chapter 7.

## 1.6 Notation Used

In the dissertation, we use  $(\cdot)^T$  to denote transpose, and  $(\cdot)^H$  the conjugate transpose. We use  $[\mathbf{B}]_{m,n}$  to denote the element in the  $m^{\text{th}}$  row and  $n^{\text{th}}$  column of  $\mathbf{B}$ , where row/column indices begin with zero. Further, we use  $\mathbf{I}_K$  to denote the  $K \times K$  identity matrix,  $\mathbf{e}_K(k)$  to denote its  $k^{\text{th}}$  column.  $\mathcal{D}(a_0, a_1, \dots, a_K)$  and  $\mathcal{D}(\mathbf{a})$  denote the square matrix constructed by placing the elements of  $\mathbf{a} = [a_0, a_1, \dots, a_K]^T$  along the leading diagonal. Similarly,  $\mathcal{D}_b(\mathbf{A}_0, \dots, \mathbf{A}_K)$  creates a block-diagonal matrix from matrices  $\{\mathbf{A}_k\}_{k=0}^K$ , where  $\{\mathbf{A}_k\}_{k=0}^K$  have the same dimensions. On the other hand,  $\text{diag}_k(\mathbf{A})$  stacks the elements on the  $k$ -th diagonal of matrix  $\mathbf{A}$  into a column vector. The trace and Frobenius norm of a matrix and the norm of a vector are denoted by  $\text{tr}(\cdot)$ ,  $\|\cdot\|_F$  and  $\|\cdot\|$ , respectively. The Kronecker product and element-wise product of two matrices are denoted by operators  $\otimes$  and  $\odot$ , respectively. Also,  $\delta_l$  denotes the Kronecker delta with argument  $l$ ,  $\langle \cdot \rangle_N$  the modulo- $N$  operation, and  $\mathbb{C}$  the set of all complex numbers. Expectation is denoted by  $E(\cdot)$ , autocovariance by  $\mathbf{\Sigma}_{\mathbf{b}} := E(\mathbf{b}\mathbf{b}^H) - E(\mathbf{b})E(\mathbf{b}^H)$ , and crosscovariance by  $\mathbf{\Sigma}_{\mathbf{a},\mathbf{b}} := E(\mathbf{a}\mathbf{b}^H) - E(\mathbf{a})E(\mathbf{b}^H)$ . The mutual information (MI) between two random entities  $X$  and  $Y$  is denoted by  $I(X; Y)$ . When conditioned on a realization of the random entity  $Z$ , MI is denoted by  $I(X; Y | Z)$ . Finally, we abbreviate the phrase “without loss of generality” by “w.l.o.g.”, and the phrase “with probability one” by “w.p.1” for brevity.

## CHAPTER 2

### SYSTEM MODEL

We develop the mathematical model for communication using a generic MCM system. We first present details of the modulation strategy used in MCM transmitters. The transmitted MCM signal experiences a DS channel en route to the receiver. We describe two popular channel models used for such DS channels. The receiver has to demodulate a received sequence that is a noisy and faded version of the MCM transmit sequence. The details of this demodulation procedure are presented next. Finally, we describe extensions to this model to incorporate multiple receive antennas and a multi-antenna transmitter employing beamforming.

#### 2.1 Multicarrier Modulation

At each multicarrier symbol index  $i \in \mathbb{Z}$ ,  $N$  scalar symbols  $\{s_k(i)\}_{k=0}^{N-1}$  are collected to form a multicarrier symbol  $\mathbf{s}(i) = [s_0(i), \dots, s_{N-1}(i)]^T$ , where  $s_k(i)$  is modulated on the  $k$ -th subcarrier. These scalar symbols are obtained by coding and interleaving the information sequence and inserting pilot symbols for purposes of channel estimation. The details of coding and interleaving are presented in Section 3.1, whereas the details of the channel estimation procedure are presented in Section 3.3. The modulated subcarriers are then shaped by a (possibly non-rectangular) modulation pulse  $\{a_n\}$

and transmitted. The transmitted sequence  $\{t_n\}$  can then be expressed as:

$$t_n = \sum_{i=-\infty}^{\infty} a_{n-iN_s} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(i) e^{j\frac{2\pi}{N}(n-iN_s-N_o)k}. \quad (2.1)$$

In (2.1),  $N_s$  is the symbol interval, so that a new multicarrier is transmitted every  $N_s$  samples. Note that  $N_s > N$  corresponds to adding time domain guards of  $N_s - N$  samples at the expense of a loss in spectral efficiency,  $N_s = N$  corresponds to full Nyquist rate signaling and  $N_s < N$  corresponds to a overloaded system. Additionally in (2.1),  $N_o \in \{0, \dots, N - 1\}$  delays the subcarrier origin relative to the pulse origin.

## 2.2 Doubly-Selective Channel Models

The transmitted sequence encounters a doubly selective (DS) wireless channel en route to the receiver. This channel is described by its time-variant discrete impulse response  $h_{n,l}$ , defined as the time- $n$  response to an impulse applied at time  $n - l$ . We assume a causal impulse response that has a maximum length of  $N_h$  samples. The additive noise corrupted and faded transmit sequence observed at the receiver can then be expressed as

$$x_n = \sum_{l=0}^{N_h-1} h_{n,l} t_{n-l} + \nu_n, \quad (2.2)$$

where  $\{\nu_n\}$  is a circular white Gaussian noise (CWGN) sequence such that  $E\{\nu_{n_1}\nu_{n_2}\} = \sigma^2\delta_{n_1-n_2}$ .

Two DS channel models are used in this dissertation.

### 2.2.1 WSSUS Rayleigh-Faded Model

The wide-sense stationary uncorrelated scattering (WSSUS) Rayleigh-faded channel model is used in Chapters 3 through 5. In this model, each lag-tap of the channel is

modeled as complex Gaussian random sequence with wide-sense stationary statistics, i.e., stationary first- and second order statistics. The statistics of the DS channel change rather slowly, and wide-sense stationarity provides a good “local” statistical characterization of each lag-tap of the channel. Moreover, a rich scattering environment is considered that leads to mutually uncorrelated channel lag-taps. Then the WSSUS channels can be statistically characterized by

$$\mathbb{E}\{h_{n_1,l_1}h_{n_2,l_2}^*\} = \sigma_{l_1}^2\phi_{n_1-n_2}\delta_{l_1-l_2}. \quad (2.3)$$

In (2.3),  $\sigma_l^2 := \mathbb{E}\{|h_{n,l}|^2\}$  denotes the variance (or, power) of the  $l$ -th lag-tap. Furthermore,  $\{\phi_n\}$  is the normalized auto-correlation sequence, i.e.,  $\phi_0 = 1$ , and is dependent on the propagation model used. For instance, when isotropic scattering [1] with a single-sided maximum Doppler spread of  $F_{\text{Dop}}$  is assumed,  $\phi_n = J_0(2\pi F_{\text{Dop}}T_s n)$ , where  $J_0$  denotes the zero-th order Bessel function and  $T_s$  the sampling interval.

### 2.2.2 CE-BEM Model

We also consider DS channels which obey a complex-exponential basis expansion model (CE-BEM) in Chapter 6. In this model, each lag is represented in terms of complex-exponential basis via a truncated Fourier series expansion. The Fourier series coefficients, so obtained, are modeled as uncorrelated complex Gaussian random variables. These coefficients are commonly called the CE-BEM coefficients. Again, a rich scattering environment is considered, which implies that the CE-BEM coefficients are uncorrelated across lag-taps. The DS channel is then characterized by the mutually uncorrelated complex-Gaussian CE-BEM coefficients  $\{\xi_{m,l} : m \in \{-D, \dots, D\}, l \in \{0, \dots, N_h - 1\}\}$ , i.e,

$$\mathbb{E}[\xi_{m_1,l_1}\xi_{m_2,l_2}^*] = \sigma_{m_1,l_1}^2\delta_{m_1-m_2}\delta_{l_1-l_2}. \quad (2.4)$$

Here,  $\sigma_{m,l}^2$  is the variance of the  $m$ -th CE-BEM coefficient for the  $l$ -th lag-tap of the DS channel. The impulse response of the DS channel can then be defined in terms of the CE-BEM coefficients by

$$h_{n,l} = \frac{1}{\sqrt{N}} \sum_{m=-D}^D \xi_{m,l} e^{j\frac{2\pi}{N}m(n-l)}. \quad (2.5)$$

It is clear from (2.5) that  $D$  measures the maximum single-sided frequency dispersion produced by DS channel and can be related to the single-sided maximum Doppler spread  $F_{\text{Dop}}$  via

$$D \approx \lceil F_{\text{Dop}} T_s N \rceil. \quad (2.6)$$

Recall that  $T_s$  represents the sampling interval for the communication system.

We assume the wireless DS channels to be energy preserving throughout the dissertation. For WSSUS modeled wireless DS channels, this implies that

$$\sum_{l=0}^{N_h-1} \sigma_l^2 = 1. \quad (2.7)$$

When we have a uniform power-delay profile, as is assumed for simplicity through most of the dissertation, (2.7) implies that

$$\sigma_l^2 = \frac{1}{N_h}, \quad l \in \{0, \dots, N_h - 1\}. \quad (2.8)$$

Similarly for energy preserving CE-BEM DS channels, we have

$$\sum_{l=0}^{N_h-1} \sum_{m=-D}^D \sigma_{l,m}^2 = 1. \quad (2.9)$$

Moreover, when we have a uniform Doppler-delay profile, (2.9) implies that

$$\sigma_{l,m}^2 = \frac{1}{N_h(2D+1)}, \quad (2.10)$$

for all  $l \in \{0, \dots, N_h - 1\}$  and  $m \in \{-D, \dots, D\}$ .

## 2.3 Multicarrier Demodulation

The received sequence is demodulated during each multicarrier symbol interval at the receiver. In order to understand the demodulation process, we define  $x_n(i) := x_{iN_s+n}$ ,  $\nu_n(i) := \nu_{iN_s+n}$ , and  $h_{n,l}(i) := h_{iN_s+n,l}$  as the received sequence, the CWGN sequence and the channel impulse response for multicarrier symbol index  $i$ , respectively. With these definitions, equations (2.1) and (2.2) imply

$$x_n(i) = \nu_n(i) + \sum_{l=0}^{N_h-1} h_{n,l}(i) t_{iN_s+n-l} \quad (2.11)$$

$$\begin{aligned} &= \nu_n(i) + \sum_{l=0}^{N_h-1} h_{n,l}(i) \\ &\quad \times \sum_{q=-\infty}^{\infty} a_{(i-q)N_s+n-l} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k^{(q)} e^{j\frac{2\pi}{N}(n-l+iN_s-qN_s-N_o)k} \end{aligned} \quad (2.12)$$

$$\begin{aligned} &= \nu_n(i) + \sum_{l=0}^{N_h-1} h_{n,l}(i) \\ &\quad \times \sum_{q=-\infty}^{\infty} a_{qN_s+n-l} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(i-q) e^{j\frac{2\pi}{N}(n-l+qN_s-N_o)k}. \end{aligned} \quad (2.13)$$

The receiver employs the demodulation pulse  $\{b_n\}$  to generate  $\{r_d(i)\}_{d=0}^{N-1}$  the observation at subcarrier index  $d$  during multicarrier symbol index  $i$ :

$$r_d(i) = \frac{1}{\sqrt{N}} \sum_{n=-\infty}^{\infty} x_n(i) b_n e^{-j\frac{2\pi}{N}d(n-N_o)}. \quad (2.14)$$

Plugging (2.13) into (2.14), we find

$$r_d(i) = \sum_{q=-\infty}^{\infty} \sum_{k=0}^{N-1} H_{d-k,k}(i, q) s_k(i-q) + w_d(i). \quad (2.15)$$

In (2.15),

$$w_d(i) = \frac{1}{\sqrt{N}} \sum_{n=-\infty}^{\infty} b_n \nu_n(i) e^{-j\frac{2\pi}{N}d(n-N_o)} \quad (2.16)$$

is colored frequency-domain noise, and

$$H_{d,k}(i, q) = \frac{1}{N} \sum_{n=-\infty}^{\infty} \sum_{l=0}^{N_h-1} h_{n,l}(i) b_n a_{qN_s+n-l} e^{-j\frac{2\pi}{N}d(n-N_o)} e^{-j\frac{2\pi}{N}k(l-qN_s)} \quad (2.17)$$

subcarrier coupling coefficients. Equation (2.15) indicates that the subcarrier coupling coefficient  $H_{d,k}(i, q)$  can be interpreted as the response, at multicarrier symbol index  $i$  at subcarrier index  $k + d$ , to an impulse applied at multicarrier symbol index  $i - q$  and subcarrier  $k$ . Observe that  $H_{d,k}(i, q)$  is a function the modulation and demodulation pulses  $\{a_n\}$  and  $\{b_n\}$ . Note that  $\{H_{d,k}(i, q)\}_{q \neq 0}$  models the interference *across multicarrier symbols* and is referred to as inter symbol interference (ISI) coefficients, whereas,  $\{H_{d,k}(i, 0)\}_{d \neq 0}$  models the interference *across subcarriers of an MCM symbol* and is referred to as inter-carrier interference (ICI) coefficients.

In practice, we use finite-duration causal pulses  $\{a_n\}$  and  $\{b_n\}$  of length  $N_a$  and  $N_b$ , respectively. This implies that only a finite number of terms in the sequence  $\{H_{d,k}(i, q)\}_{q \in \mathbb{Z}}$  will be non-zero. In general,  $H_{d,k}^{(m)}(i, q)$  may be non-zero only for  $q \in \{-L_{\text{pre}}, \dots, L_{\text{pst}}\}$ , where

$$\begin{aligned} -L_{\text{pst}}N_s + N_a - 1 + N_h - 1 &\geq 0 \\ \Rightarrow L_{\text{pst}} &\leq \frac{N_a + N_h - 2}{N_s} \\ \Rightarrow L_{\text{pst}} &= \left\lfloor \frac{N_a + N_h - 2}{N_s} \right\rfloor, \end{aligned} \quad (2.18)$$

and

$$\begin{aligned} L_{\text{pre}}N_s &\leq N_b - 1 \\ \Rightarrow L_{\text{pre}} &= \left\lfloor \frac{N_b - 1}{N_s} \right\rfloor. \end{aligned} \quad (2.19)$$

We can now define the observation vector for the  $i$ -th multicarrier symbol index  $\mathbf{r}(i) := [r_0(i), \dots, r_{N-1}(i)]^T$ , the frequency domain noise vector for the  $i$ -th multicarrier symbol index  $\mathbf{w}(i) := [w_0(i), \dots, w_{N-1}(i)]^T$ , and subcarrier coupling matrices for the  $i$ -th multicarrier symbol index  $[\mathbf{H}(i, q)]_{d,k} := H_{d-k,k}(i, q)$ . Using these definitions, (2.15) can be re-written in matrix-vector form as

$$\mathbf{r}(i) = \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \mathbf{s}(i - q) + \mathbf{w}(i) \quad (2.20)$$

It is straightforward to observe from (2.20) that  $\{\mathbf{H}(i, q)\}_{q \neq 0}$  represent ISI whereas off-diagonal elements of  $\mathbf{H}(i, 0)$  represent ICI. Note that, with

$$\boldsymbol{\nu}(i) := [\nu_0(i), \dots, \nu_{N_b-1}(i)]^T, \quad (2.21)$$

we have

$$\mathbf{w}(i) = \mathbf{F}_N \mathbf{J} \mathcal{D}(\mathbf{b}) \boldsymbol{\nu}(i) \quad (2.22)$$

$$\mathbf{J} := \begin{bmatrix} \mathbf{0}_{N-N_o \times N_o} & \mathbf{I}_N & \cdots & \mathbf{I}_N \\ \mathbf{I}_{N_o} & & & \end{bmatrix} \begin{bmatrix} \mathbf{I}_{\bar{N}_o} \\ \mathbf{0}_{N-\bar{N}_o \times \bar{N}_o} \end{bmatrix}, \quad (2.23)$$

where  $\mathbf{F}_N$  denotes the unitary  $N$ -DFT matrix,  $\bar{N}_o := \langle N_b - N_o \rangle_N$ , and the number of  $\mathbf{I}_N$  matrices in  $\mathbf{J}$  is  $\lfloor \frac{N_b - N_o}{N} \rfloor$ . Thus the covariance of the frequency domain noise vector is

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{w}(i)} &:= \mathbb{E}\{\mathbf{w}(i)\mathbf{w}(i)^H\} \\ &= \mathbf{F}_N \mathbf{J} \mathcal{D}(\mathbf{b}) \mathbb{E}\{\boldsymbol{\nu}(i)\boldsymbol{\nu}(i)^H\} \mathcal{D}(\mathbf{b}^*) \mathbf{J}^H \mathbf{F}_N^H \\ &= \sigma_w^2 \mathbf{F}_N \mathbf{J} \mathcal{D}(\mathbf{b})^2 \mathbf{J}^H \mathbf{F}_N^H. \end{aligned} \quad (2.24)$$

## 2.4 CP-OFDM: An Example MCM System

Our description of the MCM system is generic. As an example, we now show that the popular CP-OFDM system [72] follows from our MCM system model for an appropriate choice of system parameters.

Recall that the incorporation of an  $N_h$ -length cyclic-prefix implies that an MCM symbol is transmitted every  $N + N_h - 1$  samples in a CP-OFDM system, so that  $N_s = N + N_h - 1$ . The CP-OFDM modulator uses a rectangular modulation pulse with  $N_a = N_s$  so that

$$a_n = \begin{cases} 1 & 0 \leq n \leq N_s - 1, \\ 0 & \text{else.} \end{cases} \quad (2.25)$$

The cyclic prefix is incorporated by setting  $N_o = N_s - N$ . At the receiver, the cyclic prefix is removed prior to demodulation with a  $N_b = N$ -length rectangular demodulation pulse

$$b_n = \begin{cases} 1 & N_o \leq n \leq N_s - 1, \\ 0 & \text{else.} \end{cases} \quad (2.26)$$

Note that  $N_g = N_s - N$  is analogous to the CP-OFDM time domain guard interval. It is straightforward to show that the modulation pulse in (2.25) and demodulation pulse in (2.26) imply that  $L_{\text{pre}} = L_{\text{pst}} = 0$ , i.e., ISI is completely suppressed. Thus for the CP-OFDM case, (2.20) reduces to

$$\mathbf{r}(i) = \mathbf{H}(i, 0)\mathbf{s}(i) + \mathbf{w}(i). \quad (2.27)$$

It is generally assumed that CP-OFDM encounters white frequency domain noise. By plugging (2.26) in (2.24), we verify that indeed  $\boldsymbol{\Sigma}_{\mathbf{w}(i)} = \sigma^2 \mathbf{I}_N$  for CP-OFDM.

## 2.5 Incorporating Multiple Antennas

In this section, we show that multiple receive or transmit antennas can be incorporated into our MCM system model through simple modifications.

### 2.5.1 Single-Input Multiple-Output (SIMO) Systems

We first look at the scenario when transmission from a single transmit antenna is received through  $N_r > 1$  receive antennas. Such systems are popularly known as single-input multiple-output (SIMO) systems. The received sequence  $x_n^{(n_r)}(i)$  at the  $n_r$ -th receive antenna for multicarrier symbol index  $i$  can be defined in similar fashion to (2.2) as

$$x_n^{(n_r)}(i) := \sum_{l=0}^{N_h-1} h_{n,l}^{(n_r)}(i) t_{n+iN_s-l} + \nu_n^{(n_r)}(i). \quad (2.28)$$

In (2.28),  $h_{n,l}^{(n_r)}(i)$  denotes the response at the  $n_r$ -th receive antenna at time  $n+iN_s$  to an impulse transmitted at time  $n+iN_s-l$  and  $\{\nu_n^{(n_r)}(i)\}$  denotes samples of additive circular white Gaussian noise (CWGN) at the  $n_r$ -th receive antenna. The received sequence is demodulated using the demodulation pulse  $\{b_n\}$  at each receive antenna. After demodulation with  $\{b_n\}$  at the  $n_r$ -th receive antenna, the observation at the  $d$ -th subcarrier output  $r_d^{(n_r)}(i)$  for  $d \in \{0, \dots, N-1\}$  can be expressed in similar fashion to (2.20) as

$$r_d^{(n_r)}(i) = \sum_{q=-\infty}^{\infty} \sum_{k=0}^{N-1} H_{d-k,k}^{(n_r)}(i, q) s_k(i-q) + w_d^{(n_r)}(i). \quad (2.29)$$

In (2.29), the subcarrier coupling coefficient for the  $n_r$ -th receive antenna  $H_{d,k}^{(n_r)}(i, q)$  is the response at subcarrier index  $k+d$  at the  $n_r$ -th antenna to an impulse transmitted on subcarrier  $k$  and multicarrier symbol  $i-q$ , whereas,  $w_d^{(n_r)}(i)$  is the frequency

domain noise observed at subcarrier index  $d$  at the  $n_r$ -th receive antenna. These can be defined in identical fashion to (2.17) and (2.16), respectively, as:

$$H_{d,k}^{(n_r)}(i, q) = \frac{1}{N} \sum_{n=-\infty}^{\infty} \sum_{l=0}^{N_h-1} h_{n,l}^{(n_r)}(i) b_n a_{qN_s+n-l} e^{-j\frac{2\pi}{N}d(n-N_o)} e^{-j\frac{2\pi}{N}k(l-qN_s)}, \quad (2.30)$$

$$w_d^{(n_r)}(i) = \frac{1}{\sqrt{N}} \sum_{n=-\infty}^{\infty} b_n \nu_n^{(n_r)}(i) e^{-j\frac{2\pi}{N}d(n-N_o)}. \quad (2.31)$$

Using (2.30), we define the subcarrier coupling matrix for the  $n_r$ -th receive antenna element-wise as  $[\mathbf{H}^{(n_r)}(i, q)]_{m_1, m_2} = H_{m_1, m_1 - m_2}^{(n_r)}(i, q)$ , and the SIMO subcarrier coupling matrix as

$$\mathbf{H}(i, q) = \sum_{n_r=1}^{N_r} \mathbf{H}^{(n_r)}(i, q) \otimes \mathbf{e}_{N_r}(n_r). \quad (2.32)$$

Grouping the observations at each subcarrier output at the  $n_r$ -th receive antenna into  $\mathbf{r}^{(n_r)}(i) := [r_0^{(n_r)}(i), \dots, r_{N-1}^{(n_r)}(i)]^T$ , we can define the observation for the  $i$ -th multicarrier symbol as

$$\mathbf{r}(i) = \sum_{n_r=1}^{N_r} \mathbf{r}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}(n_r). \quad (2.33)$$

Note that in  $\mathbf{r}(i)$ , observations are grouped by subcarrier index, i.e., in  $\mathbf{r}(i)$ , the observations at the  $N_r$  receive antenna at the  $k = 0$ -th subcarrier appear first, followed by the  $N_r$  observations at the  $k = 1$ -st subcarrier, and so on. In similar fashion, we group the noise terms into  $\mathbf{w}^{(n_r)}(i) := [w_0^{(n_r)}(i), \dots, w_{N-1}^{(n_r)}(i)]^T$  and define the noise vector

$$\mathbf{w}(i) = \sum_{n_r=1}^{N_r} \mathbf{w}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}(n_r). \quad (2.34)$$

Using these definitions, the observation  $\mathbf{r}(i)$  can be expressed in terms of the transmitted MCM symbol vectors as

$$\mathbf{r}(i) = \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \mathbf{s}(i - q) + \mathbf{w}(i). \quad (2.35)$$

Finally, we describe the statistical assumptions for the SIMO systems. We assume that all spatial paths have independent fades, so that

$$\mathbb{E}[h_{n_1, l_1}^{(n_{r_1})} h_{n_2, l_2}^{(n_{r_2})*}] = \sigma_{l_1}^2 \phi_{(n_1-n_2)} \delta_{(l_1-l_2)} \delta_{(n_{r_1}-n_{r_2})}. \quad (2.36)$$

Furthermore, we assume that the noise at different receive antennas are also independent, implying that

$$\mathbb{E}[\mathbf{w}^{(n_{r_1})}(i) \mathbf{w}^{(n_{r_2})}(i)^H] = \Sigma_{\mathbf{w}^{(n_r)}(i)} \delta_{(n_{r_1}-n_{r_2})}, \quad (2.37)$$

where  $\Sigma_{\mathbf{w}^{(n_r)}(i)}$  is given by (2.24). We assume that noise observed at each receive antenna has identical statistics. Then using (2.34) and (2.37), it is straightforward to show that

$$\mathbb{E}[\mathbf{w}(i) \mathbf{w}(i)^H] = \Sigma_{\mathbf{w}^{(0)}(i)} \otimes \mathbf{I}_{N_r}. \quad (2.38)$$

This concludes our discussion on modeling SIMO systems.

## 2.5.2 Beamforming MIMO Systems

We now describe an MCM system with multiple antennas at both transmitter and receiver. Such a system is commonly referred to as a multiple-input multiple-output (MIMO) system. In this dissertation, we only deal with beamforming MIMO-MCM systems. At a beamforming multi-antenna MCM (MA-MCM) transmitter, the  $i$ -th MA-MCM symbol  $\underline{\mathbf{s}}^{\text{ma}}(i) \in \mathbb{C}^{NN_t}$  is defined as  $\underline{\mathbf{s}}^{\text{ma}}(i) := [\mathbf{s}_0^{\text{ma}}(i)^T, \mathbf{s}_1^{\text{ma}}(i)^T, \dots, \mathbf{s}_{N-1}^{\text{ma}}(i)^T]^T$ , where the entries of vector  $\mathbf{s}_k^{\text{ma}}(i) \in \mathbb{C}^{N_t}$  are modulated onto the  $k$ -th subcarrier at the  $N_t$  transmit antennas. The set of beamforming vectors (BVs),  $\{\mathbf{c}_k(i) \in \mathbb{C}^{N_t}\}_{k=0}^{N-1}$ , relate the  $i$ -th MCM symbol to the  $i$ -th MA-MCM symbol by

$$\underline{\mathbf{s}}^{\text{ma}}(i) = \mathcal{D}_b(\mathbf{c}_0(i), \mathbf{c}_1(i), \dots, \mathbf{c}_{N-1}(i)) \mathbf{s}(i). \quad (2.39)$$

Components of  $\underline{\mathbf{s}}^{\text{ma}}(i)$  are collected along the respective transmit antenna, modulated using the MCM modulation technique described in Section 2.1 and transmitted. We note here that to make transmit power constraints easier to handle, we use unit power BF vectors, so that  $\|\mathbf{c}_k(i)\| = 1$ , for all  $k \in \{0, \dots, N-1\}$  and for all  $i \in \mathbb{Z}$ .

The received sequence at the  $n_r$ -th receive antenna can be expressed in terms of the  $N_t$  transmitted sequences  $\{t_n^{(n_t)}\}_{n_t=1}^{N_t}$  in similar fashion as (2.28) by

$$x_n^{(n_r)}(i) := \sum_{n_t=1}^{N_t} \sum_{l=0}^{N_h-1} h_{n,l}^{(n_r, n_t)}(i) t_{n+iN_s-l}^{(n_t)} + \nu_n^{(n_r)}(i). \quad (2.40)$$

In (2.40),  $h_{n,l}^{(n_r, n_t)}(i)$  denotes the response at the  $n_r$ -th receive antenna at time  $n+iN_s$  to an impulse transmitted on the  $n_t$ -th transmit antenna at time  $n+iN_s-l$ . Demodulation is identical to the SIMO system in Section 2.5.1. Then the frequency domain observation  $r_d^{(n_r)}(i)$  for  $d \in \{0, \dots, N-1\}$  can be expressed in similar fashion to (2.20) as

$$r_d^{(n_r)}(i) = \sum_{n_t=1}^{N_t} \sum_{q=-\infty}^{\infty} \sum_{k=0}^{N-1} H_{d-k, k}^{(n_r, n_t)}(i, q) [\mathbf{c}_k(i-q)]_{n_t-1} s_k(i-q) + w_d^{(n_r)}(i). \quad (2.41)$$

In (2.41), the subcarrier coupling coefficient  $H_{d,k}^{(n_r, n_t)}(i, q)$  is the response at subcarrier index  $k+d$  at the  $n_r$ -th antenna to an impulse transmitted on subcarrier  $k$  and multicarrier symbol  $i-q$  at the  $n_t$ -th transmit antenna. Then in similar fashion as (2.17), we obtain

$$H_{d,k}^{(n_r, n_t)}(i, q) = \frac{1}{N} \sum_{n=-\infty}^{\infty} \sum_{l=0}^{N_h-1} h_{n,l}^{(n_r, n_t)}(i) b_n a_{qN_s+n-l} e^{-j\frac{2\pi}{N} [d(n-N_o)+k(l-qN_s)]}. \quad (2.42)$$

Using (2.42), we define the subcarrier coupling matrix for the  $n_r$ -th receive and  $n_t$ -th transmit antenna pair element-wise as  $[\mathbf{H}^{(n_r, n_t)}(i, q)]_{m_1, m_2} = H_{m_1, m_1-m_2}^{(n_r, n_t)}(i, q)$ , and the MIMO subcarrier coupling matrix  $\mathbf{H}(i, q)$  by

$$\mathbf{H}(i, q) = \sum_{n_t=1}^{N_t} \sum_{n_r=1}^{N_r} \mathbf{H}^{(n_r, n_t)}(i, q) \otimes [\mathbf{e}_{N_r}(n_r) \mathbf{e}_{N_t}(n_t)^T]. \quad (2.43)$$

With the definitions of  $\mathbf{r}(i)$ ,  $\mathbf{w}(i)$ , and  $\underline{\mathbf{s}}^{\text{ma}}(i)$  in (2.33),(2.34) and (2.39), respectively, we obtain

$$\mathbf{r}(i) = \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \underline{\mathbf{s}}^{\text{ma}}(i - q) + \mathbf{w}(i) \quad (2.44)$$

$$= \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \mathcal{D}_b(\mathbf{c}_0(i - q), \dots, \mathbf{c}_{N-1}(i - q)) \mathbf{s}(i - q) + \mathbf{w}(i). \quad (2.45)$$

We again assume that all spatial paths experience uncorrelated fades, so that

$$\mathbb{E}[h_{n_1, l_1}^{(n_{r_1}, n_{t_1})} h_{n_2, l_2}^{(n_{r_2}, n_{t_2})}] = \sigma_{l_1}^2 \phi_{(n_1 - n_2)} \delta_{(l_1 - l_2)} \delta_{(n_{r_1} - n_{r_2})} \delta_{(n_{t_1} - n_{t_2})}. \quad (2.46)$$

This concludes our discussion on modeling beamforming MIMO systems.

## CHAPTER 3

### TRANSMISSION AND RECEPTION STRATEGY

This chapter is devoted to describe processing done at the transmitter and receiver of a SIMO-MCM system. We begin by a description of the pilot and code structure, and follow it up by describing the reception strategy. The penultimate section presents details of the channel estimation algorithm used. Finally, we formally define the LSP constraint and develop a model to investigate its performance.

#### 3.1 Code and Pilot Structure

We assume that the transmitter knows the channel and noise statistics, but not their realizations. The transmitter codes information using  $Q_d N$  zero-mean i.i.d Gaussian codebooks of suitably long codeword length  $N_c$ . Rate allocation is detailed in Section 6.2.2. For now, it suffices to say that the rates, functions of the channel and noise statistics, are chosen to facilitate error-free decoding when  $N_c$  is sufficiently large. The generated codewords are then interleaved. The interleaved codewords are modulated into multicarrier symbols using the modulation technique described in Section 2.1. These data-bearing multicarrier symbols are structured into frames by embedding pilot multicarrier symbols. The constructed frames are then transmitted. The details of this transmission strategy will now be provided.

Recall that the channel coefficients are correlated across time. The interleaving is assumed to be of sufficient depth to mitigate the effects of this time domain correlation between the channel coefficients at the input to the decoder.

We define one transmission epoch (depicted in Fig. 3.1) to consists of a sequence of  $N_c$  frames, where each frame comprises of  $Q_c$  multicarrier symbols. To facilitate the acquisition of channel state information (CSI) at the receiver,  $Q_p = Q_c - Q_d$  multicarrier symbols out of the  $Q_c$  multicarrier symbols in a frame are assumed to be known a priori at the receiver as pilot symbols, whereas the other  $Q_d$  multicarrier symbols in the frame bear coded data. The arrangement of the pilot and data multicarrier symbols in each frame can be described as follows: The  $Q_d N$  scalar data symbols transmitted during a frame consist of one symbol from each of the  $Q_d N$  codewords. The pilot symbol indices,  $\mathcal{I}_p$ , and data symbol indices,  $\mathcal{I}_d$ , form a partition of  $\{0, \dots, Q_c - 1\}$  with  $|\mathcal{I}_p| = Q_p$  and  $|\mathcal{I}_d| = Q_d$ . We assume w.l.o.g. that the first multicarrier symbol in each frame is a pilot, i.e.,  $0 \in \mathcal{I}_p$ , and that there exists at least  $\max(L_{\text{pre}}, L_{\text{pst}})$  data multicarrier symbols between every pair of pilot multicarrier symbols. This should be easy to satisfy since typically  $L_{\text{pre}} = L_{\text{pst}} = 1$ . The arrangement of scalar symbols can be described as follows:  $s_k(i_f Q_c + i_o)$ , the scalar symbol modulated on the  $k$ -th subcarrier of the  $i_o$ -th multicarrier symbol in the  $i_f$ -th frame, is a scalar symbol of the  $(\psi(i_o)N + k)$ -th codeword when  $i_o \in \mathcal{I}_d$ . Otherwise, when  $i_o \in \mathcal{I}_p$ ,  $s_k(i_f Q_c + i_o)$  is a pilot symbol. Here  $\psi : \mathcal{I}_d \rightarrow \{0, \dots, Q_d - 1\}$  is a one-to-one mapping that enumerates the data multicarrier symbols in a frame. Clearly, then,  $Q_d N$  complete codewords are transmitted over each epoch of  $N_c$  frames.

In this dissertation, a transmit power constraint ensures that  $\text{E}\{\mathbf{s}(i_1)\mathbf{s}(i_2)^H\} = \mathbf{I}_N \delta_{i_1 - i_2}$  when  $i_1, i_2 \in \mathcal{I}_d$  and  $\|\mathbf{s}(i)\|^2 = N$  when  $i \in \mathcal{I}_p$ .

## 3.2 Pilot-Aided Reception

The receiver is assumed to know the channel and noise statistics, but not their realizations, and as well as the transmitted pilot symbols. At each multicarrier symbol interval, the receiver collects a multicarrier observation vector  $\mathbf{r}(i)$ . Once the multicarrier observation vectors  $\{\mathbf{r}(i)\}_{i=0}^{Q_c N_c - 1}$  for an  $N_c$  frame epoch has been collected, the receiver generates pilot-aided LMMSE estimates of the dominant ICI coefficients of data-bearing multicarrier symbols, i.e.,  $\{\hat{\mathbf{H}}(i_f Q_c + i_o) : 0 \leq i_f < N_c, i_o \in \mathcal{I}_d\}$  using the procedure detailed in Section 3.3. These estimates are then used to (coherently) decode data on the subcarriers sequentially (i.e., in the order  $k = 0, 1, 2, \dots$ ). To decode the  $Q_d$  codewords transmitted on the  $k^{\text{th}}$  subcarrier, the receiver generates estimates of  $\{s_k(i_f Q_c + i_o) : 0 \leq i_f < N_c, i_o \in \mathcal{I}_d\}$ , one-at-a-time under an LSP constraint (described in Section 3.4), deinterleaves the sequence of estimates and, finally, sorts and sends these estimates to  $Q_d$  optimal decoders. Recall that the code rates are chosen with the aim of facilitating error-free decoding. Hence, previously decoded subcarriers are accurately known. The detection of data transmitted on the  $k$ -th subcarrier  $\{s_k(i_f Q_c + i_o) : 0 \leq i_f < N_c, i_o \in \mathcal{I}_d\}$  is consequently aided by the knowledge of these previously decoded subcarriers  $\{s_d(i_f Q_c + i_o) : 0 \leq i_f < N_c, i_o \in \mathcal{I}_d\}_{d=0}^{k-1}$ .

## 3.3 Channel Estimation

Here we outline the salient features of pilot-aided LMMSE estimation of the ICI coefficients. Recall that we assumed uncorrelated spatial paths as well as independent noise at each receive antenna. Then the estimation of channel coefficients can be decoupled into  $N_r$  estimations, one for each receive antenna. Here, we consider the

estimation of channel coefficients for the  $n_r$ -th receive antenna,  $n_r \in \{1, \dots, N_r\}$ , for generality.

For channel estimation purposes, we find it useful to define

$$\mathbf{g}^{(n_r)}(i) := [\text{diag}_0(\mathbf{H}^{(n_r)}(i, 0))^T \cdots \text{diag}_{N-1}(\mathbf{H}^{(n_r)}(i, 0))^T]^T, \quad (3.1)$$

where  $\text{diag}_k(\cdot)$  creates a vector from the  $k^{\text{th}}$  sub-diagonal of its matrix argument via  $\text{diag}_k(\mathbf{H}) := [[\mathbf{H}]_{k,0}, [\mathbf{H}]_{k+1,1}, \dots, [\mathbf{H}]_{k+N-1,N-1}]^T$  with modulo- $N$  indexing assumed. Using  $\hat{\mathbf{g}}^{(n_r)}(i)$  to denote the pilot-aided LMMSE estimate of  $\mathbf{g}^{(n_r)}(i)$ , and  $\tilde{\mathbf{g}}^{(n_r)}(i) := \mathbf{g}^{(n_r)}(i) - \hat{\mathbf{g}}^{(n_r)}(i)$  to denote the corresponding estimation error, the CEE term can be written as

$$\mathbf{w}_{\text{cee}}^{(n_r)}(i) = \mathbf{S}(i)\tilde{\mathbf{g}}^{(n_r)}(i) \quad (3.2)$$

$$\mathbf{S}(i) := [\Theta^0 \mathcal{D}(\mathbf{s}(i)) \cdots \Theta^{N-1} \mathcal{D}(\mathbf{s}(i))], \quad (3.3)$$

where

$$\Theta := \begin{bmatrix} \mathbf{0}_{N-1}^T & 1 \\ \mathbf{I}_{N-1} & \mathbf{0}_{N-1} \end{bmatrix} \quad (3.4)$$

is the cyclic down-shift matrix, and  $\mathcal{D}(\cdot)$  creates a diagonal matrix from its vector argument.

The ICI coefficients for the data multicarrier symbols in each frame are estimated jointly, and each frame is handled individually. Recall from Section 3.1 the assumption that the first multicarrier symbol in each frame is a pilot multicarrier symbol. We use the pilots in the  $i_f$ -th frame as well as the first (pilot) multicarrier symbol of the  $(i_f + 1)$ -th frame to estimate the ICI coefficients for the data multicarrier symbols in the  $i_f$ -th frame. Specifically, for the  $i_f$ -th frame, we estimate  $\{\mathbf{g}^{(n_r)}(i) : i = i_f Q_c + i_o, i_o \in \mathcal{I}_d\}$  from the pilot observations  $\{\mathbf{r}^{(n_r)}(i) : i = i_f Q_c + i_o, i_o \in \mathcal{I}'_p\}$ , where

$\mathcal{I}'_p := \mathcal{I}_p \cup \{Q_c\}$ . Since each frame is treated identically, we consider the  $i_f = 0$ -th frame w.l.o.g. here on.

Note that, with a causal length- $N_a$  modulation pulse, the transmitted sequence  $\{t_n\}$  for the  $Q_c + 1$  multicarrier symbols with index  $i \in \{0, \dots, Q_c\}$  is causal with length  $N_t := N_a + Q_c N_s$ . Thus, from (2.2), the antenna- $n_r$  time-domain channel coefficients that affect these  $(Q_c + 1)$  multicarrier symbols are  $\{h_{n,l}^{(n_r)}\}_{n=0}^{N_t-1}$  for  $l \in \{0, \dots, N_h - 1\}$ . For later use, we collect them into  $\underline{\mathbf{h}}^{(n_r)} \in \mathbb{C}^{N_t N_h}$ , defined element-wise as

$$[\underline{\mathbf{h}}^{(n_r)}]_n = h_{\langle n \rangle_{N_t}, \lfloor \frac{n}{N_t} \rfloor}^{(n_r)}. \quad (3.5)$$

Note that  $\Sigma_{\underline{\mathbf{h}}^{(n_r)}} := \mathbb{E}\{\underline{\mathbf{h}}^{(n_r)} \underline{\mathbf{h}}^{(n_r)H}\}$  is easily written in terms of  $\{\sigma_l^2\}_{l=0}^{N_h-1}$  and  $\{\phi_n\}_{n=0}^{N_t-1}$ . For  $i \in \mathcal{I}'_p$ ,  $\mathbf{S}(i)$  is a known pilot, and (2.20) and (A.9) imply

$$\mathbf{r}^{(n_r)}(i) = \mathbf{S}(i) \mathbf{g}^{(n_r)}(i) + \mathbf{w}_{\text{isi}}^{(n_r)}(i) + \mathbf{w}^{(n_r)}(i). \quad (3.6)$$

The ICI coefficients can be related to the time-domain coefficients via

$$\mathbf{g}^{(n_r)}(i) = \mathbf{C}(i) \underline{\mathbf{h}}^{(n_r)}, \quad (3.7)$$

where (2.17) implies that  $\mathbf{C}(i) \in \mathbb{C}^{N^2 \times N_t N_h}$  obeys

$$[\mathbf{C}(i)]_{n,l} := \frac{1}{N} b_{\langle l \rangle_{N_t} - i N_s} a_{\langle l \rangle_{N_t} - i N_s - \lfloor \frac{l}{N_t} \rfloor} e^{-j \frac{2\pi}{N} \lfloor \frac{l}{N_t} \rfloor n} e^{-j \frac{2\pi}{N} \lfloor \frac{n}{N} \rfloor (\langle l \rangle_{N_t} - i N_s - N_o)}. \quad (3.8)$$

Using  $\mathbf{P}(i) := \mathbf{S}(i) \mathbf{C}(i)$ , we can rewrite (2.20) as

$$\mathbf{r}^{(n_r)}(i) = \mathbf{P}(i) \underline{\mathbf{h}}^{(n_r)} + \mathbf{w}_{\text{isi}}^{(n_r)}(i) + \mathbf{w}^{(n_r)}(i). \quad (3.9)$$

Stacking the pilot observations  $\{\mathbf{r}^{(n_r)}(i)\}_{i \in \mathcal{I}'_p}$  into the vector  $\underline{\mathbf{r}}^{(n_r)} \in \mathbb{C}^{N(Q_p+1)}$ , the ISI and noise terms into  $\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}$  and  $\underline{\mathbf{w}}^{(n_r)}$ , respectively, and  $\{\mathbf{P}(i)\}_{i \in \mathcal{I}'_p}$  into the block

diagonal matrix  $\underline{\mathbf{P}} \in \mathbb{C}^{N(Q_p+1) \times N_t N_h}$ , we have

$$\underline{\mathbf{r}}^{(n_r)} = \underline{\mathbf{P}}\underline{\mathbf{h}}^{(n_r)} + \underline{\mathbf{w}}_{\text{isi}}^{(n_r)} + \underline{\mathbf{w}}^{(n_r)}. \quad (3.10)$$

Recall that the goal is the estimation of  $\{\mathbf{g}^{(n_r)}(i)\}_{i \in \mathcal{I}_d}$  from  $\underline{\mathbf{r}}^{(n_r)}$ . Due to the linear relationship in (3.7), the LMMSE estimates  $\hat{\mathbf{g}}^{(n_r)}(i)$  and  $\hat{\underline{\mathbf{h}}}^{(n_r)}$  are related via

$$\hat{\mathbf{g}}^{(n_r)}(i) = \mathbf{C}(i)\hat{\underline{\mathbf{h}}}^{(n_r)}. \quad (3.11)$$

Thus, we first focus on the estimation of  $\underline{\mathbf{h}}^{(n_r)}$ . It is well known [73] that the LMMSE estimate

$$\hat{\underline{\mathbf{h}}}^{(n_r)} = \underline{\Sigma}_{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{r}}^{(n_r)}} \underline{\Sigma}_{\underline{\mathbf{r}}^{(n_r)}}^{-1} \underline{\mathbf{r}}^{(n_r)}, \quad (3.12)$$

where

$$\underline{\Sigma}_{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{r}}^{(n_r)}} := \mathbb{E}\{\underline{\mathbf{h}}^{(n_r)} \underline{\mathbf{r}}^{(n_r)H}\}, \quad (3.13)$$

$$\underline{\Sigma}_{\underline{\mathbf{r}}^{(n_r)}} := \mathbb{E}\{\underline{\mathbf{r}}^{(n_r)} \underline{\mathbf{r}}^{(n_r)H}\}. \quad (3.14)$$

Consequently, the estimation error covariance is given by

$$\underline{\Sigma}_{\hat{\underline{\mathbf{h}}}^{(n_r)}} = \underline{\Sigma}_{\underline{\mathbf{h}}^{(n_r)}} - \underline{\Sigma}_{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{r}}^{(n_r)}} \underline{\Sigma}_{\underline{\mathbf{r}}^{(n_r)}}^{-1} \underline{\Sigma}_{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{r}}^{(n_r)}}^H. \quad (3.15)$$

To proceed further, we must investigate the joint statistics of  $\{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{w}}_{\text{isi}}^{(n_r)}, \underline{\mathbf{w}}^{(n_r)}\}$ . Recall the assumption that there exist at least  $\max(L_{\text{pre}}, L_{\text{pst}})$  data symbols between any pair of pilot symbols. Then, for any  $i \in \mathcal{I}'_p$ ,  $\mathbf{w}_{\text{isi}}^{(n_r)}(i)$  is a function of the data multicarrier symbols  $\{\mathbf{s}(i)\}_{i \in \mathcal{I}_d}$  but not the pilot multicarrier symbols  $\{\mathbf{s}(i)\}_{i \in \mathcal{I}'_p}$ . Also recall that the channel, the codebooks and the additive noise are zero mean Gaussian distributed and are mutually independent. Then,  $\{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{w}}_{\text{isi}}^{(n_r)}, \underline{\mathbf{w}}^{(n_r)}\}$  are mutually uncorrelated. Moreover,  $\underline{\Sigma}_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} := \mathbb{E}\{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)} \underline{\mathbf{w}}_{\text{isi}}^{(n_r)H}\}$  can be constructed from covariance matrices in the set  $\{\underline{\Sigma}_{\mathbf{w}_{\text{isi}}^{(n_r)}(i_1), \mathbf{w}_{\text{isi}}^{(n_r)}(i_2)} := \mathbb{E}\{\mathbf{w}_{\text{isi}}^{(n_r)}(i_1) \mathbf{w}_{\text{isi}}^{(n_r)}(i_2)H\} : i_1, i_2 \in \mathcal{I}'_p\}$ .

Note that  $\Sigma_{\mathbf{w}_{\text{isi}}^{(n_r)}(i_1), \mathbf{w}_{\text{isi}}^{(n_r)}(i_2)} \neq \mathbf{0}$  only when  $|i_1 - i_2| < L_{\text{pre}} + L_{\text{pst}}$ . Now, consider constructing  $\Sigma_{\underline{\mathbf{w}}^{(n_r)}} := \text{E}\{\underline{\mathbf{w}}^{(n_r)} \underline{\mathbf{w}}^{(n_r)H}\}$ . Noting that  $\text{E}\{\mathbf{w}^{(n_r)}(i_1) \mathbf{w}^{(n_r)}(i_2)^H\} = \mathbf{0}$  if  $|i_1 - i_2| \geq L_{\text{pre}}$  and that there are at least  $L_{\text{pre}} \leq \max(L_{\text{pre}}, L_{\text{pst}})$  data multicarrier symbols between every pair of pilot multicarrier symbols,  $\text{E}\{\mathbf{w}^{(n_r)}(i_1) \mathbf{w}^{(n_r)}(i_2)^H\} = \mathbf{0}$  for all  $\{i_1, i_2 : i_1 \neq i_2\} \subset \mathcal{I}'_p$ . Then

$$\Sigma_{\underline{\mathbf{w}}^{(n_r)}} = \mathbf{I}_{(Q_p+1)} \otimes \Sigma_{\mathbf{w}^{(n_r)}}. \quad (3.16)$$

The MMSE quantities can then be written

$$\Sigma_{\underline{\mathbf{h}}^{(n_r)}, \underline{\mathbf{r}}^{(n_r)}} = \Sigma_{\underline{\mathbf{h}}^{(n_r)}} \underline{\mathbf{P}}^H, \quad (3.17)$$

$$\Sigma_{\underline{\mathbf{r}}^{(n_r)}} = \underline{\mathbf{P}} \Sigma_{\underline{\mathbf{h}}^{(n_r)}} \underline{\mathbf{P}}^H + \Sigma_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} + \Sigma_{\underline{\mathbf{w}}^{(n_r)}}. \quad (3.18)$$

Taking the eigenvalue decomposition  $\Sigma_{\underline{\mathbf{h}}^{(n_r)}} = \mathbf{U}_h \Lambda_h \mathbf{U}_h^H$ , with invertible  $\Lambda_h$  and  $\mathbf{U}_h^H \mathbf{U}_h = \mathbf{I}$ , we see that the estimation error covariance can be obtained as

$$\begin{aligned} \Sigma_{\tilde{\underline{\mathbf{h}}}^{(n_r)}} &= \Sigma_{\underline{\mathbf{h}}^{(n_r)}} - \Sigma_{\underline{\mathbf{h}}^{(n_r)}} \underline{\mathbf{P}}^H \times \\ &\quad \left( \underline{\mathbf{P}} \Sigma_{\underline{\mathbf{h}}^{(n_r)}} \underline{\mathbf{P}}^H + \Sigma_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} + \Sigma_{\underline{\mathbf{w}}^{(n_r)}} \right)^{-1} \underline{\mathbf{P}} \Sigma_{\underline{\mathbf{h}}^{(n_r)}} \end{aligned} \quad (3.19)$$

$$\begin{aligned} &= \mathbf{U}_h \left( \Lambda_h - \Lambda_h \mathbf{U}_h^H \underline{\mathbf{P}}^H \times \right. \\ &\quad \left. \left( \underline{\mathbf{P}} \mathbf{U}_h \Lambda_h \mathbf{U}_h^H \underline{\mathbf{P}}^H + \Sigma_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} + \Sigma_{\underline{\mathbf{w}}^{(n_r)}} \right)^{-1} \underline{\mathbf{P}} \mathbf{U}_h \Lambda_h \right) \mathbf{U}_h^H \end{aligned} \quad (3.20)$$

$$= \mathbf{U}_h \left( \Lambda_h^{-1} + \mathbf{U}_h^H \underline{\mathbf{P}}^H \left( \Sigma_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} + \Sigma_{\underline{\mathbf{w}}^{(n_r)}} \right)^{-1} \underline{\mathbf{P}} \mathbf{U}_h \right)^{-1} \mathbf{U}_h^H, \quad (3.21)$$

where for (3.21) we used the matrix inversion lemma. To ensure that the estimation error varies inversely with SNR in Regime 1 (see Section 4.2), we assume that pilots are chosen so that  $\underline{\mathbf{P}} \mathbf{U}_h$  is full column rank (FCR). While specifying a design procedure for  $\{\mathbf{S}(i)\}_{i \in \mathcal{I}_p}$  that guarantees FCR  $\underline{\mathbf{P}} \mathbf{U}_h$  is outside the scope of this dissertation, it is easy to verify whether a particular design will suffice, since  $\mathbf{U}_h$  and  $\{\mathbf{C}(i)\}_{i \in \mathcal{I}_p}$  are

known. Finally, the estimation error for the ICI coefficients is given by

$$\mathbb{E}\{\tilde{\mathbf{g}}^{(n_r)}(i)\tilde{\mathbf{g}}^{(n_r)}(i)^H\} = \mathbf{C}(i)\boldsymbol{\Sigma}_{\underline{\mathbf{k}}^{(n_r)}}\mathbf{C}(i)^H. \quad (3.22)$$

### 3.4 Local Subcarrier Processing

Our goal is to analyze the performance limits imposed by the *local subcarrier processing* (LSP) constraint on multicarrier reception. The LSP constraint allows use of only the local observations <sup>1</sup>  $\{r_{k-D_r}^{(n_r)}, \dots, r_{k+D_r}^{(n_r)}\}_{m=1}^{N_r}$  when estimating  $s_k(i)$ . Clearly, increasing (decreasing)  $D_r$  results in an increase (decrease) in complexity as well as performance. Hence,  $0 \leq D_r < \lfloor \frac{N-1}{2} \rfloor$  can be interpreted as a processing radius that affects both performance and complexity. This is in contrast to global subcarrier processing (GSP) reception where the “global observations”  $\{r_0^{(n_r)}, \dots, r_{N-1}^{(n_r)}\}_{m=1}^{N_r}$  are used when estimating  $s_k(i)$ . For practical reasons, we will assume the availability of pilot-aided estimates of only the ICI coefficients covering a radius of  $D_h \geq 0$  subcarriers, i.e.,  $\{H_{d,k}^{(n_r)}(i, 0)\}_{d=-D_h}^{D_h}$ . Henceforth, we refer to these channel coefficients as the *dominant* ones. We demonstrate in Section 4.2 that given a band-limited Doppler spectrum, it is possible to choose the dominant ICI (DICI) radius  $D_h$  in conjunction with the MCM design parameters  $\{b_n, a_n, N, N_s, N_o\}$  and the pilot pattern  $\{\mathbf{s}(i)\}_{i \in \mathcal{I}_p}$  to make interference from the *residual* channel coefficients arbitrarily small. In practice, these residual coefficients are ignored in receiver processing either to reduce computation or because they are below the resolution allowed by quantization for fixed point processing. From our viewpoint, dominant channel-coefficient estimation and the LSP constraint captures the essence of practical multicarrier reception. For

<sup>1</sup>Throughout, the subcarrier index is treated as cyclic with period  $N$ .

example, classical CP-OFDM reception for time-invariant channels [72] is characterized by  $D_r = D_h = 0$ , in which case the complexity of estimating  $\mathbf{s}(i)$  is  $\mathcal{O}(N)$ . Here, we do not include the complexity incurred in the computation of  $\mathbf{r}^{(n_r)}(i)$ , which typically consumes  $\mathcal{O}(N \log N)$  operations via FFT.

We now provide a more explicit description of LSP. Say that  $\hat{\mathbf{H}}^{(n_r)}(i)$  denotes a  $D_h$ -quasi-banded estimate of  $\mathbf{H}^{(n_r)}(i, 0)$ , such that elements inside the shaded region of Fig. 4.1 are estimated from pilots and those outside the shaded region are set to zero. Note that  $\hat{\mathbf{H}}^{(n_r)}(i)$  contains estimates of the dominant ICI coefficients, i.e., those within a radius of  $D_h$  subcarriers. Then (2.20) can be written as

$$\mathbf{r}^{(n_r)}(i) = \hat{\mathbf{H}}^{(n_r)}(i)\mathbf{s}(i) + \mathbf{v}^{(n_r)}(i) \quad (3.23)$$

$$\mathbf{v}^{(n_r)}(i) = \mathbf{w}_{\text{cee}}^{(n_r)}(i) + \mathbf{w}_{\text{rici}}^{(n_r)}(i) + \mathbf{w}_{\text{isi}}^{(n_r)}(i) + \mathbf{w}^{(n_r)}(i), \quad (3.24)$$

where  $\mathbf{v}^{(n_r)}(i)$  is treated as interference when decoding. In (3.24), this interference is partitioned into channel estimation error (CEE), residual ICI (RICI), ISI, and additive noise. We give explicit expressions for, and characterize the statistics of, these terms in Appendix A. As in Section 2.5.1, we can combine the observations, frequency domain noise and channel coefficients across the  $N_r$  antennas via

$$\mathbf{r}(i) := \sum_{n_r=1}^{N_r} \mathbf{r}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}, \quad (3.25)$$

$$\mathbf{v}(i) := \sum_{n_r=1}^{N_r} \mathbf{v}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}, \quad (3.26)$$

$$\hat{\mathbf{H}}(i) := \sum_{n_r=1}^{N_r} \hat{\mathbf{H}}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}. \quad (3.27)$$

Using the definitions above, we obtain

$$\mathbf{r}(i) = \hat{\mathbf{H}}(i)\mathbf{s}(i) + \mathbf{v}(i). \quad (3.28)$$

As previously defined, an LSP-constrained estimate of  $s_k(i)$  is generated using received subcarriers within a “processing radius” of  $D_r$  around the  $k^{\text{th}}$  subcarrier, i.e.,  $\mathbf{r}_k^{(n_r)}(i) := [r_{k-D_r}^{(n_r)}, \dots, r_{k+D_r}^{(n_r)}]^T$  for  $n_r \in \{1, \dots, N_r\}$ . Due to the  $D_h$ -quasi-banded structure of  $\hat{\mathbf{H}}^{(n_r)}(i)$ , only the symbols  $\mathbf{s}_k(i) := [s_{k-D_h-D_r}, \dots, s_{k+D_h+D_r}]^T$  contribute (coherently) to  $\mathbf{r}_k^{(n_r)}(i)$ , so that we can write

$$\mathbf{r}_k^{(n_r)}(i) = \hat{\mathbf{H}}_k^{(n_r)}(i)\mathbf{s}_k(i) + \mathbf{v}_k^{(n_r)}(i), \quad (3.29)$$

where  $\mathbf{v}_k^{(n_r)}(i) := [v_{k-D_r}^{(n_r)}, \dots, v_{k+D_r}^{(n_r)}]^T$  and  $\hat{\mathbf{H}}_k^{(n_r)}(i)$  is the sub-matrix composed from rows  $\{k - D_r, \dots, k + D_r\}$  and columns  $\{k - D_h - D_r, \dots, k + D_h + D_r\}$  of  $\hat{\mathbf{H}}^{(n_r)}(i)$ . Combining the receptions from the  $N_r$  antennas, we get the LSP observation

$$\mathbf{r}_k(i) = \hat{\mathbf{H}}_k(i)\mathbf{s}_k(i) + \mathbf{v}_k(i). \quad (3.30)$$

The terms in (3.30) are defined in similar fashion to the terms in (3.28) as

$$\mathbf{r}_k(i) := \sum_{n_r=1}^{N_r} \mathbf{r}_k^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}, \quad (3.31)$$

$$\mathbf{v}_k(i) := \sum_{n_r=1}^{N_r} \mathbf{v}_k^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}, \quad (3.32)$$

$$\hat{\mathbf{H}}_k(i) := \sum_{n_r=1}^{N_r} \hat{\mathbf{H}}_k^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}. \quad (3.33)$$

It will sometimes be useful to write

$$\mathbf{r}_k(i) = \hat{\mathbf{h}}_{k,k} s_k(i) + \sum_{d \in \mathcal{K}_k^d} \hat{\mathbf{h}}_{k,d}(i) s_d(i) + \mathbf{v}_k(i), \quad (3.34)$$

where  $\mathcal{K}_k^d$  contains the subcarrier indices that contribute dominant ICI (DICI) to  $\mathbf{r}_k(i)$  and can be formally expressed as

$$\mathcal{K}_k^d := \{k - D_h - D_r, \dots, k + D_h + D_r\} \setminus \{k\}. \quad (3.35)$$

Moreover,  $\hat{\mathbf{h}}_{k,d}(i)$  characterizes the influence of  $s_d(i)$  on  $\mathbf{r}_k(i)$ , and can be expressed in terms of its single antenna components as

$$\hat{\mathbf{h}}_{k,d}(i) = \sum_{m=1}^{N_r} \hat{\mathbf{h}}_{k,d}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}. \quad (3.36)$$

In (3.36),  $\hat{\mathbf{h}}_{k,d}^{(n_r)}(i) \in \mathbb{C}^{(2D_r+1)N_r}$  is composed from rows  $\{k - D_r, \dots, k + D_r\}$  and column  $d$  of  $\hat{\mathbf{H}}^{(n_r)}(i)$ . It will also be convenient to store the dominant interference subcarriers  $\{s_d(i)\}_{d \in \mathcal{K}_k^d}$  in the vector  $\bar{\mathbf{s}}_k(i) \in \mathbb{C}^{2(D_h+D_r)}$  and the DICI coefficient vectors  $\{\hat{\mathbf{h}}_{k,d}(i)\}_{d \in \mathcal{K}_k^d}$  in the corresponding dominant ICI matrix  $\hat{\mathbf{H}}_k(i)$ , so that

$$\mathbf{r}_k(i) = \hat{\mathbf{h}}_{k,k} s_k(i) + \hat{\mathbf{H}}_k(i) \bar{\mathbf{s}}_k(i) + \mathbf{v}_k(i). \quad (3.37)$$

As before, expressing in terms of single antenna components, we obtain

$$\hat{\mathbf{H}}_k(i) = \sum_{m=1}^{N_r} \hat{\mathbf{H}}_k^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}. \quad (3.38)$$

Figure 3.2 illustrates the structure of  $\hat{\mathbf{H}}_k(i)$  and its various component matrices used for modeling reception under the LSP constraint. This completes the discussion on modeling reception under the LSP constraint.

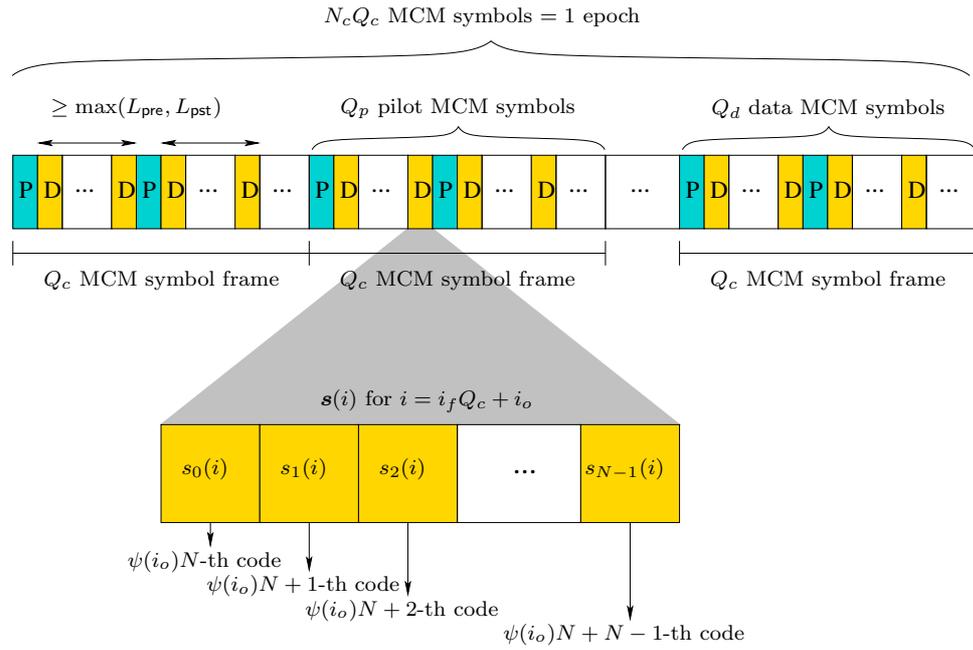


Figure 3.1: Transmission epoch and frame structure. Pilot MCM symbols and data MCM symbols are denoted by P and D, respectively.

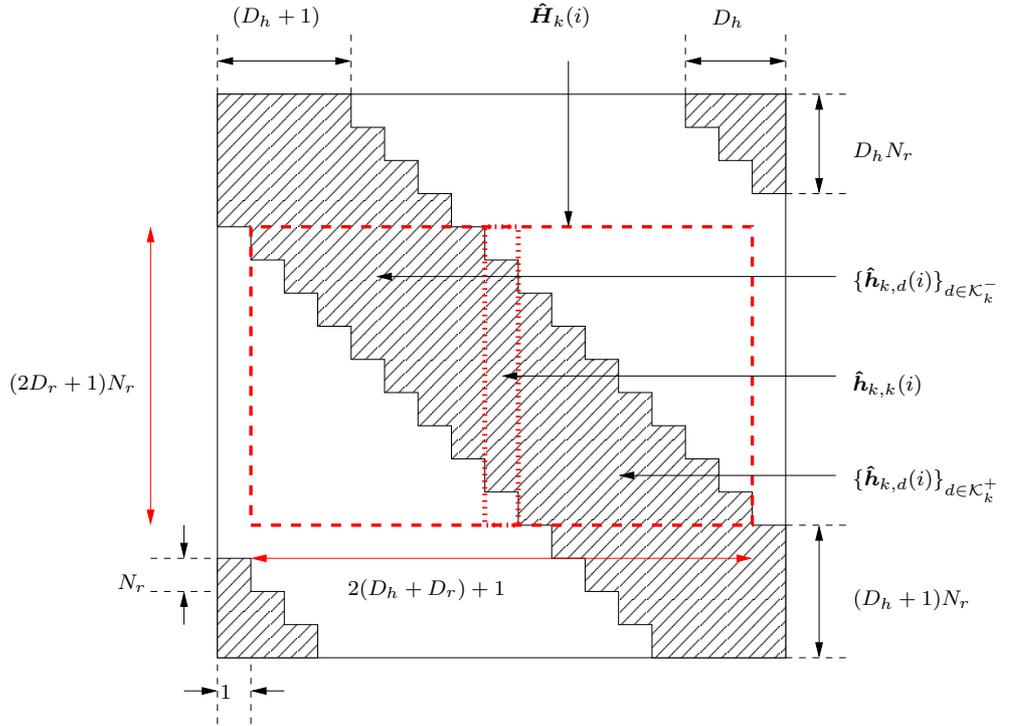


Figure 3.2: Estimation for the  $k^{\text{th}}$  subcarrier under the LSP constraint only uses coefficients within the dashed rectangle, i.e., the matrix  $\hat{\mathbf{H}}_k(i)$ . After the column vector  $\hat{\mathbf{h}}_{k,k}(i)$  (dotted rectangle) is removed, the remaining columns within  $\hat{\mathbf{H}}_k(i)$  form the dominant ICI matrix  $\hat{\hat{\mathbf{H}}}_k(i)$ .

## CHAPTER 4

### LSP-CONSTRAINED PERFORMANCE ANALYSIS

The LSP constraint for MCM reception was introduced in Chapter 3. We now characterize the ergodic achievable-rate of SIMO-MCM systems under the LSP constraint. In this direction, we first introduce a “surrogate” system model that makes this analysis tractable. Recall from Chapter 3 that the interference observed in LSP-based decoding has many constituents. We define SNR regimes to characterize the effect of the interference (and its constituents) at various SNRs. We then derive a lower-bound on the LSP constrained achievable-rate of SIMO-MCM systems, and use it to evaluate the LSP-constrained spectral efficiency of MCM transmission. We corroborate our results through numerical examples and also discuss an inherent trade-off between performance and computational complexity.

#### 4.1 Analysis Via Surrogate System

To understand the performance limitations of LSP, we pursue a lower-bound to the ergodic achievable rate of LSP-constrained schemes, motivated by the availability of “good” coding strategies. Furthermore, we focus on high-SNR since this is the practical operating regime for most MCM systems.

In deriving our lower-bounds, we assume the use of the Gaussian-coded pilot-aided transmission strategy described in Section 3.1 and the reception strategy described in Section 3.2. Even with these simplifications, the statistical properties of the interference  $\mathbf{v}^{(m)}(i)$  in (2.20) make achievable-rate analysis difficult. In particular, Appendix A shows that the sequence of interference vectors  $\{\mathbf{v}^{(m)}(i)\}$  is not quite Gaussian, mutually uncorrelated, and uncorrelated with the signal component  $\hat{\mathbf{H}}^{(m)}(i)\mathbf{s}(i)$ . Thus, to proceed, we consider a *surrogate system* wherein  $\{\mathbf{v}^{(m)}(i)\}$  is replaced by sequence of uncorrelated Gaussian random vectors that otherwise preserve the statistics of  $\mathbf{v}^{(m)}(i)$ . In Appendix B.1, we detail the steps that lead us to this surrogate-based approximation of the system model. We use the surrogate model to derive expressions for, and study the high SNR asymptotes of, lower-bounds on the LSP-constrained achievable rate via the approach of [40, 42].

## 4.2 SNR Regimes

We define the signal-to-noise ratio (SNR)  $\rho$  as the ratio of received signal power to received additive-noise power at each receive antenna. Because we have assumed a unit-power source, energy preserving modulation and demodulation pulses, an energy-preserving channel, and a  $\sigma^2$ -power additive-noise, SNR  $\rho := \sigma^{-2}$ .

As can be seen from (3.24), as SNR increases, i.e.,  $\sigma^2 \rightarrow 0$ , the interference covariance  $\Sigma_{\mathbf{v}(i)} := \mathbb{E}\{\mathbf{v}(i)\mathbf{v}(i)^H\}$  contains some components that will vanish, but other components (e.g., ISI and RICI) that will not vanish. Formally, we state this as follows.

**Lemma 1.** Consider the MCM system for pilot-aided transmission over the WSSUS doubly selective channel. Define  $\sigma$ -invariant  $C$  and  $\epsilon$  as

$$C := \text{tr}(\mathbf{J} \mathcal{D}(\mathbf{b})^2 \mathbf{J}^H) + 2\alpha_p \lambda_b, \quad (4.1)$$

$$\epsilon := \frac{\text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{nci}}(i)}) + 2\alpha_p \lambda_{\text{isi}}}{C}, \quad (4.2)$$

where  $\lambda_b$  is the largest eigenvalue of  $\mathbf{J} \mathcal{D}(\mathbf{b})^2 \mathbf{J}^H$ ,  $\lambda_{\text{isi}}$  the largest eigenvalue of  $\boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)}$  and  $\alpha_p$  is a function of the pilot pattern, the source statistics and the channel statistics.

Then, the interference covariance obeys

$$\text{tr}(\boldsymbol{\Sigma}_{\mathbf{v}(i)}) \leq C(\epsilon + \sigma^2), \quad i \in \mathcal{I}_d. \quad (4.3)$$

*Proof.* See Appendix B.2. □

The term  $C\epsilon$  bounds the power of the non-vanishing component of the interference. We show in the sequel that, over a range of  $\sigma^2$  that includes very small values (i.e., high SNRs), it is possible to guarantee  $\epsilon < \sigma^2$  via suitable choice of MCM design parameters  $\{D_h, b_n, a_n, N, N_s, N_o\}$  and pilot pattern  $\{\mathbf{s}(i) : i \in \mathcal{I}_p\}$ . Thus, the regime where simultaneously  $\epsilon < \sigma^2$  and  $\sigma^2 \rightarrow 0$  (which we call ‘‘Regime 1,’’ reminiscent of [49]) is of practical interest. However, there exists a range of extremely small values of  $\sigma^2$  for which it may be impossible to guarantee  $\epsilon < \sigma^2$  under practical MCM design parameters. Thus, the regime where  $\sigma^2 \rightarrow 0$  with fixed small  $\epsilon$  (which we call ‘‘Regime 2’’) may also be of practical interest. In analyzing the achievable rate, we consider each of these regimes separately.

Lemma 1 also shows that channel estimation error affects the SNR regimes. Consequently, choosing a good pilot pattern is of critical importance.

Note that the boundary between the SNR regimes would depend on the term  $C\epsilon$ . We now outline a procedure for choosing system parameters for a target  $C\epsilon$ .

Recall that the DICl radius  $D_h$  affects  $\mathbf{\Sigma}_{\mathbf{w}_{\text{ricl}(i)}}$ , but not  $\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}(i)}}$  and  $\alpha_p \lambda_{\text{isi}}$ . Thus, first we choose  $\{a_n, b_n, N, N_s, N_o\}$  and a pilot pattern to satisfy  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}(i)}}) + \alpha_p \lambda_{\text{isi}} \leq \frac{C\epsilon}{2}$ . Then, realizing that increasing  $D_h$  will reduce the RICl, we choose a large enough  $D_h$  to ensure that  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{ricl}(i)}}) \leq \frac{C\epsilon}{2}$ . Having seen how to design the system for a target  $C\epsilon$ , we now consider each regime individually.

When operating in Regime 1, additive noise will overshadow RICl and ISI. Thus, both the channel estimation and LSP-constrained decoding will behave as if RICl and ISI were absent and performance will improve with increasing SNR. Recalling that  $\epsilon < \sigma^2$  in Regime 1, we further bound (4.3) as

$$\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)}) \leq 2C\sigma^2, \quad i \in \mathcal{I}_d \quad (4.4)$$

and utilize (4.4) in analyzing the LSP-constrained achievable-rates in Regime 1.

On the other hand, when operating in Regime 2, RICl and ISI will dominate the additive noise. Under these circumstances, the LSP-constrained achievable rate hits a ceiling, i.e., it becomes invariant to changes in SNR. Though an achievable-rate ceiling might seem discouraging, we show by examples in Section 4.5 that Regime 2 is encountered at SNRs well outside the typical operating SNR range of practical wireless systems.

### 4.3 Lower Bound on LSP-Constrained Achievable-Rate

Recall from Section 3.2 that each frame is processed in identical fashion with LSP. Thus the achievable rate of the system can be analyzed by considering the 0-th frame w.l.o.g. The average mutual information between the observations and the

transmitted data multicarrier symbols in a frame can be written as <sup>2</sup>

$$R_{\text{sys}} = \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{r}(i); \mathbf{s}(i)) \quad \text{“per channel use”} \quad (4.5)$$

$$= \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \mathbb{E}_{\hat{\mathbf{H}}(i)} \left[ I(\mathbf{r}(i); s_0(i)) + \sum_{k=1}^{N-1} I(\mathbf{r}(i); s_k(i) | \{s_d(i)\}_{d=0}^{k-1}) \right], \quad (4.6)$$

where (4.6) is obtained by applying the chain rule [74]. However, LSP generates estimates of  $s_k(i)$  using only  $\mathbf{r}_k(i)$ , the observations within a processing radius of  $D_r$  subcarriers around the  $k^{\text{th}}$  subcarrier. Then, the LSP-constrained achievable rate with processing radius  $D_r$  is

$$R_{D_r} = \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \mathbb{E}_{\hat{\mathbf{H}}(i)} \left[ I(\mathbf{r}_0(i); s_0(i)) + \sum_{k=1}^{N-1} I(\mathbf{r}_k(i); s_k(i) | \{s_d(i)\}_{d=0}^{k-1}) \right]. \quad (4.7)$$

We lower-bound the LSP-constrained achievable rate by that of the two-step strategy now described. During the interference-cancellation (IC) step, dominant ICI from previously decoded subcarriers is removed from the observation:

$$\mathbf{y}_k(i) = \mathbf{r}_k(i) - \sum_{d \in \mathcal{K}_k^-} \hat{\mathbf{h}}_{k,d}(i) s_d(i) \quad (4.8)$$

$$= \hat{\mathbf{h}}_{k,k}(i) s_k(i) + \sum_{d \in \mathcal{K}_k^+} \hat{\mathbf{h}}_{k,d}(i) s_d(i) + \mathbf{v}_k(i). \quad (4.9)$$

In (4.8)-(4.9), the set  $\mathcal{K}_k^-$  contains indices of previously decoded subcarriers that cause dominant ICI to  $\mathbf{r}_k(i)$ , whereas  $\mathcal{K}_k^+$  contains indices of subcarriers to be decoded in the future that cause dominant ICI to  $\mathbf{r}_k(i)$ . Recall that  $\mathcal{K}_k^d$ , defined in Section 3.4 as

$$\mathcal{K}_k^d := \{k - D_h - D_r, \dots, k + D_h + D_r\} \setminus \{k\}, \quad (4.10)$$

<sup>2</sup>The mutual information calculated here is conditioned on the estimate  $\hat{\mathbf{H}}(i)$ . However, we omit it from the notation for brevity.

contains indices of all subcarriers that cause dominant ICI to  $\mathbf{r}_k(i)$ . Then  $\mathcal{K}_k^-$  and  $\mathcal{K}_k^+$  can be formally defined as

$$\mathcal{K}_k^- := \mathcal{K}_k^d \cap \{d : 0 \leq d < k\}, \quad (4.11)$$

$$\begin{aligned} \mathcal{K}_k^+ &:= \mathcal{K}_k^d \setminus \mathcal{K}_k^- \\ &= \mathcal{K}_k^d \cap \{d : k < d \leq N - 1\}. \end{aligned} \quad (4.12)$$

After IC, a linear-combination (LC) of the post-IC observation using the combiner  $\mathbf{z}_k(i)$  generates the estimate

$$\hat{\mathbf{s}}_k(i) = \mathbf{z}_k^H(i) \mathbf{y}_k(i). \quad (4.13)$$

The efficacy of the combiner  $\mathbf{z}_k(i)$  is measured by the post-LC signal-to-interference-plus-noise ratio (SINR) defined as

$$\gamma_k(\mathbf{z}_k(i)) = \frac{\mathbf{z}_k^H(i) \hat{\mathbf{h}}_{k,k}(i) \hat{\mathbf{h}}_{k,k}^H(i) \mathbf{z}_k(i)}{\mathbf{z}_k^H(i) \left( \sum_{d \in \mathcal{K}_k^+} \hat{\mathbf{h}}_{k,d}(i) \hat{\mathbf{h}}_{k,d}^H(i) + \Sigma_{\mathbf{v}_k(i)} \right) \mathbf{z}_k(i)}. \quad (4.14)$$

Henceforth, we call this strategy the IC-LC-LSP strategy. An important property of this IC-LC-LSP strategy is the following.

**Theorem 1.** *Consider the surrogate MCM system (i.e., where the interference is Gaussian and uncorrelated with the signal). The IC-LC-LSP strategy using processing radius  $D_r$  and the max-SINR combiner*

$$\mathbf{z}_k^m(i) = \alpha \left( \sum_{d \in \mathcal{K}_k^+} \hat{\mathbf{h}}_{k,d}(i) \hat{\mathbf{h}}_{k,d}^H(i) + \Sigma_{\mathbf{v}_k(i)} \right)^{-1} \hat{\mathbf{h}}_{k,k}(i), \quad (4.15)$$

where  $\alpha$  scales  $\mathbf{z}_k^m(i)$  to ensure  $\|\mathbf{z}_k^m(i)\| = 1$ , achieves the rate  $R_{D_r}$  and is achievable-rate-optimal. Then for the surrogate MCM system,

$$R_{D_r} = \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \sum_{k=0}^{N-1} \mathbb{E}_{\hat{\mathbf{H}}(i)} \log \left[ 1 + \hat{\mathbf{h}}_k(i)^H \right. \\ \left. \times \left( \sum_{d \in \mathcal{K}_k^+} \hat{\mathbf{h}}_{k,d}(i) \hat{\mathbf{h}}_{k,d}^H(i) + \Sigma_{\mathbf{v}_k(i)} \right)^{-1} \hat{\mathbf{h}}_k(i) \right]. \quad (4.16)$$

*Proof.* See Appendix B.3. □

## 4.4 Spectral Efficiency

The spectral efficiency of a system with achievable rate  $R(\rho)$  is defined as

$$\eta := \lim_{\rho \rightarrow \infty} \frac{R(\rho)}{\log \rho}. \quad (4.17)$$

We characterize the LSP-constrained spectral efficiency with the aid of the Gaussian-coded pilot-aided surrogate MCM system. Because we are interested in analyzing LSP-constrained frequency domain reception, we focus on the  $D_h > 0$  case, i.e., when there is significant coupling between subcarriers. The degenerate case of  $D_h = 0$ , i.e., when subcarrier coupling is negligible, can be investigated as a case of reception (in correlated noise) of transmission over a time- and frequency-flat fading channel.

We first characterize the LSP-constrained spectral efficiency in Regime 1. Consider the LC-LSP strategy wherein the symbol estimates for the scalar symbols modulated on the  $k$ -th subcarrier are generated by linear-combining of the LSP observation  $\mathbf{r}_k(i)$ , i.e.,  $\hat{s}_k(i) = \mathbf{z}_k(i)^H \mathbf{r}_k(i)$ . In the absence of interference cancellation, the signal, ICI and noise statistics will be uniform across subcarriers. Since the ergodic rates are only a function of channel and noise statistics, all the subcarriers should be allocated identical code rates. The achievable-rate of this LC-LSP strategy on the surrogate

system is then

$$R'_{D_r} = \frac{N}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \mathbb{E}_{\hat{\mathbf{H}}(i)} \log(1 + \gamma'_0(\mathbf{z}_0(i))) \quad (4.18)$$

$$\gamma'_k(\mathbf{z}_k(i)) = \frac{\mathbf{z}_k(i)^H \hat{\mathbf{h}}_{k,k}(i) \hat{\mathbf{h}}_{k,k}(i)^H \mathbf{z}_k(i)}{\mathbf{z}_k(i)^H \left( \hat{\mathbf{H}}_k(i) \hat{\mathbf{H}}_k(i)^H + \Sigma_{\mathbf{v}_k(i)} \right) \mathbf{z}_k(i)}, \quad (4.19)$$

where we used (3.37) to calculate the SINR. Now, consider LC-LSP with a zero-forcing (ZF) combiner  $\mathbf{z}_k^{\text{zf}}(i)$  with  $\|\mathbf{z}_k^{\text{zf}}(i)\| = 1$  that suppresses dominant ICI perfectly, i.e.,  $\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{H}}_k(i) = \mathbf{0}^T$ , and has a non-zero projection along the signal direction, i.e.,  $\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i) \neq 0$ . When such a  $\mathbf{z}_k^{\text{zf}}(i)$  exists, the post-combining SINR can be bounded using (4.4) as

$$\gamma'_k(\mathbf{z}_k^{\text{zf}}(i)) = \frac{\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i) \hat{\mathbf{h}}_{k,k}(i)^H \mathbf{z}_k^{\text{zf}}(i)}{\mathbf{z}_k^{\text{zf}}(i)^H \left( \hat{\mathbf{H}}_k(i) \hat{\mathbf{H}}_k(i)^H + \Sigma_{\mathbf{v}_k(i)} \right) \mathbf{z}_k^{\text{zf}}(i)} \quad (4.20)$$

$$= \frac{\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i) \hat{\mathbf{h}}_{k,k}(i)^H \mathbf{z}_k^{\text{zf}}(i)}{\mathbf{z}_k^{\text{zf}}(i)^H \Sigma_{\mathbf{v}_k(i)} \mathbf{z}_k^{\text{zf}}(i)} \quad (4.21)$$

$$\geq \frac{|\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i)|^2}{2C\sigma^2} \quad (4.22)$$

$$\geq \frac{|\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i)|^2 \rho}{2C}. \quad (4.23)$$

Consequently, the  $k$ -th subcarrier of the  $i$ -th data multicarrier symbol will have a spectral efficiency of 1. Assume that ZF combiners exist for each of the  $N$  subcarriers on each of the  $Q_d$  data multicarrier symbols in each frame. Then the LSP-constrained spectral efficiency of the surrogate system is

$$\lim_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} \geq \frac{Q_d N}{Q_c N_s}. \quad (4.24)$$

Now consider that the spectral efficiency of a genie-aided global subcarrier processing (GGSP) strategy that uses the entire observation  $\mathbf{r}(i)$  to decode each codeword optimally under perfect CSI cannot exceed  $\frac{Q_d N}{Q_c N_s}$  since only  $Q_d N$  codewords are transmitted per frame. Therefore, the existence of ZF combiners for each subcarrier of each

data multicarrier symbol would imply that the LSP-constrained spectral efficiency equals  $\frac{Q_d N}{Q_c N_s}$ . Next, we show that for cases of practical interest, ZF combiners exist only when multiple antennas are used.

**Theorem 2.** *Consider a surrogate MCM system with DICI radius  $D_h > 0$  employing LC-LSP with processing radius  $D_r$ . ZF combiners exist for each subcarrier  $k \in \{0, \dots, N-1\}$  and for each data multicarrier symbol  $i \in \mathcal{I}_d$  with probability one (w.p.1.) if*

$$N_r > 1 + \frac{2D_h - 1}{2D_r + 1}. \quad (4.25)$$

Moreover, for  $N_r = 1$  and any choice of  $D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ , ZF combiners do not exist w.p.1 for any subcarrier  $0 \leq k \leq N-1$  on any data multicarrier symbol  $i \in \mathcal{I}_d$ .

*Proof.* See Appendix B.4 □

Note that  $\frac{2D_h-1}{2D_r+1} > 0$  when  $D_h > 0$ . Then Theorem 2 states that multiple antennas are *sufficient* to ensure the existence of ZF combiners when  $D_h > 0$ . Further, it states that when  $N_r = 1$  and  $D_h > 0$ , there is no hope of finding ZF combiners for LC-LSP for small values of processing radius ( $2D_r + 1 \ll N$ ). When  $D_r > \lfloor \frac{N-1}{2} \rfloor - D_h$ , LSP uses more than  $N - 2D_h$  observations for each subcarrier. Given that  $D_h \ll N$ , these choices of processing radii are of little interest from the LSP perspective. In light of previous discussion, Theorem 2 implies that LSP should achieve the spectral efficiency of  $\frac{Q_d N}{Q_c N_s}$  when  $N_r > 1$  and  $D_h > 0$ . This is established by the following theorem:

**Theorem 3.** *Consider a Gaussian-coded pilot-aided surrogate MCM system with  $N_r > 1$  receive antennas. Let  $Q_d$  data multicarrier symbols be transmitted in each*

frame consisting of  $Q_c$  multicarrier symbols. Let the DICI radius be  $D_h > 0$  and the processing radius be  $D_r \geq D_h$ . Then, in Regime 1, the achievable rate  $R_{D_r}$  of the system obeys

$$\lim_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} = \frac{Q_d N}{Q_c N_s}. \quad (4.26)$$

*Proof.* See Appendix B.5. □

On the other hand, we reckon from previous discussion that the spectral efficiency of LSP on a SISO ( $N_r = 1$ ) system will be less than the corresponding SIMO ( $N_r > 1$ ) system when  $D_h > 0$ . This is established by the following theorem:

**Theorem 4.** *Consider a Gaussian-coded pilot-aided surrogate MCM system with  $N_r = 1$  receive antenna. Let  $Q_d$  data multicarrier symbols be transmitted in each  $Q_c$  multicarrier symbol long frame. Let the DICI radius be  $D_h > 0$  and the processing radius be  $D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ . Then, in Regime 1, the achievable rate  $R_{D_r}$  of this system can be bounded by*

$$\lim_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} < \frac{Q_d N}{Q_c N_s} \quad (4.27)$$

*Proof.* See Appendix B.6. □

Note that this theorem demonstrates that a SISO system has a lower spectral efficiency when Gaussian-codebooks are used. However, the proof technique reveals that this happens because the performance of the system becomes “dominant-ICI limited,” rather than due to properties of Gaussian codebooks. Hence, it is unlikely that non-Gaussian codebooks would achieve the spectral efficiency of  $\frac{Q_d N}{Q_c N_s}$ . However, a detailed proof for all continuously distributed codebooks is beyond the scope of this dissertation.

Theorem 3 and Theorem 4 reveal that when  $N_r > 1$  and  $D_h > 0$ , LSP will attain the same spectral efficiency as a global subcarrier processing (GSP) strategy. It then becomes important to study the performance gap between LSP and GSP in Regime 1 when  $N_r > 1$  and  $D_h > 0$ . We characterize this performance gap between LSP and GSP as a function of the processing radius  $D_r$  through a simulation study in Section 4.5. This concludes the discussion on Regime 1.

In Regime 2, the performance of the MCM system is determined by the power of the non-vanishing components of interference. Thus, we expect LSP to hit an achievable-rate ceiling in this regime, as established by the following theorem:

**Theorem 5.** *Consider a Gaussian-coded pilot-aided surrogate MCM system with  $N_r \geq 1$  receive antennas. Let  $Q_d$  data multicarrier symbols be transmitted in each  $Q_c$  multicarrier symbol long frame. Let the DICI radius be  $D_h > 0$  and the processing radius be  $D_r \geq 0$ . The LSP-constrained achievable-rate of the system is bounded in Regime 2. Consequently, LSP-constrained spectral efficiency in Regime 2 is zero.*

*Proof.* See Appendix B.7. □

This completes the discussion on spectral efficiency.

## 4.5 Numerical Examples

We design and simulate a MCM system to illustrate the obtained theoretical results. We design a MCM system with  $N = 64$  subcarriers without time domain guards, so that  $N_s = N = 64$ . Transmission is over spatially independent, energy preserving DS channels with delay spreads of  $N_h = 8$  chips, uniform power-delay profiles and chip-rate normalized single-sided maximum Doppler spreads of  $F_{\text{Dop}}T_s =$

0.014. These could correspond to, for instance, a system with a bandwidth of 1.5 MHz and a carrier frequency of 60 GHz operating over channels with 5.4  $\mu$ s delay spreads and mobile and reflector velocities of 138 km/h in a “triple-Doppler” scenario [75]. This triple-Doppler scenario arises, e.g., when the mobile is moving away from the transmitter, and the reflector is moving towards the transmitter at 138 km/h. For our tests, the channels are modeled to simulate Jakes’ spectrum [1].

We first design suitable modulation and demodulation pulses. For our system, we use the max-SINR pulse design technique of [23]. We design the system to ensure that Regime 1 extends to at least  $\rho_{\max} = 85$  dB. This also ensures that Regime 2 is sufficiently beyond the operating SNR range of typical wireless systems. We set the max-SINR pulse design SNR to  $\rho_{\max} = 85$  dB. Additionally, following guidelines in [23], we set the modulation and demodulation pulse lengths to  $N_a = 96$  and  $N_b = 100$  chips, respectively. We then choose the smallest DICI radius  $D_h$  for which the variance of RICICI as well as ISI (observed at each subcarrier output) is below  $-\rho_{\max} = -85$  dB, i.e.,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{rici}(i)}})/N < -\rho_{\max}$  and  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}(i)}})/N < -\rho_{\max}$ . We find the choice of DICI radius  $D_h = 4$  leads to a RICICI variance of  $-89.1$  dB and an ISI variance of  $-94.7$  dB. The resulting pulses are shown in Fig. 4.2.

We choose  $Q_c = 2$  and  $Q_p = Q_d = 1$  for our frames. By design,  $i_o = 0 \in \mathcal{I}_p$ . Then  $i_o = 1 \in \mathcal{I}_d$ . We make these choices for simplicity. Though large frames with more complex structure could lead to lower channel estimation errors, such designs would incur higher estimation complexity as well.

We choose a suitable pilot pattern next. Again, for simplicity, we transmit the same pilot pattern on each pilot multicarrier symbol. Motivated by the optimal pilot patterns for CE-BEM DS channels in [48], we choose our pilot pattern as a sequence

of Kronecker impulses spaced  $N_p$  samples apart, i.e.,

$$s_k(i_o) = \sqrt{\frac{N}{N_p}} \delta_{\langle k \rangle_{N_p}} \quad i_o \in \mathcal{I}_p. \quad (4.28)$$

Fig. 4.3 shows the variance of the effect of the estimation error observed at each subcarrier output,  $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ , for impulse spacings  $N_p = 4$ ,  $N_p = 8$ , and  $N_p = 16$ , respectively. Clearly,  $N_p = 8$  is the best choice. This agrees with [48] where the optimal impulse spacing  $N_p$  is matched to the DICI spread  $(2D_h + 1)$ .

We now examine the variance of the interference at each subcarrier output given by  $\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)})/N$ , and its constituents: the variance of the additive noise, RICI, ISI and the effect of estimation error, i.e.,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}(i)})/N$ ,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{rici}}(i)})/N$ ,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}}(i)})/N$ , and  $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ , respectively. Fig. 4.4 shows these interference components at various SNRs. Clearly, the effect of the channel estimation error,  $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ , is the dominant interference component up to SNR=120 dB. At higher SNRs,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}}(i)})/N$  and  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{rici}}(i)})/N$  dominate. Note that, for this pilot pattern, the threshold for the estimation error to decay inversely with SNR is about 35 dB and that  $\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)})/N$  decays inversely with SNR between SNR=35 dB and SNR=115 dB. Thus, Regime 1 extends to SNR = 115 dB. On the other hand, the ISI covariance  $\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}}(i)}$  is highly rank deficient for the MCM system used here. Consequently, the channel estimation error does not floor at high SNR. Thus, the boundary between the SNR regimes is not affected by the influence of ISI on channel estimation error. The boundary between Regime 1 and Regime 2 occurs around SNR= 115 dB. As the SNR grows beyond 120 dB,  $\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)})$  floors off, heralding Regime 2. A wireless system designed with the chosen parameters would almost certainly never operate in Regime 2.

We study the lower-bounds on LSP-constrained achievable-rate on this system next. Recall that Theorem 3 showed that LSP achieves the maximum spectral efficiency in Regime 1 when  $D_r \geq D_h$  and  $N_r > 1$ . Thus, for our tests, we set  $D_r = D_h$  to verify our results. As a benchmark for LSP, we consider the GSP scheme that uses the entire observations  $\mathbf{r}(i)$  in generating all symbol estimates. We simulate the surrogate MCM system. Fig. 4.5 shows the lower-bound on the LSP-constrained achievable-rate for the surrogate MCM system with (a)  $N_r = 1$ , (b)  $N_r = 2$  and (c)  $N_r = 3$  receive antennas. The traces for LSP and GSP on a surrogate system are labeled **LSP** and **GSP**, respectively. In order to measure spectral efficiency, we measure the slope of a least squares line fit on the achievable rate traces between the SNRs of 50 dB and 90dB. This analysis tells us that when  $N_r = 1$ , the measured spectral efficiency of LSP is about 88% of that for GSP. However, when  $N_r = 2$  the spectral efficiency of LSP is more than 98% of that for GSP. When the number of receive antennas is increased further to  $N_r = 3$ , the spectral efficiency of LSP is unchanged. However, the additional antenna does provide a 2 dB SNR gain. The trends demonstrate that LSP obtains the maximum spectral efficiency of  $\frac{Q_d N}{Q_c N_s}$  for  $D_h > 0$  only when multiple receive antennas are used.

## 4.6 Rate-Complexity Trade-Off

We characterize the gap between lower-bounds on achievable rates of LSP and GSP as a function of the processing radius  $D_r$  for a multi-antenna receiver. This provides a means of choosing a suitable  $D_r$  by characterizing a trade-off between IC-LC-LSP's achievable-rate and computational-complexity at a given SNR. We refer to this as the rate-complexity trade-off (RCT). The achievable-rate for IC-LC-LSP is

$R_{D_r}$ . Meanwhile, IC-LC-LSP's complexity is  $\mathcal{O}((2D_r+1)^3 N_r^3 N)$  per data multicarrier symbol, which is cubic in  $D_r$ . Thus the RCT for a MCM system at a given SNR can be described by the pair  $(R_{D_r}, D_r)$ . The RCTs for the MCM system with  $N_r = 2$  at SNRs of 40 dB and 70 dB are plotted in Fig. 4.6. The RCT trends are similar at both SNRs. When  $D_r \leq D_h$ , the achievable-rate  $R_{D_r}$  grows quickly with increasing  $D_r$ . However, when  $D_r > D_h$ ,  $R_{D_r}$  grows much more slowly with  $D_r$ . In conclusion, there exists a “sweet spot” on the RCT corresponding to  $D_r = D_h$ , and this choice of processing radius provides a good trade-off for LSP-constrained reception.

#### 4.7 Comparing MCM Schemes under A Complexity Constraint

We compare the lower-bound on LSP-constrained achievable-rate of three MCM schemes. By choosing the same processing radius, we ensure that all three schemes have identical receiver processing complexity. In Fig. 4.7, MSINR refers to the jointly optimized max-SINR pulses from [23, 53], while GP refers to the MCM scheme with Gaussian prototype pulses that are dilated to minimize out-of-target ICI/ISI, and OFDM refers to standard CP-OFDM [72] described in Section 2.4. For this experiment, transmission is over channels with  $N_h = 16$  chip delay spreads, uniform power profiles, and  $F_{\text{Dop}} T_s = 0.008$ . The DICI radius and processing radius are chosen to be unity, i.e.,  $D_h = D_r = 1$ . Each data point is an average of measurements over  $10^3$  channel realizations. Additionally, to make the results of this experiment independent of the pilot pattern chosen, we assume perfect CSI at the receiver by setting the channel estimation error to zero, i.e.,  $\mathbf{\Sigma}_{\mathbf{w}_{\text{cee}}^{(n_r)}(i)} = \mathbf{0}$  for all  $n_r \in \{1, \dots, N_r\}$ .

In Fig. 4.7, it is clear that the MSINR scheme outperforms CP-OFDM, which, in turn, performs better than the Gaussian prototype pulsed (GPP) scheme. Recall

that performance is limited by RICI and ISI in Regime 2. The results in Fig. 4.7 can directly be related to RICI/ISI suppression capabilities of these MCM schemes. For instance, the MSINR pulses provide the best RICI/ISI suppression, and consequently, support the highest rate of the three. Note that the RICI and ISI suppression ability of GPP depends on the spread of the modulation and demodulation pulses used. The GPP scheme used in Fig. 4.7 suppresses RICI better than CP-OFDM. However, CP-OFDM suppresses ISI completely at the expense of spectral efficiency, whereas, the performance of GPP is hampered by ISI. For our setup, the loss in rate due to the guards in CP-OFDM is smaller than the rate loss due to uncanceled ISI in the GPP scheme. Thus for our setup, CP-OFDM performs better than GPP MCM.

In a nutshell, the schemes compared here are designed according to different philosophies and have different spectral efficiencies. Yet, the lower-bound on LSP-constrained achievable-rate provides a *fair* means of comparing them under the practical assumption of LSP-constrained reception.

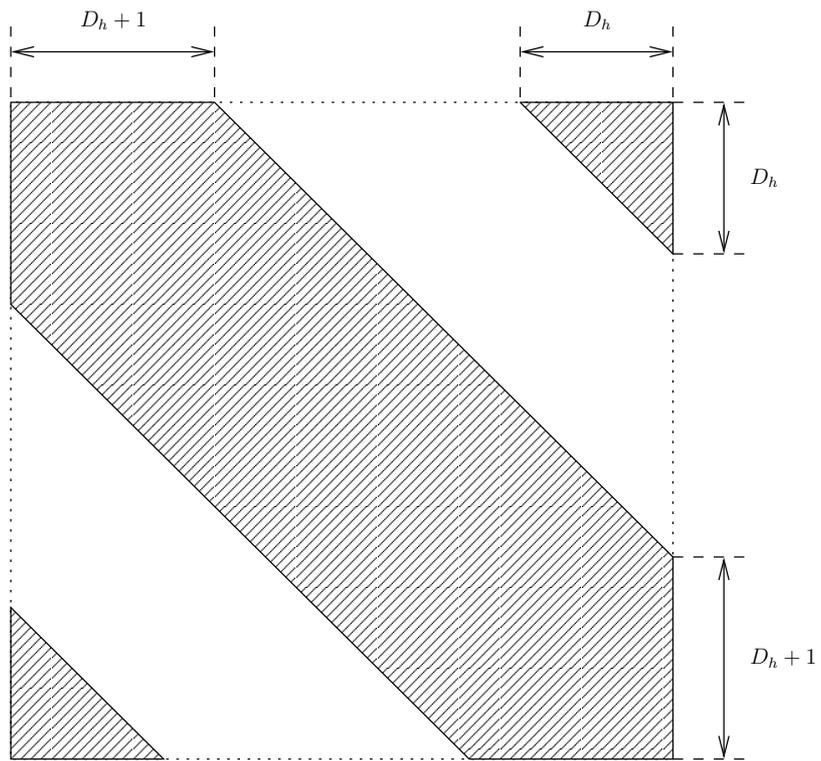


Figure 4.1: Quasi-banded channel matrix.

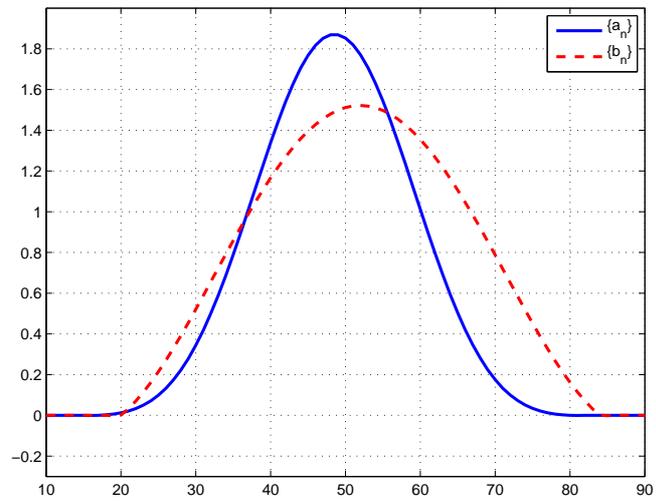


Figure 4.2: The modulation pulse  $\{a_n\}$  and demodulation pulse  $\{b_n\}$  used.

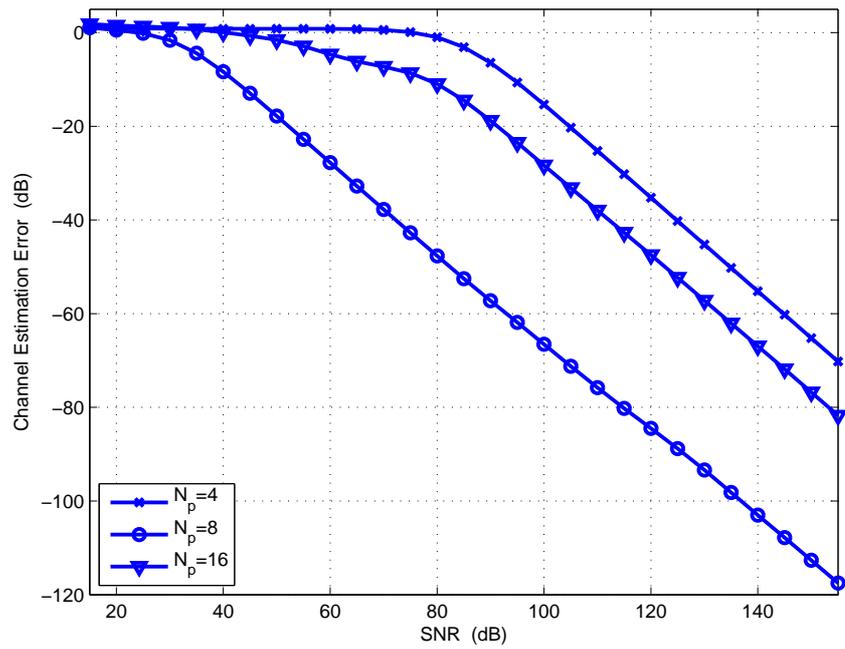


Figure 4.3: The effect of channel estimation error  $\text{tr}(\Sigma_{\text{cee}})/N$  on the observations for pilot spacing  $N_p = 4$ ,  $N_p = 8$  and  $N_p = 16$ .

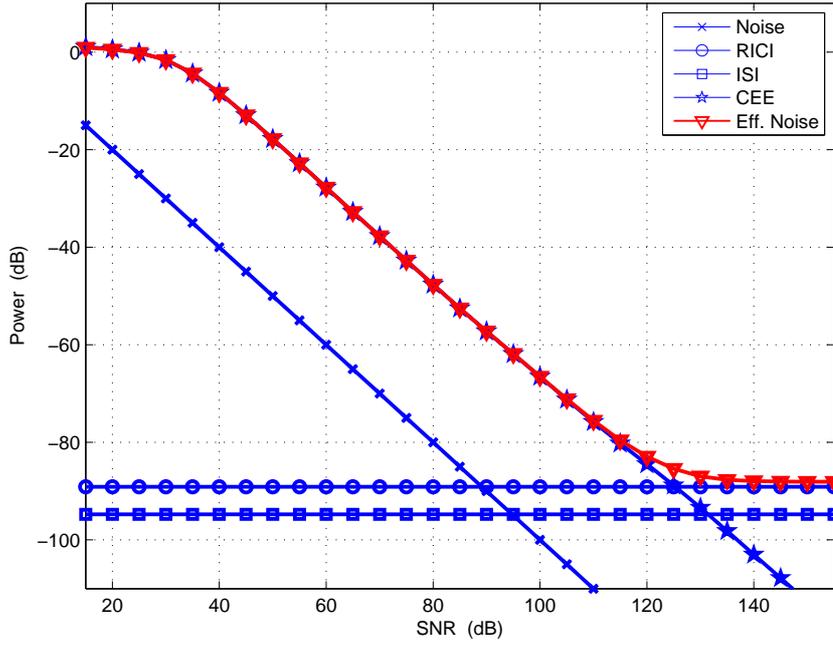


Figure 4.4: The interference variance  $\text{tr}(\mathbf{\Sigma}_{\mathbf{v}(i)})/N$ , labeled Eff. Noise, and its components: the variance of the additive noise  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}(i)})/N$ , labeled Noise; the residual ICI  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{ricl}}(i)})/N$ , labeled RICl; the ISI  $\text{tr}(\mathbf{\Sigma}_{\mathbf{w}_{\text{isi}}(i)})/N$ , labeled ISI; and the effect of channel estimation error  $\text{tr}(\mathbf{\Sigma}_{\text{cee}})/N$ , labeled CEE.

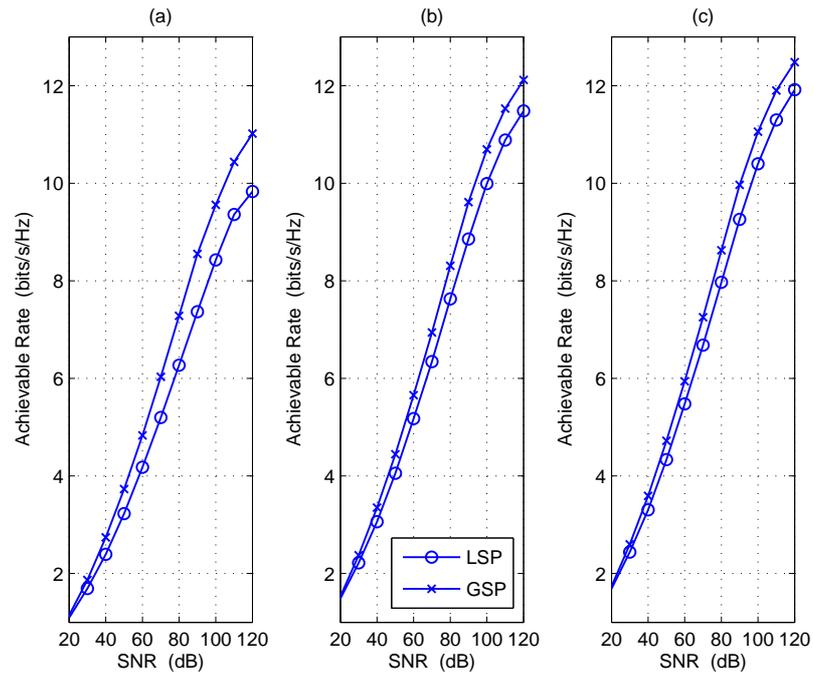


Figure 4.5: Performance of LSP and GSP on the surrogate MCM system for (a)  $N_r = 1$ , (b)  $N_r = 2$  and (c)  $N_r = 3$ .

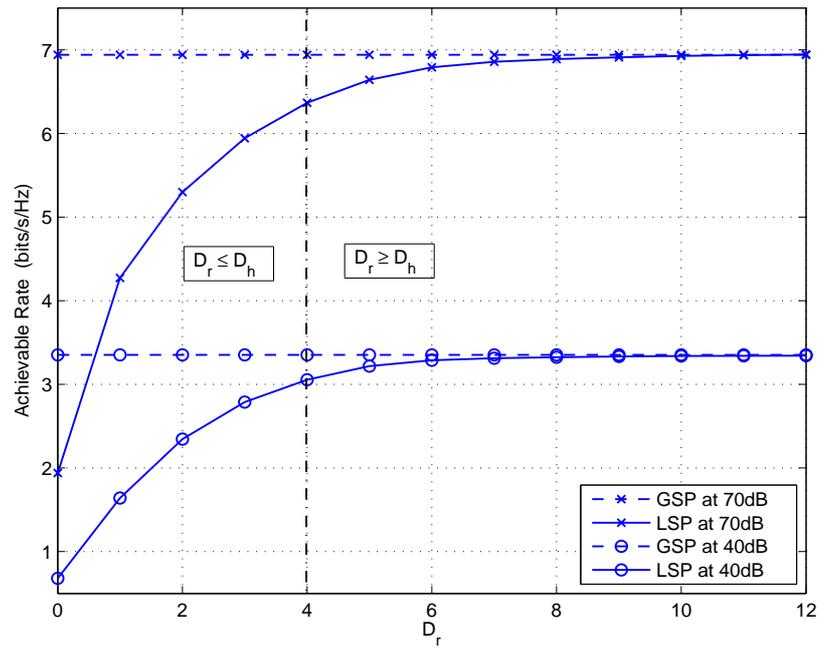


Figure 4.6: The rate-complexity trade-offs with  $N_r = 2$  at two SNRs.

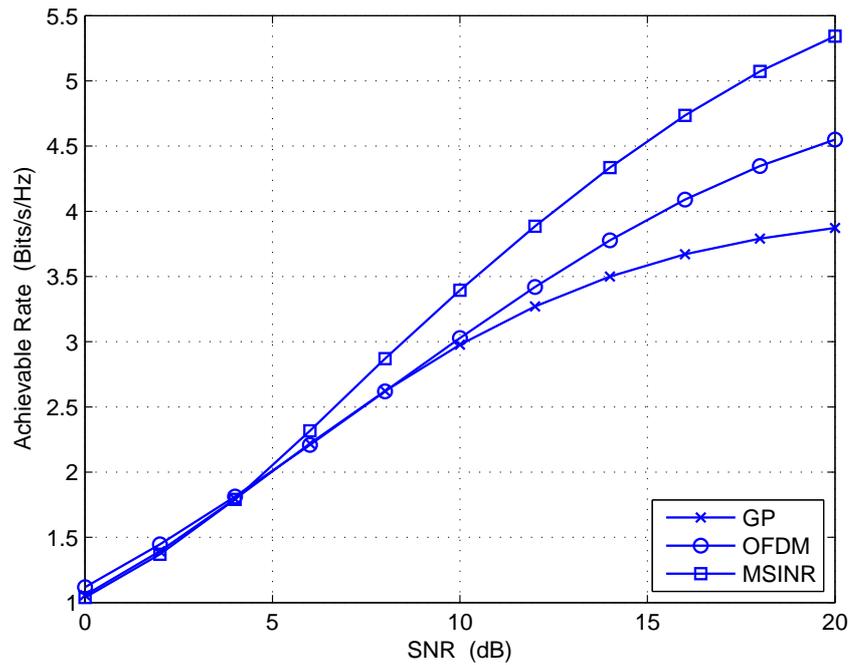


Figure 4.7: Lower-bounds on LSP-constrained achievable rate of CP-OFDM, jointly optimized max-SINR MCM, and Gaussian pulsed MCM scheme for a SIMO-MCM system with  $N_r = 2$  receive antennas and  $N = 128$  subcarriers for  $D_h = D_h = 1$ .

## CHAPTER 5

### BEAMFORMING FOR MIMO-OFDM

An extension of the LSP-constrained achievable rate idea to beamforming MIMO systems is considered here. Starting at the MIMO-MCM system model from Section 2.5.2, the MIMO-OFDM system model is obtained by making certain parameter choices. We then extend the definition of the LSP-constraint to a beamforming MIMO system, derive an expression for the lower-bound on LSP-constrained achievable rate, and use it design beamforming and combining vectors.

#### 5.1 LSP on Beamforming MIMO-OFDM System

From Chapter 2, recall (2.45) where the observation  $\mathbf{r}(i) \in \mathbb{C}^{NN_r}$  can be expressed in terms of the MIMO subcarrier coupling matrices  $\{\mathbf{H}(i, q) \in \mathbb{C}^{NN_r \times NN_t} : i \in \mathbb{Z}\}_{q=-L_{\text{pre}}}^{L_{\text{pst}}}$ , the beamforming vectors  $\{\mathbf{c}_k(i) \in \mathbb{C}^{N_t} : i \in \mathbb{Z}\}_{k=0}^{N-1}$ , and the MCM symbols  $\{\mathbf{s}(i) : i \in \mathbb{Z}\}$  by

$$\mathbf{r}(i) = \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \underline{\mathbf{s}}^{\text{ma}}(i - q) + \mathbf{w}(i) \quad (5.1)$$

$$= \sum_{q=-L_{\text{pre}}}^{L_{\text{pst}}} \mathbf{H}(i, q) \mathcal{D}_b \left( \mathbf{c}_0(i - q), \dots, \mathbf{c}_{N-1}(i - q) \right) \mathbf{s}(i - q) + \mathbf{w}(i). \quad (5.2)$$

For the discussion on beamforming, we make the following assumptions: We consider a MIMO-OFDM system, i.e., a MIMO-MCM system using the CP-OFDM modulation and demodulation technique described in Section 2.4. Moreover, recall from Section 2.4 that ISI is absent, i.e.,  $\mathbf{H}(i, q) = \mathbf{0}$ ,  $\forall q \neq 0$ . Consequently, (5.2) can be simplified to

$$\mathbf{r}(i) = \mathbf{H}(i, 0)\mathcal{D}_b(\mathbf{c}_0(i), \dots, \mathbf{c}_{N-1}(i))\mathbf{s}(i) + \mathbf{w}(i) \quad (5.3)$$

for a MIMO-OFDM system. Furthermore, recall from Section 2.4 that the MIMO-OFDM generates white frequency-domain noise, i.e.,  $\Sigma_{\mathbf{w}(i)} = \sigma^2\mathbf{I}_{NN_r}$ . Additionally, we assume that perfect channel estimates are available at the transmitter and receiver, so that the system sees no channel estimation errors, i.e.,  $\Sigma_{\mathbf{w}_{\text{cee}}(i)} = \mathbf{0}$ . With these assumptions, the surrogate model for the MIMO-OFDM system can be expressed as

$$\mathbf{r}(i) = \hat{\mathbf{H}}(i)\mathcal{D}_b(\mathbf{c}_0(i), \dots, \mathbf{c}_{N-1}(i))\mathbf{s}(i) + \mathbf{v}(i), \quad (5.4)$$

where,  $\hat{\mathbf{H}}(i) \in \mathbb{C}^{NN_r \times NN_t}$  is the  $D_h$ -quasi-block banded matrix with DICI coefficients, and our assumptions imply that the interference vector  $\mathbf{v}(i)$  can be expressed as

$$\mathbf{v}(i) = \mathbf{w}_{\text{rici}}(i) + \mathbf{w}(i). \quad (5.5)$$

Consequently,

$$\Sigma_{\mathbf{v}(i)} = \Sigma_{\mathbf{w}_{\text{rici}}(i)} + \Sigma_{\mathbf{w}(i)}. \quad (5.6)$$

Recall from Section 4.6 that matching the processing radius with the DICI radius provides a lucrative compromise between LSP performance and complexity. Thus for this chapter, we restrict our discussion of LSP to this choice,  $D_r = D_h$ . As a result,  $\mathbf{r}_k(i) \in \mathbb{C}^{(2D_h+1)N_r}$ , constructed from the single antenna components

$\mathbf{r}_k^{(n_r)}(i) := [r_{k-D_h}^{(n_r)}(i), \dots, r_{k+D_h}^{(n_r)}(i)]^T$  for  $n_r \in \{1, \dots, N_r\}$ , is used to estimate scalar data symbols on the  $k^{\text{th}}$  subcarrier. Furthermore, from the definition of  $\mathcal{K}_k^{\text{d}}$  in (3.35),  $\mathbf{r}_k(i)$  experiences DICl from subcarriers with indices in the set

$$\mathcal{K}_k^{\text{d}} = \{k \pm 1, \dots, k \pm 2D_h\} \quad (5.7)$$

only. Then  $\mathbf{r}_k(i)$  can be written in terms of the information symbols as

$$\mathbf{r}_k(i) = \sum_{k' \in \mathcal{K}_k^{\text{d}} \cup \{k\}} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) s_{k'}(i) + \mathbf{v}_k(i), \quad (5.8)$$

where  $\hat{\mathbf{H}}_{k,k'}(i) \in \mathbb{C}^{(2D_h+1)N_r \times N_t}$  represents the influence of  $\mathbf{s}_{k'}^{\text{ma}}(i) = \mathbf{c}_{k'}(i) s_{k'}(i)$  on  $\mathbf{r}_k(i)$ , as depicted in Fig. 5.1, and where  $\mathbf{v}_k(i)$  is comprised of interference samples that affect  $\mathbf{r}_k(i)$  and has covariance  $\Sigma_{\mathbf{v}_k(i)}$ . In the LSP-constrained reception strategy considered, linear combining (LC) is performed to generate estimates of scalar symbols on the first subcarrier ( $k = 0$ ). The obtained estimates are fed to the decoder(s). Assuming judicious rate allocation and consequent error-free decoding, the DICl due to the first subcarrier can be regenerated and removed from observations for neighboring subcarriers. This process is called interference cancellation (IC). These steps are then repeated for the second ( $k = 1$ ) subcarrier, and so on. The effect of IC on  $\mathbf{r}_k(i)$  can be represented as

$$\mathbf{y}_k(i) = \mathbf{r}_k(i) - \sum_{k' \in \mathcal{K}_k^-} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) s_{k'}(i) \quad (5.9)$$

$$= \sum_{k' \in \mathcal{K}_k^+ \cup \{k\}} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) s_{k'}(i) + \mathbf{v}_k(i). \quad (5.10)$$

In (5.9) and (5.10), the set of subcarrier indices  $\mathcal{K}_k^-$  and  $\mathcal{K}_k^+$  are defined as  $\mathcal{K}_k^- = \mathcal{K}_k^{\text{d}} \cap \{l : l < k\}$  and  $\mathcal{K}_k^+ = \mathcal{K}_k^{\text{d}} \cap \{l : l > k\}$ , respectively. Linear combining for the  $k^{\text{th}}$  subcarrier with combiner  $\mathbf{z}_k(i)$  can be written as

$$\hat{s}_k(i) = \mathbf{z}_k(i)^H \mathbf{y}_k(i). \quad (5.11)$$

This chapter presents novel approaches for the design of BVs  $\{\mathbf{c}_k(i)\}_{k=0}^{N-1}$  and combiners  $\{\mathbf{z}_k(i)\}_{k=0}^{N-1}$  using the simplified system model in (5.8) with the aim of (approximately) maximizing a lower-bound on the (ergodic) LSP-constrained achievable-rate (AR) for the system. As a result of using (5.8), the BV and combiner designs as well as all receiver processing use only a few neighboring DIC1 coefficients. (See Fig. 5.1.) Using details in Section 4.3, the lower-bound on LSP-constrained achievable-rate for the surrogate MIMO-OFDM system can be expressed as

$$R^{\text{lb}} = \sum_{k=0}^{N-1} \frac{1}{N + N_h - 1} \mathbb{E}_{\mathbf{H}(i,0)} \left[ \log(1 + \gamma_k(i)) \right], \quad (5.12)$$

where the signal to interference-plus-noise ratio (SINR) for the  $k^{\text{th}}$  subcarrier (after LC) is given by

$$\gamma_k(i) = \frac{\mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i) \mathbf{c}_k(i)^H \hat{\mathbf{H}}_{k,k}(i)^H \mathbf{z}_k(i)}{\mathbf{z}_k(i)^H \left( \sum_{k' \in \mathcal{K}_k^+} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) \mathbf{c}_{k'}(i)^H \hat{\mathbf{H}}_{k,k'}(i)^H + \Sigma_{\mathbf{v}_k(i)} \right) \mathbf{z}_k(i)}. \quad (5.13)$$

In the next section, we present the combiner and BV designs aimed at maximizing  $R^{\text{lb}}$ .

## 5.2 Combiner and Beamforming Vector Designs

A combiner design for a given set of BVs is presented in Section 5.2.1, a BV design is presented in Section 5.2.2, and a joint combiner/BV design is presented in Section 5.2.3.

### 5.2.1 Combiner Design

Here, we design combiners  $\{\mathbf{z}_k(i)\}_{k=0}^{N-1}$  to maximize  $R^{\text{lb}}$  given a set of BVs  $\{\mathbf{c}_k(i)\}_{k=0}^{N-1}$ . In this regard, notice that combiner  $\mathbf{z}_k(i)$  does not affect  $\{\gamma_l(i)\}_{l \neq k}$ . Then maximizing

$R^{\text{lb}}$  w.r.t. combiner  $\mathbf{z}_k(i)$  reduces to maximizing  $\gamma_k(i)$ . It is well known that  $\gamma_k(i)$  in (5.13) is maximized by the choice

$$\mathbf{z}_k(i) = \alpha_k \left( \sum_{k' \in \mathcal{K}_k^+} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) \mathbf{c}_{k'}(i)^H \hat{\mathbf{H}}_{k,k'}(i)^H + \boldsymbol{\Sigma}_{\mathbf{v}_k(i)} \right)^{-1} \times \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i), \quad (5.14)$$

where *w.l.o.g.*, we choose  $\alpha_k$  to ensure  $\|\mathbf{z}_k(i)\| = 1$ . We use (5.14) to design all combiners in this chapter. Note that this choice is often referred to as the “max-SINR” combiner in literature (e.g., in [76]).

### 5.2.2 Max-SNR Beamforming for DS Channels

Traditional BV designs for MIMO-OFDM over time-invariant channels maximize subcarrier SNRs. A similar max-SNR BV design is possible for MIMO-OFDM over DS channels, too. However, the DS channel spreads the energy of each subcarrier into neighboring subcarriers. Taking this spreading into account, the max-SNR BV can be written as the principal eigenvector of the matrix  $\hat{\mathbf{H}}_{k,k}^H(i) \hat{\mathbf{H}}_{k,k}(i)$ , i.e.

$$\mathbf{c}_k(i) = \mathbf{u}_* \left( \hat{\mathbf{H}}_{k,k}(i)^H \hat{\mathbf{H}}_{k,k}(i) \right). \quad (5.15)$$

The solution in (5.15) is called the max-SNR-DS BV design. The designed BV  $\mathbf{c}_k(i)$  maximizes the energy from  $s_k(i)$  in  $\hat{s}_k(i)$ . However, in doing so, the max-SNR-DS BVs potentially increase the ICI caused to neighboring subcarriers (i.e., energy from  $s_k(i)$  in  $\{\hat{s}_l(i)\}_{l \neq k}$ ). Therefore, performance can be improved if DICI suppression can be incorporated into the BV design process. In Section 5.2.3, we propose one such solution.

### 5.2.3 Approximate Max-AR Beamforming and Combining

In this section, BVs and combiners are jointly designed to (approximately) maximize  $R^{\text{lb}}$ . First, we consider the design of BVs given a set of combiners  $\{\mathbf{z}_k(i)\}_{k=0}^{N-1}$ . Realize that each BV  $\mathbf{c}_k(i)$  affects several  $\{\gamma_l(i)\}_{l \neq k}$  and directly computing  $\mathbf{c}_k(i)$  to maximize  $R^{\text{lb}}$  in (5.12) is difficult. Instead, we intuit properties of  $R^{\text{lb}}$ -optimal BVs and use these to construct an alternative cost function that is optimized to calculate the BVs. In this regard, notice that  $\hat{s}_k(i)$  in (5.11) has a “signal” component  $\varphi_k^s(i)$  and a “DICI plus noise” component  $\varphi_k^i(i)$  given by

$$\varphi_k^s(i) = \mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i) \mathbf{s}_k(i), \quad (5.16)$$

$$\varphi_k^i(i) = \mathbf{z}_k(i)^H \left( \sum_{k' \in \mathcal{K}_k^+} \hat{\mathbf{H}}_{k,k'}(i) \mathbf{c}_{k'}(i) s_{k'}(i) + \mathbf{v}_k(i) \right). \quad (5.17)$$

Then  $\gamma_k(i) = \text{E}(|\varphi_k^s(i)|^2) / \text{E}(|\varphi_k^i(i)|^2)$ , where the expectations are taken over the joint distribution of  $\mathbf{s}(i)$  and  $\mathbf{v}_k(i)$ . Observe that BV  $\mathbf{c}_k(i)$  appears in  $\varphi_k^s(i)$  and in  $\{\varphi_l^i(i)\}_{l \in \mathcal{K}_k^-}$ . In the low SNR regime, additive noise dominates DICI and RICl. Therefore, ignoring the DICI terms and setting  $\boldsymbol{\Sigma}_{\mathbf{v}_k(i)} \approx \sigma^2 \mathbf{I}$  at low-SNR, (5.13) can be approximated by

$$\gamma_k(i) \approx \frac{\mathbf{c}_k(i)^H \hat{\mathbf{H}}_{k,k}(i)^H \mathbf{z}_k(i) \mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i)}{\sigma^2 \|\mathbf{z}_k(i)\|^2} \quad (5.18)$$

$$= \frac{\mathbf{c}_k(i)^H \hat{\mathbf{H}}_{k,k}(i)^H \mathbf{z}_k(i) \mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i)}{\sigma^2 \|\mathbf{c}_k(i)\|^2}. \quad (5.19)$$

Here (5.18) yields (5.19) since we assumed  $\|\mathbf{z}_k(i)\| = \|\mathbf{c}_k(i)\| = 1$ . Then  $\mathbf{c}_k(i)$  only affects  $\gamma_k(i)$ . Thus, an AR-optimal BV  $\mathbf{c}_k(i)$  should maximize  $\text{E}(|\varphi_k^s(i)|^2)$ . On the other hand, uncanceled DICI outweighs noise at high SNR. In this case, an AR-optimal BV  $\mathbf{c}_k(i)$  should maximize  $\text{E}(|\varphi_k^s(i)|^2)$  and minimize each element of  $\{\text{E}(|\varphi_l^i(i)|^2)\}_{l \in \mathcal{K}_k^-}$  simultaneously.

These intuitions suggest that a “good” BV should maximize  $\mathcal{E}_k^s(i)$  and minimize  $\mathcal{E}_k^i(i)$  simultaneously, where

$$\mathcal{E}_k^s(i) = \mathbb{E} \left( |\mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i) s_k(i)|^2 \right), \quad (5.20)$$

$$\mathcal{E}_k^i(i) = \mathbb{E} \left( \left| \sum_{k' \in \mathcal{K}_k^-} \mathbf{z}_{k'}(i)^H \hat{\mathbf{H}}_{k',k}(i) \mathbf{c}_k(i) s_k(i) + \mathbf{z}_k(i)^H \mathbf{v}_k(i) \right|^2 \right). \quad (5.21)$$

This prompts us to define the cost function  $\gamma_k^{\text{bv}}(i) = \mathcal{E}_k^s(i)/\mathcal{E}_k^i(i)$ . Recalling that  $\|\mathbf{z}_k(i)\| = 1 = \|\mathbf{c}_k(i)\|$ , the cost function can be simplified to

$$\gamma_k^{\text{bv}}(i) = \frac{\mathbf{c}_k(i)^H \hat{\mathbf{H}}_{k,k}(i)^H \mathbf{z}_k(i) \mathbf{z}_k(i)^H \hat{\mathbf{H}}_{k,k}(i) \mathbf{c}_k(i)}{\mathbf{c}_k(i)^H \left( \sum_{k' \in \mathcal{K}_k^-} \hat{\mathbf{H}}_{k',k}(i)^H \mathbf{z}_{k'}(i) \mathbf{z}_{k'}(i)^H \hat{\mathbf{H}}_{k',k}(i) + \Sigma_{\mathbf{v}_k(i)} \right) \mathbf{c}_k(i)}. \quad (5.22)$$

Given a set of combiners  $\{\mathbf{z}_k(i)\}_{k=0}^{N-1}$ , the  $\gamma_k^{\text{bv}}(i)$ -optimal BV is

$$\mathbf{c}_k(i) = \beta_k \left( \sum_{k' \in \mathcal{K}_k^-} \hat{\mathbf{H}}_{k',k}(i)^H \mathbf{z}_{k'}(i) \mathbf{z}_{k'}(i)^H \hat{\mathbf{H}}_{k',k}(i) + \Sigma_{\mathbf{v}_k(i)} \right)^{-1} \hat{\mathbf{H}}_{k,k}(i)^H \mathbf{z}_k(i), \quad (5.23)$$

where  $\beta_k$  can be chosen *w.l.o.g.* to ensure  $\|\mathbf{c}_k(i)\| = 1$ .

Combining the combiner design from Section 5.2.1 and the BV design above, we propose an iterative approximate max-AR (AMAR) algorithm as follows. The algorithm is initialized by choosing a set of BVs  $\{\mathbf{c}_k^{(0)}(i)\}_{k=0}^{N-1}$ . For our experiments, we found that the max-SNR-DS initialization

$$\mathbf{c}_k^{(0)}(i) = \mathbf{u}_* \left( \hat{\mathbf{H}}_{k,k}(i)^H \hat{\mathbf{H}}_{k,k}(i) \right) \quad (5.24)$$

leads to good results. Each iteration  $n_i \in \{1, \dots, N_i\}$  consists of two stages. First, combiners  $\{\mathbf{z}_k^{(n_i)}(i)\}_{k=0}^{N-1}$  are computed using BVs  $\{\mathbf{c}_k^{(n_i-1)}(i)\}_{k=0}^{N-1}$  via (5.14). Next, BVs  $\{\mathbf{c}_k^{(n_i)}(i)\}_{k=0}^{N-1}$  are recalculated using the new combiners  $\{\mathbf{z}_k^{(n_i)}(i)\}_{k=0}^{N-1}$  via (5.23). The system then uses  $\{\mathbf{z}_k^{(n_i)}(i)\}_{k=0}^{N-1}$  and  $\{\mathbf{c}_k^{(N_i)}(i)\}_{k=0}^{N-1}$  as the BVs and combiners, respectively.

It is easy to see that the combiner design complexity is  $\mathcal{O}((2D_h + 1)^3 N_r^3 N N_i)$  per MCM symbol, whereas the BV design complexity is  $\mathcal{O}(N_t^3 N)$  per MCM symbol, for a MIMO-OFDM system. Thus, the complexity of the AMAR algorithm has the same scaling *w.r.t.* the number of subcarriers  $N$ , the number of transmit antennas  $N_t$ , and the number of receive antennas  $N_r$  as traditional max-SNR designs for MIMO-OFDM over time-invariant channels [61–63].

### 5.3 Numerical Results and Discussion

In this section, we present results of numerical experiments that verify the utility of our designs. Specifically, we measure, using (5.12),  $R^b$  of a MIMO-OFDM system that employs our designs. Tests are performed on a MIMO-OFDM system with  $N_r = 2$  receive antennas and  $N = 128$  subcarriers. Transmission is over channels with  $N_h = 16$  chip delay spreads and uniform power-delay profiles, and each data point is an average of measurements for  $10^3$  channel realizations. In all our experiments, we compare our schemes, i.e., the max-SNR-DS BVs from Section 5.2.2 and combiners from Section 5.2.1 (labeled as MSNR-DS) and the joint AMAR BVs and combiners from Section 5.2.3 (labeled as AMAR), to two benchmarks. First, we compare our designs with ICI-ignoring max-SNR beamforming and combining (labeled as MSNR), intended for time-invariant channels, from [61–63]. Second, we also compare our designs with an upper bound (labeled as UB). The upper bound corresponds to performance on a system using max-SNR-DS BVs, where the receiver, aided by a genie, *cancel all ICI perfectly*. Thus, such a receiver harnesses all available Doppler diversity while completely avoiding the ill effects of ICI.

Fig. 5.2 shows a plot of achievable rate versus subcarrier SNR for (a)  $N_t = 4$ , (b)  $N_t = 6$  and (c)  $N_t = 8$  transmit antennas, respectively, and a (chip normalized single-sided) maximum Doppler spread of  $F_{\text{Dop}}T_s = 0.008$ . We choose the DICI spread as  $D_h = \lceil F_{\text{Dop}}T_s N \rceil = 1$ . The results show that both of our designs are significantly superior to ICI-ignoring max-SNR designs. The iterative AMAR design provides additional rate gains over the max-SNR-DS design at the expense of slightly higher design complexity. Further, our designs perform close to the UB at low and moderate SNRs. On the other hand, neglecting strong ICI components creates a significantly lower performance ceiling for the ICI ignoring max-SNR schemes. When the SNR is high, the gap between the UB and the performance of both our proposed schemes grows due to RICI. However, RICI only produces pronounced performance degradation at high SNRs that may be beyond the normal operating SNR range of most practical systems. Finally, we observe that performance can be enhanced by increasing the number of transmit antennas. This is expected as more transmit antennas provide more freedom to the BV designs in choosing directions rich in signal energy and low in interference energy.

Next, we study the effect of Doppler spread (i.e., mobility) on our system. First, realize that when there is no Doppler spread ( $D_h = 0$ ), our designs reduce to a max-SNR design and hence, perform optimally. A plot of achievable rate versus subcarrier SNR can be found in Fig. 5.3 for  $N_t = 4$  transmit antennas and a maximum Doppler spread of (a)  $F_{\text{Dop}}T_s = 0.008$  ( $D_h = 1$ ), and (b)  $F_{\text{Dop}}T_s = 0.016$  ( $D_h = 2$ ), respectively. These could correspond to, for instance, a channel with bandwidth of 1.5 MHz, carrier frequency of 60 GHz, delay spread of 10.8  $\mu\text{s}$ , and mobile and reflector velocities of (a) 69 km/h and (b) 138 km/h, respectively, in a “triple Doppler” scenario [75].

The assumption of transmit CSI may be unrealistic for a rapidly varying channel. However, approximate transmit CSI can be attained in systems operating in a time division duplex (TDD) mode via prediction from channel measurements made during the previous TDD epoch. To test this idea, we assume that, when in reception mode, the node has near-perfect CSI (via, e.g., pilot aided or decision directed estimation). The node then predicts the channel for the next OFDM symbol duration, when it operates as a transmitter. A MMSE channel predictor that exploits the correlation structure arising from the the WSSUS Rayleigh fading is used. In Fig. 5.3, traces labeled *MSNR-DS-P* and *AMAR-P* refer to versions of the max-SNR-DS BV design and the joint AMAR design, respectively, that use predicted transmitter CSI. The general trends are similar to that of Fig. 5.2. In addition, we observe that, whereas our schemes adapt well to channels with large Doppler spreads, the max-SNR design loses significantly. This behavior results from the fact that the max-SNR scheme completely neglects ICI. Furthermore, we see that, even in a highly mobile environment with large Doppler spreading, the predicted-CSI case achieves rates only slightly less than the perfect-CSI case. This establishes the robustness of our designs to imperfect transmitter CSI.

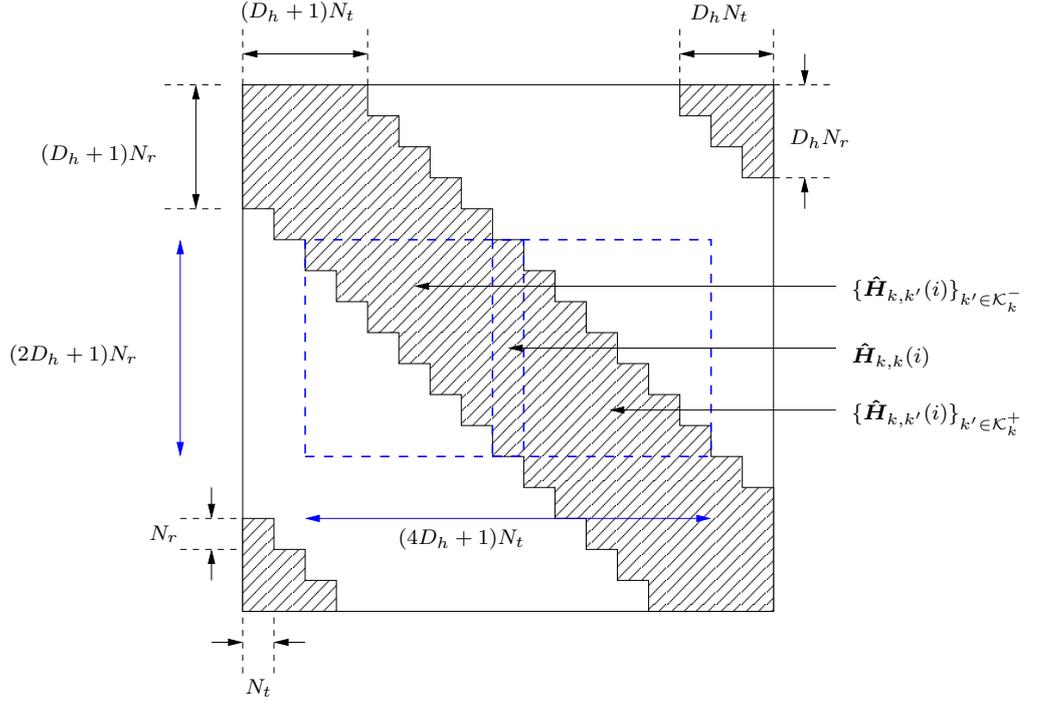


Figure 5.1: Structure of MIMO subcarrier coupling matrix  $\hat{\mathbf{H}}(i)$  for the surrogate MIMO-OFDM system. Rectangle (in dotted lines) indicates the channel coefficients used for LSP and BV design for the  $k^{\text{th}}$  subcarrier.

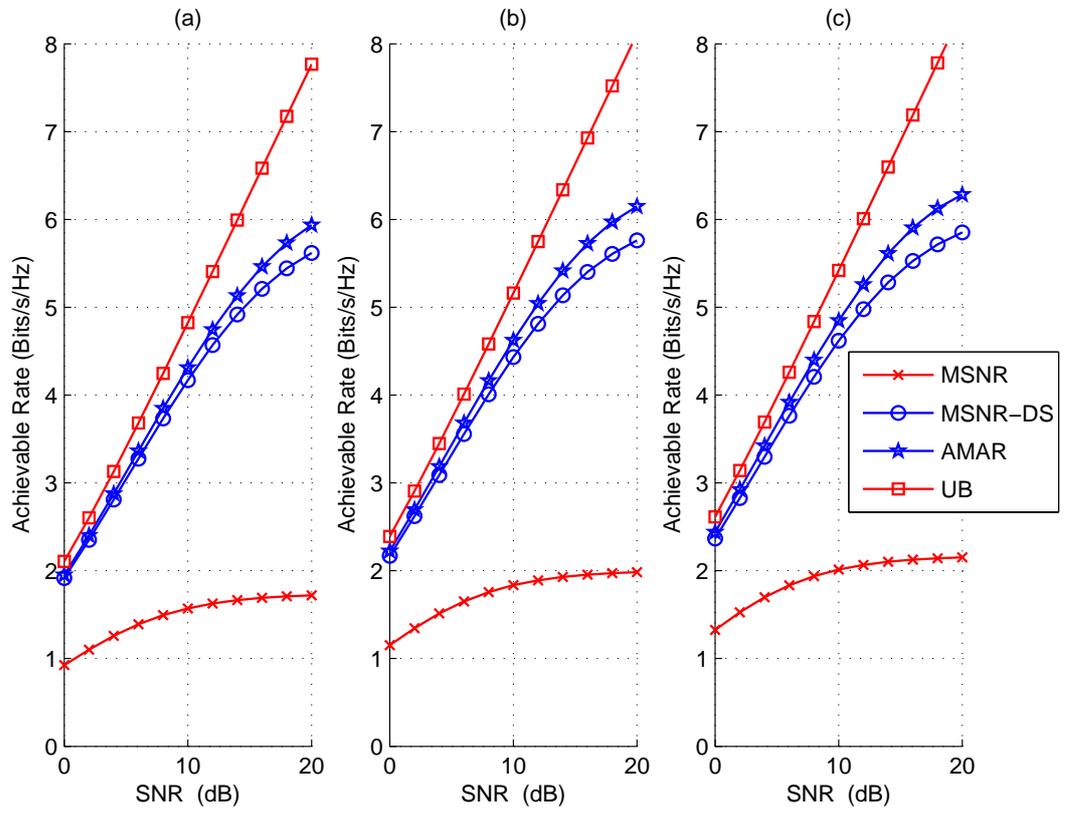


Figure 5.2: Lower-bound  $R^{\text{lb}}$  on LSP-constrained achievable rate versus SNR for (a)  $N_t = 4$ , (b)  $N_t = 6$  and (c)  $N_t = 8$  transmit antennas.

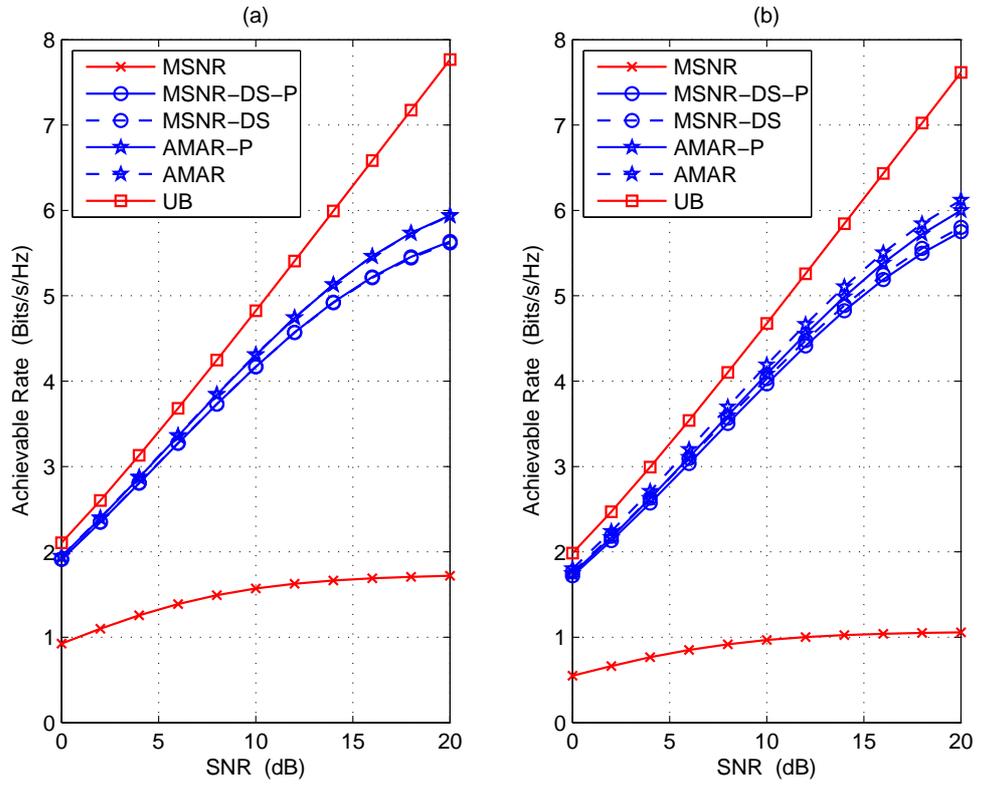


Figure 5.3: Lower-bound  $R^{\text{lb}}$  on LSP-constrained achievable rate on a system in TDD mode with predicted CSI at the transmitter for (a)  $F_{\text{Dop}} T_s = 0.008$ , and (b)  $F_{\text{Dop}} T_s = 0.016$ .

## CHAPTER 6

### INCORPORATING CHANNEL RE-ESTIMATION

Until now, we have considered the channel estimation strategy described in Section 3.3. In this penultimate chapter, channel re-estimation is incorporated into the reception strategy for a communication system operating on DS channels. The results here are developed for generic communication schemes operating over the DS channel. We formulate an achievable-rate lower-bound for reception incorporating channel re-estimation, and use it to study high-SNR spectral efficiency. A pilot-data power allocation strategy for pilot-aided transmissions over DS channels is also discussed.

#### 6.1 System Model

Because this chapter considers a generic communication system and is not limited to an MCM system, using the system model from Chapter 2 is difficult. Thus, we derive the system model from scratch and use it in this penultimate chapter only.

##### 6.1.1 Transmission Model

Consider a scheme in which information is transmitted through  $N_s$  substreams, each of which uses the channel  $N_b$  times. In particular, say that  $s_k(i)$  denotes the  $i^{\text{th}}$  sample of the  $k^{\text{th}}$  substream. The first  $N_p$  substreams (i.e.,  $\{s_k(i)\}_{i=0}^{N_b-1}$  for  $k =$

$0 \dots N_p - 1$ ) are dedicated to pilots while the remaining  $N_s - N_p$  substreams (i.e.,  $\{s_k(i)\}_{i=0}^{N_b-1}$  for  $k = N_p \dots N_s - 1$ ) are dedicated to data. The data substreams are independently encoded at rates that ensure reliable decoding, as will be discussed later. For this, we assume that the transmitter knows the channel statistics, but not the channel state.

The average transmission power is constrained to  $E_{\text{tot}}$  joules per-channel-use,  $E_p$  of which is allocated to pilots and the remainder of which is divided equally among the data substreams. Thus, each data substream has power

$$\sigma_s^2 = \frac{E_{\text{tot}} - E_p}{N_s - N_p}. \quad (6.1)$$

For analytical tractability, we assume the use of i.i.d. Gaussian codebooks. With this assumption, the power constraints can be expressed as

$$\sum_{k=0}^{N_p-1} |s_k(i)|^2 = E_p \quad \forall i \quad (6.2)$$

$$\text{E}\{\underline{\mathbf{s}}_{N_p}(i_1)\underline{\mathbf{s}}_{N_p}(i_2)^H\} = \sigma_s^2 \mathbf{I}_{N_s-N_p} \delta_{i_1-i_2} \quad (6.3)$$

where

$$\underline{\mathbf{s}}_k(i) := [s_k(i), \dots, s_{N_s-1}(i)]^T. \quad (6.4)$$

In this chapter, we will also make frequent use of the notation

$$\mathbf{s}_k(i) := [s_0(i), \dots, s_{k-1}(i)]^T \quad (6.5)$$

$$\mathbf{s}(i) := [s_0(i), \dots, s_{N_s-1}(i)]^T = [\mathbf{s}_k(i)^T, \underline{\mathbf{s}}_k(i)^T]^T. \quad (6.6)$$

### 6.1.2 Channel Model

We assume that the receiver observes the noisy inter-substream-interference (ISSI) corrupted samples  $\{y_k(i)\}_{i=0}^{N_b-1}$ , for  $k \in \{0, \dots, N-1\}$ , where

$$y_k(i) = \sum_{l=0}^{N_h-1} h_{k,l}(i) s_{k-l}(i) + w_k(i). \quad (6.7)$$

Here,  $N := N_s + N_h - 1$ , where  $N_h$  denotes the ISSI length, and  $s_k(i) := 0$  for  $k \notin \{0, \dots, N_s - 1\}$ . The observations can be written in vector form as  $\mathbf{y}(i) = [y_0(i), \dots, y_{N-1}(i)]^T$ , where

$$\mathbf{y}(i) = \mathbf{H}(i)\mathbf{s}(i) + \mathbf{w}(i) \quad (6.8)$$

for  $\mathbf{w}(i) = [w_0(i), \dots, w_{N-1}(i)]^T$  and

$$\mathbf{H}(i) = \begin{bmatrix} h_{0,0}(i) & & & & \\ & \vdots & & & \\ & & \ddots & & \\ h_{N_h-1, N_h-1}(i) & & & h_{N_s-1, 0}(i) & \\ & & & \vdots & \\ & & & & h_{N-1, N_h-1}(i) \end{bmatrix}. \quad (6.9)$$

Notice that the channel model suffices to describe either single-carrier or multi-carrier transmission. In the single-carrier case, (6.7) corresponds to  $N_b$  blocks of  $(N_h - 1)$ -zero-padded  $N_s$ -block transmission through a doubly selective fading channel with time-varying inter-symbol interference (ISI) of length  $N_h$ . In the multi-carrier case, (6.7) corresponds to  $N_b$  symbols of an  $N$ -subcarrier system with  $N_s$  active subcarriers and an inter-carrier interference (ICI) response of length  $N_h$ .

It will sometimes be convenient to write the system model as

$$\mathbf{y}(i) = \mathbf{S}(i)\mathbf{h}(i) + \mathbf{w}(i) \quad (6.10)$$

with  $\mathbf{h}(i) \in \mathbb{C}^{N_s N_h}$  such that

$$\mathbf{h}(i) := [h_{0,0}(i), \dots, h_{N_s-1,0}(i), h_{1,1}(i), \dots, h_{N_s,1}(i), \dots, h_{N_h-1,N_h-1}(i), \dots, h_{N-1,N_h-1}(i)]^T, \quad (6.11)$$

and with

$$\mathbf{S}(i) := \begin{bmatrix} s_0(i) & & & 0 & \cdots & 0 & & 0 & \cdots & 0 \\ & \ddots & & s_0(i) & & & \cdots & \vdots & & \vdots \\ & & s_{N_s-1}(i) & & \ddots & & & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & & s_{N_s-1}(i) & & s_0(i) & & \\ \vdots & & \vdots & 0 & \cdots & 0 & \cdots & & \ddots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & & & & s_{N_s-1}(i) \end{bmatrix}. \quad (6.12)$$

In this chapter, we make frequent use of

$$\mathbf{y}_k(i) := [y_0(i), \dots, y_{k-1}(i)]^T, \quad (6.13)$$

$$\mathbf{w}_k(i) := [w_0(i), \dots, w_{k-1}(i)]^T, \quad (6.14)$$

and  $\mathbf{S}_k(i)$  defined as the matrix formed by the first  $k$  rows of  $\mathbf{S}(i)$ . Note that

$$\mathbf{y}_k(i) = \mathbf{S}_k(i)\mathbf{h}(i) + \mathbf{w}_k(i), \quad (6.15)$$

where the entries in  $\mathbf{S}_k(i)$  come from  $\mathbf{s}_k(i)$  but not from  $\underline{\mathbf{s}}_k(i)$ .

We assume that the channel coefficients  $\mathbf{h}(i)$  are zero-mean circular Gaussian with  $\mathbb{E}\{\mathbf{h}(i)\mathbf{h}(i)^H\} = \mathbf{\Sigma}_h$ , where  $\text{rank}(\mathbf{\Sigma}_h) = N_m < N_s$  and where  $\text{tr}(\mathbf{\Sigma}_h) = N_s$  (i.e., the channel is energy-preserving). Similarly, we assume that the noise coefficients are zero-mean circular Gaussian with  $\mathbb{E}\{\mathbf{w}(i)\mathbf{w}(i)^H\} = \sigma^2 \mathbf{I}_N$  and independent across  $i$ .

## 6.2 Non-Coherent Pilot-and-Data-Aided Communication

### 6.2.1 Description of Scheme

We now summarize the non-coherent communication scheme, elaborating on the details after the summary.

0. The MMSE estimate of  $\mathbf{h}(i)$  from  $\mathbf{y}_{N_p}(i)$  is computed for each  $i \in \{0, \dots, N_b - 1\}$ , leveraging the fact that  $\mathbf{y}_{N_p}(i)$  is a function of the pilots  $\mathbf{s}_{N_p}(i)$  but not the unknown data  $\mathbf{x}_{N_p}(i)$ . Denoting this pilot-aided channel estimate by  $\hat{\mathbf{h}}^{(N_p)}(i)$ , the first data substream  $\{s_{N_p}(i)\}_{i=0}^{N_b-1}$  is then coherently decoded using the pilot-aided channel estimate  $\hat{\mathbf{h}}^{(N_p)}(i)$ . With large enough  $N_b$  and suitable choice of code rate, this data substream can be reliably decoded.
1. Using the decoded substream in conjunction with pilots, a refined MMSE channel estimate  $\hat{\mathbf{h}}^{(N_p+1)}(i)$  is computed from  $\mathbf{y}_{N_p+1}(i)$  for each  $i \in \{0, \dots, N_b - 1\}$ , leveraging the fact that  $\mathbf{y}_{N_p+1}(i)$  is not a function of the not-yet-decoded data. The next data substream  $\{s_{N_p+1}(i)\}_{i=0}^{N_b-1}$  is then coherently decoded using the refined channel estimate  $\hat{\mathbf{h}}^{(N_p+1)}(i)$ . With a suitable choice of code rate, this second data substream can also be reliably decoded.
2. Using the two decoded substreams in conjunction with pilots, the refined MMSE channel estimate  $\hat{\mathbf{h}}^{(N_p+2)}(i)$  is computed from  $\mathbf{y}_{N_p+2}(i)$  for each  $i \in \{0, \dots, N_b - 1\}$ . The next data substream  $\{s_{N_p+2}(i)\}_{i=0}^{N_b-1}$  is then coherently decoded using the most recent channel estimate  $\hat{\mathbf{h}}^{(N_p+2)}(i)$ , where decoding can be made reliable via proper rate selection.

\*. The procedure continues this way until all  $N_s$  data substreams have been decoded.

Next we elaborate on the channel estimation and data decoding procedures. Rate allocation will be detailed in Section 6.2.2.

The MMSE channel estimate  $\hat{\mathbf{h}}^{(k)}(i)$ , i.e., the estimate of  $\mathbf{h}(i)$  from  $\mathbf{y}_k(i)$  given perfect knowledge of  $\mathbf{S}_k(i)$ , can be written as [73]

$$\begin{aligned} \hat{\mathbf{h}}^{(k)}(i) &= \boldsymbol{\Sigma}_{\mathbf{h}} \mathbf{S}_k(i)^H \\ &\quad \times (\mathbf{S}_k(i) \boldsymbol{\Sigma}_{\mathbf{h}} \mathbf{S}_k(i)^H + \sigma^2 \mathbf{I}_k)^{-1} \mathbf{y}_k(i). \end{aligned} \quad (6.16)$$

Conditioned on  $\mathbf{s}_k(i)$ , the estimation error  $\tilde{\mathbf{h}}^{(k)}(i) := \mathbf{h}(i) - \hat{\mathbf{h}}^{(k)}(i)$  has covariance [73]

$$\begin{aligned} \boldsymbol{\Sigma}_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} &= \boldsymbol{\Sigma}_{\mathbf{h}} - \boldsymbol{\Sigma}_{\mathbf{h}} \mathbf{S}_k(i)^H \\ &\quad \times (\mathbf{S}_k(i) \boldsymbol{\Sigma}_{\mathbf{h}} \mathbf{S}_k(i)^H + \sigma^2 \mathbf{I}_k)^{-1} \mathbf{S}_k(i) \boldsymbol{\Sigma}_{\mathbf{h}}. \end{aligned} \quad (6.17)$$

Now we describe the decoding of data substream  $\{s_k(i)\}_{i=0}^{N_b-1}$  for  $k \in \{N_p, \dots, N_s - 1\}$ . In doing so, we make use of the partition  $\mathbf{H}(i) = [\mathbf{H}_k(i), \mathbf{h}_k(i), \underline{\mathbf{H}}_{k+1}(i)]$ , where  $\mathbf{H}_k(i) \in \mathbb{C}^{N \times k}$ ,  $\mathbf{h}_k(i) \in \mathbb{C}^{N \times 1}$ , and  $\underline{\mathbf{H}}_{k+1}(i) \in \mathbb{C}^{N \times (N_s - k - 1)}$ , so that (6.8) becomes

$$\begin{aligned} \mathbf{y}(i) &= \mathbf{H}_k(i) \mathbf{s}_k(i) + \mathbf{h}_k(i) s_k(i) \\ &\quad + \underline{\mathbf{H}}_{k+1}(i) \underline{\mathbf{s}}_{k+1}(i) + \mathbf{w}(i). \end{aligned} \quad (6.18)$$

In addition, we construct  $\hat{\mathbf{H}}^{(k)}(i)$  from  $\hat{\mathbf{h}}^{(k)}(i)$ , and  $\tilde{\mathbf{H}}^{(k)}(i)$  from  $\tilde{\mathbf{h}}^{(k)}(i)$ , in the same way that we constructed  $\mathbf{H}^{(k)}(i)$  from  $\mathbf{h}^{(k)}(i)$ , and we make the corresponding partition  $\hat{\mathbf{H}}^{(k)}(i) = [\hat{\mathbf{H}}_k^{(k)}(i), \hat{\mathbf{h}}_k^{(k)}(i), \underline{\hat{\mathbf{H}}}_{k+1}^{(k)}(i)]$ . The first stage of decoding involves interference cancellation and linear combining:

$$\mathbf{r}^{(k)}(i) = \mathbf{y}(i) - \hat{\mathbf{H}}_k^{(k)}(i) \mathbf{s}_k(i) \quad (6.19)$$

$$z_k(i) = \mathbf{c}^{(k)}(i)^H \mathbf{r}^{(k)}(i). \quad (6.20)$$

Recall that we have assumed that, at the time of decoding  $\{s_k(i)\}_{i=0}^{N_b-1}$ , the substreams  $\{\mathbf{s}_k(i)\}_{i=0}^{N_b-1}$  are known through reliable decoding or as pilots. Using (6.18) and (6.19), we see that

$$\mathbf{r}^{(k)}(i) = \hat{\mathbf{h}}_k^{(k)}(i)s_k(i) + \mathbf{v}^{(k)}(i), \quad (6.21)$$

$$\mathbf{v}^{(k)}(i) = \underline{\hat{\mathbf{H}}}_{k+1}^{(k)}(i)\underline{\mathbf{s}}_{k+1}(i) + \tilde{\mathbf{H}}^{(k)}(i)\mathbf{s}(i) + \mathbf{w}(i) \quad (6.22)$$

from which the post-combining SINR  $\gamma^{(k)}(i)$  becomes

$$\gamma^{(k)}(i) = \frac{|\mathbf{c}^{(k)}(i)^H \hat{\mathbf{h}}_k^{(k)}(i)|^2 \sigma_s^2}{\mathbf{c}^{(k)}(i)^H \Sigma_{\mathbf{v}^{(k)}(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)} \mathbf{c}^{(k)}(i)} \quad (6.23)$$

for  $\Sigma_{\mathbf{v}^{(k)}(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)} := \mathbb{E}\{\mathbf{v}^{(k)}(i)\mathbf{v}^{(k)}(i)^H | \mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)\}$ . The combiner output can then be written as

$$z_k(i) = g_k(i)s_k(i) + n_k(i), \quad (6.24)$$

for  $g_k(i) := \mathbf{c}^{(k)}(i)^H \hat{\mathbf{h}}_k^{(k)}(i)$  and  $n_k(i) := \mathbf{c}^{(k)}(i)^H \mathbf{v}^{(k)}(i)$ . After having computed  $\{z_k(i)\}_{i=0}^{N_b-1}$  and  $\{g_k(i)\}_{i=0}^{N_b-1}$ , coherent decoding based on (6.24) can be applied.

## 6.2.2 Achievable-Rate Analysis

Notice that, for  $k \in \{N_p, \dots, N_s - 1\}$ , the effective noise  $n_k(i)$  is non-Gaussian:

$$\begin{aligned} n_k(i) &= \mathbf{c}^{(k)}(i)^H \left( \underline{\hat{\mathbf{H}}}_{k+1}^{(k)}(i)\underline{\mathbf{s}}_{k+1}(i) \right. \\ &\quad \left. + \underbrace{\tilde{\mathbf{H}}^{(k)}(i)\mathbf{s}(i) + \mathbf{w}(i)}_{\text{non-Gaussian}} \right). \end{aligned} \quad (6.25)$$

Medard [40] and Hassibi [42] showed that the worst-case noise distribution with respect to mutual information is the Gaussian one. Additionally, the achievable rate for Gaussian signaling in the presence of Gaussian distributed noise can be expressed in terms of the post-combining SINR [76]. Therefore, assuming adequately large  $N_b$ ,

the achievable rate across data substream  $k \in \{N_p, \dots, N_s - 1\}$  can be bounded from below, in units of nats-per-(scalar)-channel-use, as

$$R_k \geq \mathbb{E} \left\{ \log(1 + \gamma^{(k)}(i)) \right\}, \quad (6.26)$$

where the SINR  $\gamma^{(k)}(i)$  was given in (6.23) and the expectation in (6.26) is taken over the joint distribution of  $\hat{\mathbf{h}}^{(k)}(i)$  and  $\mathbf{s}_k(i)$ . Note that the bound in (6.26) holds for general linear combiners  $\mathbf{c}^{(k)}(i)$ . The tightest bound can be obtained by choosing the max-SINR combiner

$$\mathbf{c}^{(k)}(i) = \boldsymbol{\Sigma}_{\mathbf{v}^{(k)}(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)}^{-1} \hat{\mathbf{h}}_k^{(k)}(i), \quad (6.27)$$

in which case the bound (6.26) becomes

$$R_k \geq \mathbb{E} \left\{ \log \left( 1 + \sigma_s^2 \hat{\mathbf{h}}_k^{(k)}(i)^H \boldsymbol{\Sigma}_{\mathbf{v}^{(k)}(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)}^{-1} \hat{\mathbf{h}}_k^{(k)}(i) \right) \right\} \quad (6.28)$$

leading to the following bound on the overall achievable rate, given in nats-per-channel-use.

$$R_{\text{tot}} \geq \frac{1}{N} \sum_{k=N_p}^{N_s-1} \mathbb{E} \left\{ \log \left( 1 + \sigma_s^2 \hat{\mathbf{h}}_k^{(k)}(i)^H \boldsymbol{\Sigma}_{\mathbf{v}^{(k)}(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)}^{-1} \hat{\mathbf{h}}_k^{(k)}(i) \right) \right\}. \quad (6.29)$$

To facilitate reliable decoding, the data substream rates should be chosen in accordance with (6.28).

### 6.2.3 Asymptotic Achievable-Rate Analysis

In this section, we analyze the achievable rate  $R_{\text{tot}}$  at high SNR. For this, we define the SNR  $\rho := \frac{E_{\text{tot}}}{N\sigma^2}$  and examine  $R_{\text{tot}}(\rho)$  as  $\rho \rightarrow \infty$ .

To provide some intuition on the high-SNR behavior, consider for the moment choosing a zero-forcing (ZF) combiner  $\mathbf{c}^{(k)}(i)$ ,  $\|\mathbf{c}^{(k)}(i)\| = 1$ , such that

$$\mathbf{c}^{(k)}(i)^H \hat{\mathbf{H}}_{k+1}^{(k)}(i) = \mathbf{0}. \quad (6.30)$$

which implies that  $g_k(i) = \mathbf{c}^{(k)}(i)^H \hat{\mathbf{h}}_k^{(k)}(i)$  and (via (6.25))

$$n_k(i) = \mathbf{c}^{(k)}(i)^H (\tilde{\mathbf{H}}^{(k)}(i) \mathbf{s}(i) + \mathbf{w}^{(k)}(i)). \quad (6.31)$$

Note that this ZF combiner exists w.p.1. We reason that, for a “well-designed” channel estimation procedure, the covariance of the channel estimation error  $\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)}$  (defined in (6.17)) should vanish as  $\sigma^2 \rightarrow 0$ . In particular, there should exist a pilot pattern  $\mathbf{s}_{N_p}$  that, for fixed  $E_{\text{tot}}$  and  $N_p < N_s$ , guarantees the existence of  $\sigma$ -invariant  $\mathbf{A}$  such that  $\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \leq \sigma^2 \mathbf{A}$  for all  $\sigma > 0$  and for all  $k \in \{N_p, \dots, N_s - 1\}$ . When this is the case, the use of a zero-forcing combiner ensures that  $\sigma^2_{n_k(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)} := \mathbb{E}\{|n_k(i)|^2 | \mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)\}$  will also vanish as  $\sigma^2 \rightarrow 0$  for each information substream. In particular, there will exist  $\sigma$ -invariant  $\alpha$  such that  $\sigma^2_{n_k(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)} \leq \sigma^2 \alpha$  for all  $\sigma > 0$  and for all  $k \in \{N_p, \dots, N_s - 1\}$ . Then we see that  $\gamma^{(k)}(i) \geq \frac{|g_k(i)|^2 \sigma_s^2}{\alpha \sigma^2} = \frac{|g_k(i)|^2 N}{\alpha(N_s - N_p)} (1 - \frac{E_p}{E_{\text{tot}}}) \rho$  for each  $k \in \{N_p, \dots, N_s - 1\}$ . When this is the case, (6.26) implies

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} \geq \frac{N_s - N_p}{N}. \quad (6.32)$$

We now make these statements more precise.

The condition under which the covariance of the estimation error  $\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)}$  vanishes with increasing SNR is given by the following lemma.

**Lemma 2.** *Let the columns of  $\mathbf{B} \in \mathbb{C}^{N_h N_s \times N_m}$  be the eigenvectors corresponding to non-zero eigenvalues of  $\Sigma_{\mathbf{h}}$ . Then there exists  $\sigma$ -invariant  $\mathbf{A}$  such that, for every  $k \in \{N_p, \dots, N_s - 1\}$ ,*

$$\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \leq \sigma^2 \mathbf{A} \quad \forall \sigma > 0 \quad (6.33)$$

if and only if

$$\text{rank}(\mathbf{S}_{N_p}(i)\mathbf{B}) = N_m. \quad (6.34)$$

*Proof.* See Appendix C.1 □

In the sequel, we refer to condition (6.34) as the “rank condition.” Lemma 2 says that the  $N_p$  pilot substreams must excite all  $N_m$  channel modes in order to obtain channel estimates whose error vanishes with increasing SNR.

**Theorem 6.** *For the class of channels that enable pilots  $\mathbf{S}_{N_m}(i)$  to yield*

$$\text{rank}(\mathbf{S}_{N_m}(i)\mathbf{B}) = N_m, \quad (6.35)$$

*the achievable rate of our scheme obeys*

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} = \frac{N_s - N_m}{N}. \quad (6.36)$$

*Proof.* See Appendix C.2 □

### 6.3 Illustrative Example

We now consider the specific case of a doubly selective fading channel which obeys a complex-exponential basis expansion model (CE-BEM) [7, 9]. In particular, the channel coefficients in  $\mathbf{h}(i)$  are parameterized by  $N_m = (2D + 1)N_h$  uncorrelated Gaussian random variables  $\{\phi_{m,l}(i) : m \in \{-D, \dots, D\}, l \in \{0, \dots, N_h - 1\}\}$  via

$$h_{k,l}(i) = \frac{1}{\sqrt{N_s}} \sum_{m=-D}^D \phi_{m,l}(i) e^{j\frac{2\pi}{N_s}m(k-l)}. \quad (6.37)$$

For simplicity, we assume that the random variables  $\{\phi_{m,l}(i)\}_{m,l}$  have equal variance.

Hence  $\text{E}[\phi_{m_1,l_1}(i_1)\phi_{m_2,l_2}^*(i_2)] = \frac{N_s}{(2D+1)N_h} \delta_{m_1-m_2} \delta_{l_1-l_2} \delta_{i_1-i_2}$ . In (6.37),  $D \approx \lceil F_{\text{Dop}}T_s N_s \rceil$

where  $F_{\text{Dop}}T_s$  is the single-sided normalized Doppler spread. For this CE-BEM channel, the eigenvector matrix  $\mathbf{B}$  (defined in Lemma 2) has the form  $\mathbf{B} = \mathbf{I}_{N_h} \otimes \mathbf{F}$ , where the  $N_s \times (2D + 1)$  matrix  $\mathbf{F}$  is defined element-wise as  $[\mathbf{F}]_{m_1, m_2} = \frac{1}{\sqrt{N_s}} e^{j \frac{2\pi}{N_s} m_1(m_2 - D)}$ .

We first address the issue of choosing a suitable pilot pattern for this example. Estimating the channel coefficients  $\{h_{k,l}(i)\}_{k=l}^{N_s+l-1}$  for each  $l$  is equivalent to estimating the  $(2D + 1)$  random variables  $\{\phi_{m,l}\}_{m=-D}^D$ . This can be accomplished by exciting the channel with a set of  $(2D + 1)$  impulse sequences of length  $N_h - 1$ . This motivates the use of the  $N_p = N_m = (2D + 1)N_h$ -length pilot pattern

$$s_k(i) = \sqrt{\frac{N_h E_p}{N_m}} \delta_{(k)N_h}, \quad 0 \leq k < N_m. \quad (6.38)$$

The proposed scheme, in conjunction with this pilot pattern, leads to the following achievable-rate characterization.

**Proposition 1.** *For the CE-BEM doubly selective fading channel, the achievable rate of our scheme obeys*

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} = \frac{N_s - N_m}{N}. \quad (6.39)$$

*Proof.* See Appendix C.3. □

Interestingly, [47, 71] has shown that, under continuously distributed inputs, the maximum spectral efficiency than can be achieved on the CE-BEM doubly selective block-fading channel is  $\frac{N_s - N_m}{N}$ . Thus, using the pilot pattern (6.38), the proposed scheme becomes “spectrally-efficient”.

Next, we tackle the issue of power allocation between pilot and data substreams. Let  $E_p = \alpha_p E_{\text{tot}}$  for some  $\alpha_p \in (0, 1)$ . Then  $\sigma_s^2 = (1 - \alpha_p)E_{\text{tot}}/(N_s - N_p)$ . We propose a “minimax” approach whereby we choose  $\alpha_p$  to maximize a lower-bound

on the achievable rate of the weakest data substream. Though this power allocation strategy can be used with an arbitrarily chosen pilot pattern, we restrict ourselves to the pilot pattern (6.38) for simplicity. Recall that the channel estimate is refined after decoding each data substream, thereby increasing the effective SINR. Thus, the first data substream  $\{s_{N_p}(i)\}_{i=0}^{N_b-1}$  must be the weakest. Recalling that  $\mathbf{e}_N^{(N_p)}$  denotes the  $N_p^{\text{th}}$  column of  $\mathbf{I}_N$ , we have the following result.

**Proposition 2.** *For the CE-BEM doubly selective fading channel, the pilot power allocation*

$$\alpha_{p,*} = \arg \max_{\alpha_p \in (0,1)} \frac{\hat{\sigma}_{N_p}^2 \sigma_s^2}{\tilde{\sigma}_{N_p}^2 \sigma_s^2 + \sigma^2}, \quad (6.40)$$

where

$$\hat{\sigma}_{N_p}^2 := \left[ \boldsymbol{\Sigma}_{\hat{\mathbf{h}}^{(N_p)}(i) | \mathbf{s}_{N_p}(i)} \right]_{N_p, N_p} \quad (6.41)$$

$$\tilde{\sigma}_{N_p}^2 := \left[ \boldsymbol{\Sigma}_{\tilde{\mathbf{h}}^{(N_p)}(i) | \mathbf{s}_{N_p}(i)} \right]_{N_p, N_p}, \quad (6.42)$$

maximizes a lower-bound on the achievable-rate of the weakest data substream, in particular, the lower bound that follows from the use of the (sub-optimal) combiner  $\mathbf{c}^{(N_p)}(i) = \mathbf{e}_N^{(N_p)}$  in (6.26).

*Proof.* See Appendix C.4 □

Figure 6.1 plots the power allocation parameter  $\alpha_{p,*}$  versus SNR for single-carrier transmission with  $N_s = 128$  substreams across the CE-BEM doubly selective fading channel with  $N_h = 8$  taps of ISSI and  $D \in \{1, 2\}$ . These parameters correspond to, e.g., a channel with bandwidth 1.5 MHz, carrier frequency 60 GHz, delay spread 5.4  $\mu\text{s}$ , and mobile and reflector velocities of  $\{69, 138\}$  km/h, in a “triple Doppler” scenario [75]. Figure 6.1 suggests that, at low values of SNR, additive noise level

dictates performance and more power is allocated to the data substreams. However, as the SNR grows, the effect of channel estimation error on performance becomes more pronounced, and the pilots are given more power to keep the estimation error in check.

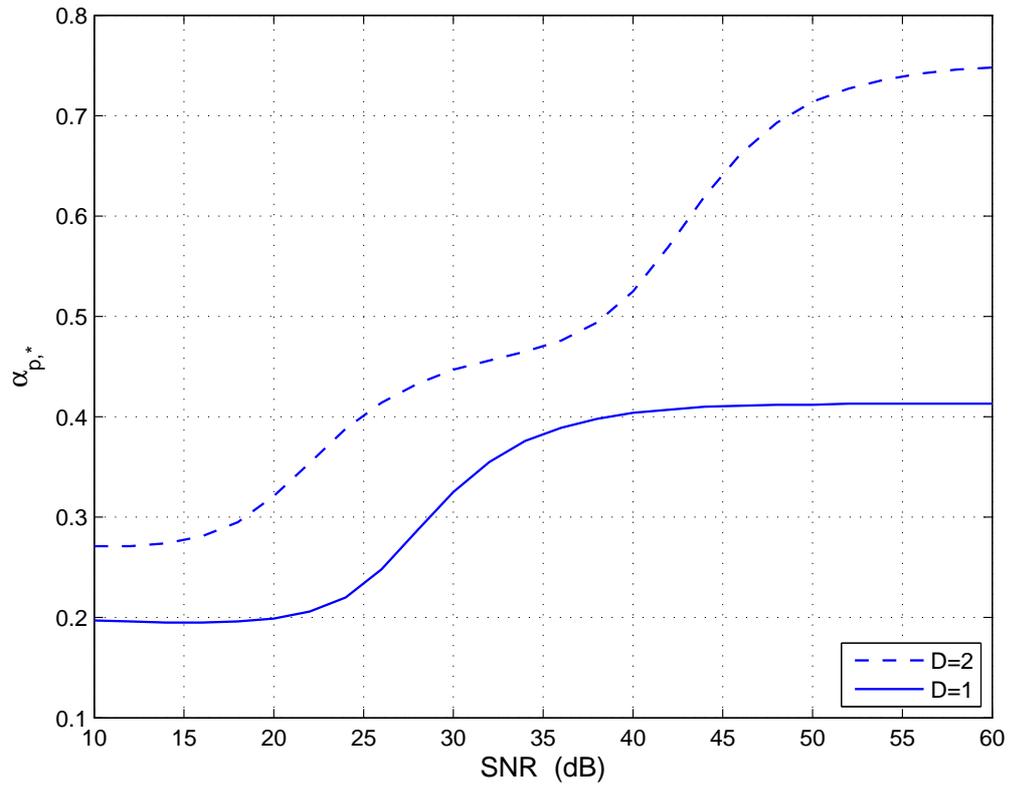


Figure 6.1: Power allocation parameter  $\alpha_{p,*}$  at various SNRs for a  $N_s = 128$ -substream spectrally efficient transmission over a CE-BEM doubly selective fading channel with  $N_h = 8$  taps of ISSI and  $D \in \{1, 2\}$ .

## CHAPTER 7

### CONCLUSIONS

#### 7.1 Summary of Our Work

We characterized the performance of local subcarrier processing (LSP)-constrained MCM reception on multi-antenna channel-estimation based reception for MCM pilot-aided transmission over WSSUS DS channels. It was found that the statistics of the observed interference make information-theoretic analysis difficult. We derived a surrogate model as an approximation for the MCM system.

With the use of i.i.d. Gaussian codebooks, we derived a lower-bound on the LSP-constrained achievable-rate of the surrogate MCM system for pilot aided MCM transmissions. This lower-bound was used to analyze the performance of the surrogate MCM system under the LSP constraint at high-SNR.

In doing so, we found it useful to split the analysis into two regimes of high-SNR based on the characteristics of observed interference. Most practical wireless systems are likely to operate in Regime 1, where additive noise and channel estimation errors dominate residual ICI and ISI components. In this regime, we established that multiple receive antennas are *necessary and sufficient* for spectrally efficient LSP-constrained MCM reception when i.i.d. Gaussian codebooks are used. Furthermore,

we demonstrated that the lower-bound could be tightened by increasing the LSP processing radius. However, tightening the bound also resulted in an increase in LSP implementation complexity since more observations and ICI coefficients were being used. The resulting trade-off suggested matching LSP's processing radius with the DICI radius for a lucrative compromise between performance and complexity. There also exists Regime 2 of yet higher SNRs where practical wireless systems are unlikely to operate in. In Regime 2, non-vanishing RICI and ISI components of interference dominate over additive noise. We showed that an achievable-rate ceiling for LSP-constrained reception existed in Regime 2.

The derived LSP-constrained achievable-rate metric provides a valuable MCM system design tool. For instance, it was used to compare MCM schemes over DS channels under a constraint on receiver complexity.

As another instance of the utility of the LSP-constrained achievable-rate metric, we presented beamforming vector and linear combiner designs for a MIMO-OFDM system appropriate for high mobility scenarios. Three novel designs: a SNR maximizing beamforming vector design, a linear combiner design, and a joint approximate max-achievable-rate design, were discussed. Results of numerical experiments suggested that our designs provided large gains over traditional designs, and remained robust to large Doppler spreads and predicted transmitter CSI, in spite of having the same complexity orders as traditional designs. Thus, they provide attractive alternatives to traditional ICI-ignoring beam formers/ combiners for MIMO-OFDM systems.

We also designed and analyzed a communication scheme for the DS channel based on pilot-aided transmission and successive-decoding with channel re-estimation. We

derived a lower-bound on the achievable rate and characterized the pre-log factor of the high-SNR achievable rate expression. For the special case of the CE-BEM doubly selective fading channel, we found the proposed communication system to be spectrally efficient. Finally, we designed a pilot/data power allocation strategy based on the maximization of an achievable-rate lower-bound.

## 7.2 Future Research Directions

In this dissertation, we derived a lower-bound on the LSP-constrained achievable rate of Gaussian coded Pilot-Aided MCM systems, and evaluated the performance gap between this lower-bound on LSP-constrained performance and global subcarrier processing (GSP) by means of simulations. However, in order to get a clearer picture of LSP-constrained performance, a theoretical analysis of the coding gain could be carried out. Such an analysis, for instance, could yield valuable insights into the design of MCM system parameters. A similar characterization of the coding gain would be interesting for the channel re-estimation and successive-decoding based reception strategy presented in Chapter 6.

The lower-bound on LSP-constrained achievable rates was used to design beamforming vectors and combiners for a MIMO-OFDM system. It is well known that a beamforming system does not fully exploit all the “degrees of freedom” provided by the MIMO setup. Thus, another possible extension of this work would be to extend the theoretical formulation and results from Chapter 4 from a single-input multiple-output (SIMO) setup to a spatial-multiplexing MIMO setup.

In practice, MCM is often used as a part of a communication system serving multiple users simultaneously. Information theory suggests that orthogonal multi-access,

e.g., OFDMA, is sub-optimal (see [76] and references therein). In this direction, inter-carrier interference (ICI) on MCM systems over DS channel could be advantageous if considered carefully. For instance, ICI could also be interpreted as spreading the inputs over subcarriers or the “frequency” resource. However, non-trivial multi-user receivers will be required in the presence of ICI. LSP-constrained multi-user reception on the up-link of multi-user MCM systems operating over DS channels could be investigated using the tools developed in this dissertation.

## APPENDIX A

### DERIVATION OF INTERFERENCE STATISTICS

In this appendix, we examine the properties of the interference term  $\mathbf{v}^{(n_r)}(i)$  in (3.23). Before expanding the terms in (3.24), we find it useful to recall

$$\mathbf{g}^{(n_r)}(i) := [\text{diag}_0(\mathbf{H}^{(n_r)}(i, 0))^T \cdots \text{diag}_{N-1}(\mathbf{H}^{(n_r)}(i, 0))^T]^T, \quad (\text{A.1})$$

where  $\text{diag}_k(\cdot)$  creates a vector from the  $k^{\text{th}}$  sub-diagonal of its matrix argument via  $\text{diag}_k(\mathbf{H}) := [[\mathbf{H}]_{k,0}, [\mathbf{H}]_{k+1,1}, \dots, [\mathbf{H}]_{k+N-1,N-1}]^T$  with modulo- $N$  indexing assumed. Using  $\hat{\mathbf{g}}^{(n_r)}(i)$  to denote the pilot-aided LMMSE estimate of  $\mathbf{g}^{(n_r)}(i)$ , and  $\tilde{\mathbf{g}}^{(n_r)}(i) := \mathbf{g}^{(n_r)}(i) - \hat{\mathbf{g}}^{(n_r)}(i)$  to denote the corresponding estimation error, the CEE term can be written as

$$\mathbf{w}_{\text{cee}}^{(n_r)}(i) = \mathbf{S}(i)\tilde{\mathbf{g}}^{(n_r)}(i) \quad (\text{A.2})$$

$$\mathbf{S}(i) := [\Theta^0 \mathcal{D}(\mathbf{s}(i)) \cdots \Theta^{N-1} \mathcal{D}(\mathbf{s}(i))], \quad (\text{A.3})$$

where

$$\Theta := \begin{bmatrix} \mathbf{0}_{N-1}^T & 1 \\ \mathbf{I}_{N-1} & \mathbf{0}_{N-1} \end{bmatrix} \quad (\text{A.4})$$

is the cyclic down-shift matrix, and  $\mathcal{D}(\cdot)$  creates a diagonal matrix from its vector argument. Using the expression derived for the covariance of  $\tilde{\mathbf{g}}^{(n_r)}(i)$ , we find that

for  $i \in \mathcal{I}_d$ ,

$$\begin{aligned}
\Sigma_{\mathbf{w}_{\text{cee}}^{(n_r)}(i)} &:= \mathbb{E} \left\{ \mathbf{w}_{\text{cee}}^{(n_r)}(i) \mathbf{w}_{\text{cee}}^{(n_r)}(i)^H \right\} \\
&= \mathbb{E} \left\{ \mathbf{S}(i) \tilde{\mathbf{g}}^{(n_r)}(i) \tilde{\mathbf{g}}^{(n_r)}(i)^H \mathbf{S}(i)^H \right\} \\
&= \mathbb{E} \left\{ \mathbf{S}(i) \mathbf{U}_h \left( \Lambda_h^{-1} + \mathbf{U}_h^H \mathbf{P}^H (\Sigma_{\underline{\mathbf{w}}_{\text{isi}}^{(n_r)}} + \Sigma_{\underline{\mathbf{w}}^{(n_r)}})^{-1} \mathbf{P} \mathbf{U}_h \right)^{-1} \right. \\
&\quad \left. \times \mathbf{U}_h^H \mathbf{S}(i)^H \right\}. \tag{A.5}
\end{aligned}$$

The RICCI term can then be written as

$$\mathbf{w}_{\text{rici}}^{(n_r)}(i) = \mathbf{S}_{D_h}(i) \hat{\mathbf{g}}_{D_h}^{(n_r)}(i) \tag{A.6}$$

$$\mathbf{S}_{D_h}(i) := [\Theta^{D_h+1} \mathcal{D}(\mathbf{s}(i)) \quad \dots \quad \Theta^{N-1-D_h} \mathcal{D}(\mathbf{s}(i))] \tag{A.7}$$

$$\hat{\mathbf{g}}_{D_h}^{(n_r)}(i) := [[\hat{\mathbf{g}}^{(n_r)}(i)]_{N(D_h+1)}, \dots, [\hat{\mathbf{g}}^{(n_r)}(i)]_{N(N-D_h)-1}]^T. \tag{A.8}$$

Alternatively, we could also have defined  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$  from the true RICCI coefficients  $\{\text{diag}_k(\mathbf{H}^{(n_r)}(i, 0))\}_{k=D_h+1}^{N-D_h-1}$  and defined  $\mathbf{w}_{\text{cee}}^{(n_r)}(i)$  from DICCI coefficient estimation error, we found it convenient to include the fictitious RICCI estimation error component  $\{\text{diag}_k(\tilde{\mathbf{H}}^{(n_r)}(i, 0))\}_{k=D_h+1}^{N-1-D_h}$  in  $\mathbf{w}_{\text{cee}}^{(n_r)}(i)$  instead of  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$ . We refer to this quantity as ‘‘fictitious’’ because, in practice, the RICCI coefficient estimates would never be computed.

The ISI term as

$$\mathbf{w}_{\text{isi}}^{(n_r)}(i) = \sum_{q \in \{-L_{\text{pre}}, \dots, L_{\text{pst}}\} \setminus 0} \mathbf{H}^{(n_r)}(i, q) \mathbf{s}(i - q). \tag{A.9}$$

Note that, in general,  $\mathbf{w}_{\text{cee}}^{(n_r)}(i)$  and  $\mathbf{w}_{\text{isi}}^{(n_r)}(i)$  are non-Gaussian when  $i \in \mathcal{I}_d$ .

We now examine the statistics of  $\mathbf{v}^{(n_r)}(i)$ . Due to the orthogonality principle of MMSE estimation,  $\mathbf{w}_{\text{cee}}^{(n_r)}(i)$  is uncorrelated with  $\hat{\mathbf{H}}^{(n_r)}(i) \mathbf{s}(i)$  as well as  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$ , and, due to zero-mean uncorrelated symbols, it is also uncorrelated with  $\mathbf{w}_{\text{isi}}^{(n_r)}(i)$

and  $\mathbf{w}^{(n_r)}(i)$ . However, the autocovariance of  $\mathbf{w}_{\text{cee}}^{(n_r)}(i)$  may vary with  $i$ , as seen from (A.5). Zero-mean uncorrelated symbols also imply that  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$  is uncorrelated with  $\mathbf{w}_{\text{isi}}^{(n_r)}(i)$  and  $\mathbf{w}^{(n_r)}(i)$ . But,  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$  is correlated with  $\hat{\mathbf{H}}^{(n_r)}(i)\mathbf{s}(i)$  via the common symbol  $\mathbf{s}(i)$  and correlation among the ICI coefficients. Finally,  $\mathbf{w}_{\text{isi}}^{(n_r)}(i)$  and  $\mathbf{w}^{(n_r)}(i)$  are uncorrelated with each other as well as with  $\hat{\mathbf{H}}^{(n_r)}(i)\mathbf{s}(i)$ . In summary, the four components of  $\mathbf{v}^{(n_r)}(i)$  identified in (3.24) are mutually uncorrelated, so that, with

$$\Sigma_{\mathbf{v}^{(n_r)}(i)} := \mathbb{E}\{\mathbf{v}^{(n_r)}(i)\mathbf{v}^{(n_r)}(i)^H\}, \quad (\text{A.10})$$

$$\Sigma_{\mathbf{w}_{\text{cee}}^{(n_r)}(i)} := \mathbb{E}\{\mathbf{w}_{\text{cee}}^{(n_r)}(i)\mathbf{w}_{\text{cee}}^{(n_r)}(i)^H\}, \quad (\text{A.11})$$

$$\Sigma_{\mathbf{w}_{\text{isi}}^{(n_r)}(i)} := \mathbb{E}\{\mathbf{w}_{\text{isi}}^{(n_r)}(i)\mathbf{w}_{\text{isi}}^{(n_r)}(i)^H\}, \quad (\text{A.12})$$

$$\Sigma_{\mathbf{w}_{\text{rici}}^{(n_r)}(i)} := \mathbb{E}\{\mathbf{w}_{\text{rici}}^{(n_r)}(i)\mathbf{w}_{\text{rici}}^{(n_r)}(i)^H\}, \quad (\text{A.13})$$

$$\Sigma_{\mathbf{w}^{(n_r)}(i)} := \mathbb{E}\{\mathbf{w}^{(n_r)}(i)\mathbf{w}^{(n_r)}(i)^H\}, \quad (\text{A.14})$$

we can write

$$\Sigma_{\mathbf{v}^{(n_r)}(i)} = \Sigma_{\mathbf{w}_{\text{cee}}^{(n_r)}(i)} + \Sigma_{\mathbf{w}_{\text{isi}}^{(n_r)}(i)} + \Sigma_{\mathbf{w}_{\text{rici}}^{(n_r)}(i)} + \Sigma_{\mathbf{w}^{(n_r)}(i)}. \quad (\text{A.15})$$

Note, however, that  $\mathbf{v}^{(n_r)}(i)$  is correlated with the signal term  $\hat{\mathbf{H}}^{(n_r)}(i)\mathbf{s}(i)$  through  $\mathbf{w}_{\text{rici}}^{(n_r)}(i)$ . In addition,  $\mathbf{v}^{(n_r)}(i_1)$  may be correlated with  $\mathbf{v}^{(n_r)}(i_2)$  for  $i_1 \neq i_2$ .

Finally note that the true channel coefficients as well as the noise at each receive antenna is uncorrelated. Then all the constituents of  $\mathbf{v}^{(n_r)}(i)$  are also uncorrelated across receive antennas. Consequently, noting that  $\Sigma_{\mathbf{v}^{(n_r)}(i)}$  is invariant with  $n_r$ , we obtain

$$\begin{aligned} \Sigma_{\mathbf{v}(i)} &= \mathbb{E}\left\{\left(\sum_{n_r=1}^{N_r} \mathbf{v}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}\right)\left(\sum_{n_r=1}^{N_r} \mathbf{v}^{(n_r)}(i) \otimes \mathbf{e}_{N_r}^{(n_r)}\right)^H\right\} \\ &= \Sigma_{\mathbf{v}^{(0)}(i)} \otimes \mathbf{I}_{N_r}. \end{aligned} \quad (\text{A.16})$$

## APPENDIX B

### LSP-CONSTRAINED ACHIEVABLE RATE ANALYSIS

#### B.1 Surrogate System

We can rewrite (3.28) using (3.23) and (3.24) as

$$\mathbf{r}(i) = \hat{\mathbf{H}}(i)\mathbf{s}(i) + \hat{\mathbf{H}}_{\text{rici}}(i)\mathbf{s}(i) + \underbrace{\tilde{\mathbf{H}}(i)\mathbf{s}(i) + \sum_{q \neq 0} \mathbf{H}(i, q)\mathbf{s}(i - q)}_{\text{non-Gaussian}} + \mathbf{w}(i), \quad (\text{B.1})$$

where,  $\hat{\mathbf{H}}(i)$  contains estimates of the dominant ICI coefficients,  $\hat{\mathbf{H}}_{\text{rici}}$  the estimates of RICI coefficients, and  $\tilde{\mathbf{H}}(i)$  is the respective estimation error. The last two terms on the r.h.s. show the effect of ISI and additive noise, respectively. We now wish to characterize  $I(\mathbf{r}(i); \mathbf{s}(i) | \hat{\mathbf{H}}(i))$ . ISI affects performance in two ways. It acts as uncorrelated interference during decoding as seen in (B.1). It also acts as a source of uncorrelated non-Gaussian noise during channel estimation as seen in (3.9). By replacing the ISI term by a Gaussian random vector that otherwise preserves the statistics, we ensure that the channel estimate is a MMSE estimate. We note from [42] that when MMSE channel estimates are used, a lower-bound on the mutual information can be calculated by replacing the non-Gaussian effect of estimation error,  $\tilde{\mathbf{H}}(i)\mathbf{s}(i)$ , by a Gaussian random vector that otherwise preserve the statistics. Hence,  $I(\mathbf{r}(i); \mathbf{s}(i) | \hat{\mathbf{H}}(i)) \geq I(\mathbf{r}'(i); \mathbf{s}(i) | \hat{\mathbf{H}}(i))$  with  $\mathbf{r}'(i)$  defined for a Gaussian distributed

$\mathbf{v}'(i)$  as

$$\mathbf{r}'(i) = \hat{\mathbf{H}}(i)\mathbf{s}(i) + \hat{\mathbf{H}}_{\text{rici}}(i)\mathbf{s}(i) + \mathbf{v}'(i), \quad (\text{B.2})$$

$$\boldsymbol{\Sigma}_{\mathbf{v}'(i)} = \boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}(i)}. \quad (\text{B.3})$$

As mentioned before, RICCI is treated as interference in LSP. However, this is difficult to incorporate in achievable rate calculations since it is correlated with the signal component  $\hat{\mathbf{H}}(i)\mathbf{s}(i)$ . Hence, to make the analysis tractable, we replace the RICCI by a Gaussian random vector uncorrelated with the signal component that otherwise preserves the statistics of the RICCI. With this substitution, we obtain the surrogate system

$$\mathbf{r}''(i) = \hat{\mathbf{H}}(i)\mathbf{s}(i) + \mathbf{v}''(i), \quad (\text{B.4})$$

$$\boldsymbol{\Sigma}_{\mathbf{v}''(i)} = \boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{rici}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}(i)}. \quad (\text{B.5})$$

For the surrogate system (B.4) to well approximate (B.2), the power of the RICCI must be significantly lower than that of signal component. This approximation works well for our setup since, by definition, the dominant ICI coefficients contain almost all of the channel energy, and only traces remain in the RICCI coefficients. For instance, in the test system used in Section 4.5, the ratio of powers of the dominant and residual ICI coefficients is in excess of 70 dB.

## B.2 Proof of Lemma 1

In this proof we consider  $N_r = 1$  receive antennas. The extension to  $N_r > 1$  receive antennas requires identical steps but with more complex expressions. Recall

from (3.24) that for  $i \in \mathcal{I}_d$

$$\text{tr}(\boldsymbol{\Sigma}_{\mathbf{v}(i)}) = \text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{rici}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)}) \quad (\text{B.6})$$

$$= \sigma^2 \text{tr}(\mathbf{J} \mathcal{D}(\mathbf{b})^2 \mathbf{J}^H) + \text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{rici}}(i)}) + \text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)}). \quad (\text{B.7})$$

Furthermore, recall from Section 3.3 that

$$\boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)} = \text{E} \left[ \mathbf{P}(i) \mathbf{U}_h (\boldsymbol{\Lambda}_h^{-1} + \mathbf{U}_h^H \underline{\mathbf{P}}^H (\boldsymbol{\Sigma}_{\underline{\mathbf{w}}_{\text{isi}}^{(m)}} + \boldsymbol{\Sigma}_{\underline{\mathbf{w}}^{(m)}})^{-1} \underline{\mathbf{P}} \mathbf{U}_h)^{-1} \mathbf{U}_h^H \mathbf{P}(i)^H \right]. \quad (\text{B.8})$$

Here the expectation is with respect to the distribution of the data symbols. Note that estimation in the presence of stronger noise will result in an increase in estimation error. Since  $2\lambda_{\text{isi}} \mathbf{I} > \boldsymbol{\Sigma}_{\underline{\mathbf{w}}_{\text{isi}}^{(m)}}$  and  $2\lambda_b \sigma^2 \mathbf{I} > \boldsymbol{\Sigma}_{\underline{\mathbf{w}}^{(m)}}$ , we have  $2(\lambda_{\text{isi}} + \lambda_b \sigma^2) \mathbf{I} > \boldsymbol{\Sigma}_{\underline{\mathbf{w}}_{\text{isi}}^{(m)}} + \boldsymbol{\Sigma}_{\underline{\mathbf{w}}^{(m)}}$ . Then channel-estimation in the presence of noise with covariance  $2(\lambda_{\text{isi}} + \lambda_b \sigma^2) \mathbf{I}$  would result in higher estimation error than that for noise with covariance  $\boldsymbol{\Sigma}_{\underline{\mathbf{w}}_{\text{isi}}^{(m)}} + \boldsymbol{\Sigma}_{\underline{\mathbf{w}}^{(m)}}$ . This implies that

$$\begin{aligned} \text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}_{\text{cee}}(i)}) &\leq \text{tr} \text{E} \left[ \mathbf{P}(i) \mathbf{U}_h (\boldsymbol{\Lambda}_h^{-1} + \mathbf{U}_h^H \underline{\mathbf{P}}^H \underline{\mathbf{P}} \mathbf{U}_h \right. \\ &\quad \left. \times \frac{(\lambda_{\text{isi}} + \lambda_b \sigma^2)^{-1}}{2})^{-1} \mathbf{U}_h^H \mathbf{P}(i)^H \right], \end{aligned} \quad (\text{B.9})$$

$$\begin{aligned} &\leq 2(\lambda_{\text{isi}} + \lambda_b \sigma^2) \text{tr} \left[ \text{E} \{ \mathbf{P}(i)^H \mathbf{P}(i) \} \right. \\ &\quad \left. \times \mathbf{U}_h (\mathbf{U}_h^H \underline{\mathbf{P}}^H \underline{\mathbf{P}} \mathbf{U}_h)^{-1} \mathbf{U}_h^H \right]. \end{aligned} \quad (\text{B.10})$$

Equation (B.10) is obtained by applying Lemma 1 from [70] since we assumed in Section 3.3 that  $\underline{\mathbf{P}} \mathbf{U}_h$  is full column rank. Then defining

$$\alpha_p := \text{tr} \left[ \text{E} \{ \mathbf{P}(i)^H \mathbf{P}(i) \} \mathbf{U}_h (\mathbf{U}_h^H \underline{\mathbf{P}}^H \underline{\mathbf{P}} \mathbf{U}_h)^{-1} \mathbf{U}_h^H \right], \quad (\text{B.11})$$

and substituting (B.10) in (B.7), we obtain

$$\begin{aligned} \text{tr}(\boldsymbol{\Sigma}_{\mathbf{v}(i)}) &\leq \sigma^2 [\text{tr}(\mathbf{J} \mathcal{D}(\mathbf{b})^2 \mathbf{J}^H) + 2\alpha_p \lambda_b] + \text{tr}(\boldsymbol{\Sigma}_{\mathbf{w}_{\text{isi}}(i)} + \boldsymbol{\Sigma}_{\mathbf{w}_{\text{rici}}(i)}) \\ &\quad + 2\alpha_p \lambda_{\text{isi}}. \end{aligned} \quad (\text{B.12})$$

The result follows. This concludes the proof.

### B.3 Proof for Theorem 1

Under the assumptions of Section 3.2, codewords on each subcarrier are decoded individually, and subcarriers are decoded in sequence. The rate allocated to the code whose codewords are modulated on the  $k^{\text{th}}$  subcarrier of the  $i^{\text{th}}$  data multicarrier symbol of each frame is given by  $\mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{r}_k(i); s_k(i) | \{s_d(i)\}_{d=0}^{k-1})$ . Then recalling IC from (4.8) and noting that, by definition,  $\mathcal{K}_k^- \subset \{d : 0 \leq d < k\}$ , we can write

$$\begin{aligned} \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{r}_k(i); s_k(i) | \{s_d(i)\}_{d=0}^{k-1}) &= \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{r}_k(i) \\ &\quad - \sum_{d \in \mathcal{K}_k^-} \hat{\mathbf{h}}_{k,d}(i) s_d(i) \Big| \{s_d(i)\}_{d=0}^{k-1}) \end{aligned} \quad (\text{B.13})$$

$$= \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{y}_k(i); s_k(i)). \quad (\text{B.14})$$

This shows that IC is information-lossless. Further, with the surrogate system, we have Gaussian signaling in the presence of Gaussian interference and noise, in which case max-SINR combining is shown to be information-lossless in [76, 77]. Then using the LC-optimal max-SINR combiner  $\mathbf{z}_k^m(i)$ ,

$$\mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{y}_k(i); s_k(i)) = \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\mathbf{z}_k^m(i)^H \mathbf{y}_k(i); s_k(i)) \quad (\text{B.15})$$

$$= \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\hat{s}_k(i); s_k(i)). \quad (\text{B.16})$$

Substituting (B.16) in (B.14) and the result in the definition of LSP-constrained achievable rate in (4.7), we obtain

$$R_{D_r} = \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \sum_{k=0}^{N-1} \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\hat{s}_k(i); s_k(i)). \quad (\text{B.17})$$

Clearly, IC-LC-LSP preserves mutual information and is achievable-rate-optimal for the surrogate system. Further, mutual information can be related to post-combining SINR for Gaussian signaling in the presence of Gaussian interference and noise [76],

so that  $\mathbb{E}_{\hat{\mathbf{H}}(i)} I(\hat{s}_k(i); s_k(i)) = \mathbb{E}_{\hat{\mathbf{H}}(i)} \log(1 + \gamma_k(\mathbf{z}_k^m(i)))$ . Then from the definition of max-SINR LC in (4.15),

$$\begin{aligned} \mathbb{E}_{\hat{\mathbf{H}}(i)} I(\hat{s}_k(i); s_k(i)) &= \mathbb{E}_{\hat{\mathbf{H}}(i)} \log \left( 1 + \hat{\mathbf{h}}_k(i)^H \right. \\ &\quad \left. \times \left( \sum_{m \in \mathcal{K}_k^+} \hat{\mathbf{h}}_{k,m}(i) \hat{\mathbf{h}}_{k,m}^H(i) + \boldsymbol{\Sigma}_{\mathbf{v}_k(i)} \right)^{-1} \hat{\mathbf{h}}_k(i) \right) \end{aligned} \quad (\text{B.18})$$

Substituting (B.18) in (B.17), we arrive at (4.16). This completes the proof.

## B.4 Proof for Theorem 2

First we show that when (4.25) holds, ZF combiners exist w.p.1 for every combination of  $k \in \{0, \dots, N-1\}$  and  $i \in \mathcal{I}_d$ . Let us denote the left null space of the dominant ICI matrix,  $\hat{\mathbf{H}}_k(i)$ , by  $\mathcal{N}(\hat{\mathbf{H}}_k(i))$ , and its column space by  $\mathcal{C}(\hat{\mathbf{H}}_k(i))$ . Recall that  $\hat{\mathbf{H}}_k(i) \in \mathbb{C}^{(2D_r+1)N_r \times 2(D_r+D_h)}$ . Rewriting (4.25) as  $(2D_r+1)N_r > 2(D_r+D_h)$ , we find that  $\hat{\mathbf{H}}_k$  has more rows than columns, and hence that  $\dim(\mathcal{N}(\hat{\mathbf{H}}_k(i))) \geq (2D_r+1)N_r - 2(D_r+D_h) > 0$ . Thus, by choosing  $\mathbf{z}_k^{\text{zf}}(i) \in \mathcal{N}(\hat{\mathbf{H}}_k(i))$  with  $\|\mathbf{z}_k^{\text{zf}}(i)\| = 1$ , we ensure that  $\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{H}}_k(i) = \mathbf{0}$ .

We now need to show that  $\exists \mathbf{z}_k^{\text{zf}}(i) \in \mathcal{N}(\hat{\mathbf{H}}_k(i))$  for which  $\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i) > 0$  w.p.1. Recall that every vector in  $\mathbb{C}^{(2D_r+1)N_r}$  can be decomposed into orthogonal components in  $\mathcal{N}(\hat{\mathbf{H}}_k(i))$  and  $\mathcal{C}(\hat{\mathbf{H}}_k(i))$ . Thus, we equivalently show that  $\hat{\mathbf{h}}_{k,k}(i) \notin \mathcal{C}(\hat{\mathbf{H}}_k(i))$  w.p.1 below. Clearly, from the definitions of  $\hat{\mathbf{h}}_{k,k}(i)$  and  $\hat{\mathbf{H}}_k(i)$  in Section 3.4, there exists a permutation matrix  $\mathbf{P}$  for which  $\mathbf{P}\hat{\mathbf{h}}_{k,k}(i) = [\hat{\mathbf{h}}_{k,k}^{(1)}(i)^T, \hat{\mathbf{h}}_{k,k}^{(2)}(i)^T, \dots, \hat{\mathbf{h}}_{k,k}^{(N_r)}(i)^T]^T$ , and  $\mathbf{P}\hat{\mathbf{H}}_k(i) = [\hat{\mathbf{H}}_k^{(1)}(i)^T, \hat{\mathbf{H}}_k^{(2)}(i)^T, \dots, \hat{\mathbf{H}}_k^{(N_r)}(i)^T]^T$ . Then  $\hat{\mathbf{h}}_{k,k}^{(m)}(i) \in \mathcal{C}(\hat{\mathbf{H}}_k^{(m)}(i))$ ,  $\forall m \in \{1, \dots, N_r\}$  are necessary conditions for  $\hat{\mathbf{h}}_{k,k}(i) \in \mathcal{C}(\hat{\mathbf{H}}_k(i))$ . These necessary conditions imply that  $\exists \mathbf{b}_k^{(m)}(i) \neq \mathbf{0}$  such that  $\hat{\mathbf{h}}_{k,k}^{(m)}(i) = \hat{\mathbf{H}}_k^{(m)}(i) \mathbf{b}_k^{(m)}(i)$ , for each  $m \in \{1, \dots, N_r\}$ . Notice that  $\hat{\mathbf{h}}_{k,k}^{(m_1)}(i)$  and  $\hat{\mathbf{h}}_{k,k}^{(m_2)}(i)$  (as well as  $\hat{\mathbf{H}}_k^{(m_1)}(i)$  and  $\hat{\mathbf{H}}_k^{(m_2)}(i)$ )

are independent when  $m_1 \neq m_2$  due to independence of the spatial paths. Then the random vectors  $\mathbf{b}_k^{(m_1)}(i)$  and  $\mathbf{b}_k^{(m_2)}(i)$  are independent when  $m_1 \neq m_2$  as well. However, for  $\hat{\mathbf{h}}_{k,k}(i) \in \mathcal{C}(\hat{\mathbf{H}}_k(i))$ ,  $\exists \mathbf{b}_k(i)$ , such that  $\hat{\mathbf{h}}_{k,k}(i) = \hat{\mathbf{H}}_k(i)\mathbf{b}_k(i)$ . This can only happen if  $\mathbf{b}_k^{(1)}(i) = \mathbf{b}_k^{(2)}(i) = \dots = \mathbf{b}_k^{(N_r)}(i) = \mathbf{b}_k(i)$ . But  $P(\mathbf{b}_k^{(1)}(i) = \mathbf{b}_k^{(2)}(i) = \dots = \mathbf{b}_k^{(N_r)}(i)) = 0$  as vectors in the set  $\{\mathbf{b}_k^{(m)}(i)\}_{m=1}^{N_r}$  are mutually independent. Hence  $\hat{\mathbf{h}}_{k,k}(i) \in \mathcal{C}(\hat{\mathbf{H}}_k(i))$  with probability zero, or equivalently  $\hat{\mathbf{h}}_{k,k}(i) \notin \mathcal{C}(\hat{\mathbf{H}}_k(i))$  w.p.1.

Next, for the  $N_r = 1$  case, we show in the following that  $\dim(\mathcal{N}(\hat{\mathbf{H}}_k(i))) = 0$  w.p.1, which implies that ZF combiners do not exist for  $N_r = 1$  w.p.1. Note that  $\hat{\mathbf{H}}_k(i) \in \mathbb{C}^{(2D_r+1) \times 2(D_r+D_h)}$  when  $N_r = 1$ . However,  $2D_r + 1 \leq 2(D_r + D_h)$  if  $D_h > 0$ . Then  $\hat{\mathbf{H}}_k(i)$  has more columns than rows, and it suffices to show that the rows of  $\hat{\mathbf{H}}_k(i)$  are linearly independent w.p.1. In this proof, we refer to the  $m$ -th row of  $\hat{\mathbf{H}}_k(i)$  as  $\check{\mathbf{h}}_{k,m}(i)^T$  so that  $\hat{\mathbf{H}}_k(i) = [\check{\mathbf{h}}_{k,0}(i), \dots, \check{\mathbf{h}}_{k,2D_r}(i)]^T$ . For ease of presentation, we split the proof into two cases:

When  $D_h \leq D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ : The structure of  $\hat{\mathbf{H}}_k(i)$  is shown in Fig. B.1 where non-zero entries are only located in the shaded regions. We observe from Fig. B.1 that the  $(D_r - D_h + 1)$  topmost rows (i.e., rows above the dashed horizontal line in Fig. B.1) are linearly independent w.p.1 because of the upper-triangular structure. Next, observe that the  $(D_r + D_h)$  bottom-most rows (i.e., rows below horizontal dashed line in Fig. B.1) are also linearly independent w.p.1 because of the lower-triangular structure. Moreover, notice that on one hand, every element in and beyond the  $(D_r + D_h)$ -th column (i.e., elements right of the vertical dashed line in Fig. B.1) in each of the topmost  $(D_r - D_h + 1)$  rows is zero. On the other hand, each of the bottom-most  $(D_r + D_h)$  rows has at least one element in and beyond the  $(D_r + D_h)$ -th column (i.e., elements right of the vertical dashed line in Fig. B.1) that is non-zero

w.p.1. Thus, we conclude that the rows of  $\hat{\mathbf{H}}_k(i)$  are linearly independent w.p.1. for  $D_h \leq D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ .

When  $D_r < D_h$ : The structure of  $\hat{\mathbf{H}}_k(i)$  is shown in Fig. B.2 where non-zero entries are only located in the shaded regions. It is straightforward to see that the  $(2D_r + 1)$  rows of  $\hat{\mathbf{H}}_k(i)$  are linearly independent w.p.1 because of the upper-triangular structure. This completes the proof.

## B.5 Proof for Theorem 3

We first show that the spectral efficiency of IC-LC-LSP is at least  $\frac{Q_d N}{Q_c N_s}$  on the surrogate system. In order to do so, we first consider the LC-LSP strategy with ZF combiners.

Let the rate allocated to the code whose symbols are transmitted on the  $k$ -th subcarrier of the  $i$ -th data multicarrier symbol in every frame be denoted by  $R''_{k,i,D_r}$ . Let us denote a ZF combiner for the  $k$ -th subcarrier of the  $i$ -th data multicarrier symbol by  $\mathbf{z}_k^{\text{zf}}(i)$ . Since we are considering LC-LSP with the ZF combiner on the surrogate system,

$$R''_{k,i,D_r} = \mathbb{E}_{\hat{\mathbf{H}}(i)} \left[ \log(1 + \gamma'_k(\mathbf{z}_k(i))) \right]. \quad (\text{B.19})$$

Dominant ICI is perfectly suppressed by ZF combiners. Then, (4.19) indicates that

$$\gamma'_k(\mathbf{z}_k^{\text{zf}}(i)) = \frac{\left| \mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i) \right|^2}{\mathbf{z}_k^{\text{zf}}(i)^H \boldsymbol{\Sigma}_{\mathbf{v}_k(i)} \mathbf{z}_k^{\text{zf}}(i)}, \quad (\text{B.20})$$

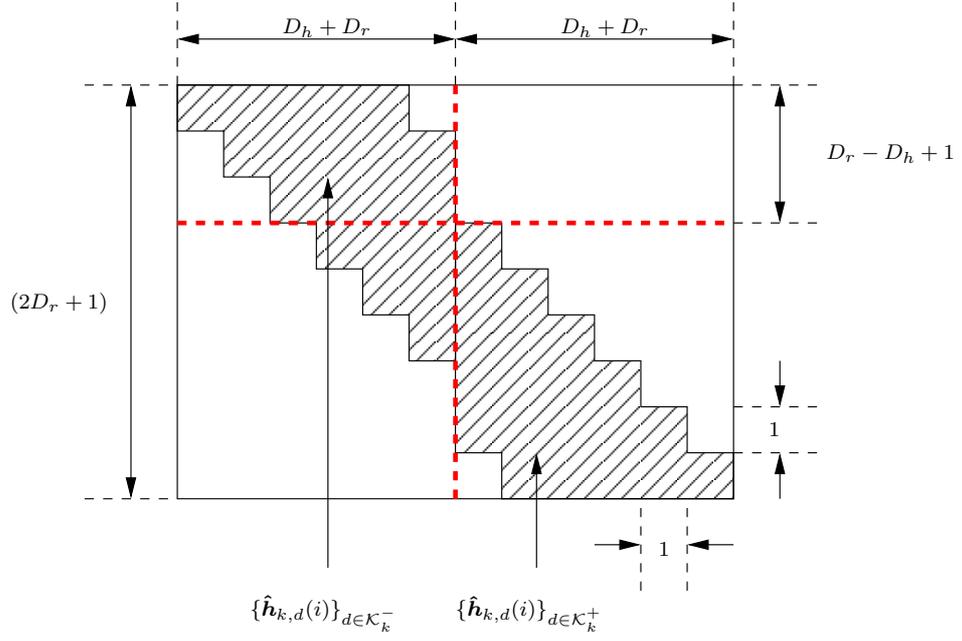


Figure B.1: Structure of the dominant ICI matrix  $\hat{\mathbf{H}}_k(i)$  when  $N_r = 1$  and  $D_h \leq D_r \leq \lfloor \frac{N-1}{2} \rfloor - D_h$ .

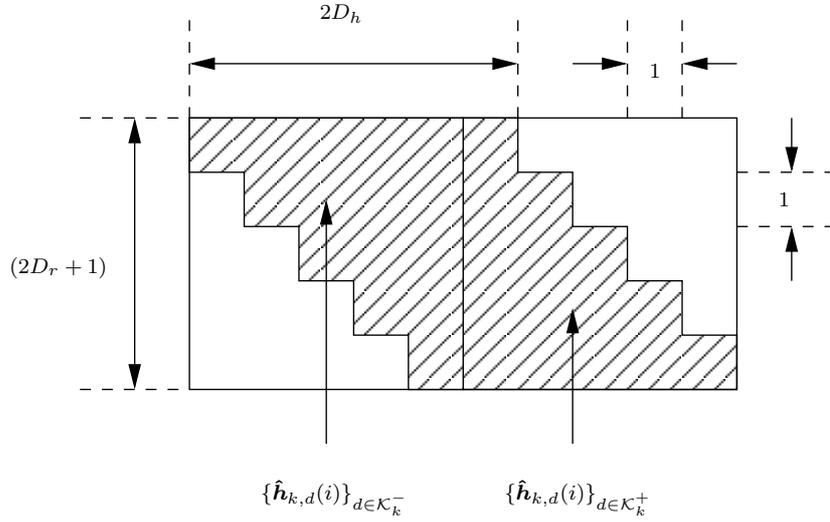


Figure B.2: Structure of the dominant ICI matrix  $\hat{\mathbf{H}}_k(i)$  when  $N_r = 1$  and  $D_r < D_h$ .

Note that applying (4.4) for Regime 1, we obtain  $\gamma'_k(\mathbf{z}_k^{\text{zf}}(i)) \geq |\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i)|^2 / (2C\sigma^2)$ . Defining  $\tau_k(i) := |\mathbf{z}_k^{\text{zf}}(i)^H \hat{\mathbf{h}}_{k,k}(i)|^2 / (2C)$ , we can say that  $\gamma_k(\mathbf{z}_k^{\text{zf}}(i)) \geq \tau_k(i)\rho$ . Replacing in (B.19) and dividing both sides by  $\log \rho$ , we obtain

$$\frac{R''_{k,i,D_r}(\rho)}{\log \rho} \geq \frac{\mathbb{E}_{\hat{\mathbf{H}}(i)} \log(1 + \tau_k(i)\rho)}{\log \rho} \quad (\text{B.21})$$

$$\geq 1 + \frac{\mathbb{E}_{\hat{\mathbf{H}}(i)} \log(\rho^{-1} + \tau_k(i))}{\log \rho}. \quad (\text{B.22})$$

From the definition of ZF combiners,

$$\lim_{\rho \rightarrow \infty} \frac{\log(\rho^{-1} + \tau_k(i))}{\log \rho} = \begin{cases} 0, & \tau_k(i) > 0 \\ -1, & \tau_k(i) = 0 \end{cases} \quad (\text{B.23})$$

$$= 0 \quad \text{w.p.1} . \quad (\text{B.24})$$

Now, taking limit infimum on both sides of (B.22),

$$\liminf_{\rho \rightarrow \infty} \frac{R''_{k,i,D_r}(\rho)}{\log \rho} = 1 + \liminf_{\rho \rightarrow \infty} \frac{\mathbb{E}_{\hat{\mathbf{H}}(i)} \log(\rho^{-1} + \tau_k(i))}{\log \rho} \quad (\text{B.25})$$

$$\geq 1 + \mathbb{E}_{\hat{\mathbf{H}}(i)} \left( \liminf_{\rho \rightarrow \infty} \frac{\log(\rho^{-1} + \tau_k(i))}{\log \rho} \right) \quad (\text{B.26})$$

$$\geq 1. \quad (\text{B.27})$$

In the above, the penultimate step is an application of Fatou's Lemma [78] and the last step applies (B.24).

Note that interference-cancellation (IC) will only enhance performance and the achievable-rate for IC-LC-LSP will exceed that of LC-LSP, implying that

$$R_{D_r}(\rho) \geq \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \sum_{k=0}^{N-1} R''_{k,i,D_r}(\rho). \quad (\text{B.28})$$

Therefore, we conclude from (B.27) and (B.28) that

$$\liminf_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} \geq \frac{Q_d N}{Q_c N_s}. \quad (\text{B.29})$$

Now consider the genie-aided global subcarrier processing (GGSP) strategy where each codeword is optimally decoded under perfect CSI. Since we transmit  $Q_d N$  codes from a single transmit antenna over  $Q_c N_s$  signaling dimensions, the spectral efficiency of GGSP is less than  $\frac{Q_d N}{Q_c N_s}$  from standard results in [76]. But, the spectral efficiency of GGSP bounds that of LSP from above. Thus we arrive at the following upper bound for the spectral efficiency of LSP:

$$\limsup_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} \leq \frac{Q_d N}{Q_c N_s}. \quad (\text{B.30})$$

Finally, (B.29) and (B.30) leads us to the required result. This completes the proof.

## B.6 Proof of Theorem 4

We consider LSP under the assumption of perfect CSI for this proof by setting the estimation error to zero. Note that even with perfect CSI, the properties of ZF combiners derived in Theorem 2 hold, and ZF combiners do not exist for the MCM system with  $N_r = 1$  since  $\hat{\mathbf{H}}_k(i)$  is full row-rank w.p.1. Additionally, we assume that (aided by a genie), residual ICI and ISI are perfectly canceled. This implies that  $\mathbf{v}_k(i) = \mathbf{w}_k(i)$  and hence that  $\Sigma_{\mathbf{v}_k(i)} = \sigma^2 \Sigma'$ , where  $\Sigma'$  is a sub-matrix of  $\mathbf{F}_N \mathbf{J} \text{diag}(\mathbf{b})^2 \mathbf{J}^H \mathbf{F}_N^H$ . Then  $\mathbf{v}_k(i)$  is Gaussian distributed and from Theorem 1, IC-LC-LSP is the achievable-rate optimal LSP strategy.

Let the rate allocated (to ensure reliable decoding with IC-LC-LSP) to the code whose symbols are transmitted on the 0-th subcarrier of the  $i$ -th data multicarrier symbol in every frame be denoted by  $R''_{0,i,D_r}$ . Recall that no interference cancellation is possible for the  $k = 0$ -th subcarrier, and IC-LC-LSP boils down to LC-LSP on the  $k = 0$ -th subcarrier. Since the interference  $\mathbf{v}_k(i)$  is Gaussian distributed, Theorem 1

tells us that

$$R''_{0,i,D_r} = \mathbb{E}_{\hat{\mathbf{H}}(i)} \left[ \log(1 + \hat{\mathbf{h}}_{0,0}(i)^H (\hat{\mathbf{H}}_0(i) \hat{\mathbf{H}}_0(i)^H + \boldsymbol{\Sigma}_{\mathbf{v}_k(i)})^{-1} \hat{\mathbf{h}}_{0,0}(i)) \right] \quad (\text{B.31})$$

$$\leq \log \left[ \mathbb{E}_{\hat{\mathbf{H}}(i)} \left( 1 + \underbrace{\hat{\mathbf{h}}_{0,0}(i)^H (\hat{\mathbf{H}}_0(i) \hat{\mathbf{H}}_0(i)^H + \rho^{-1} \boldsymbol{\Sigma}'_{\mathbf{w}})^{-1} \hat{\mathbf{h}}_{0,0}(i)}_{Q_0(i)} \right) \right], \quad (\text{B.32})$$

where (B.32) is obtained by applying Jensen's inequality [74]. Let the eigen decomposition of  $\boldsymbol{\Sigma}'_{\mathbf{w}} = \mathbf{U}_w \boldsymbol{\Lambda}_w \mathbf{U}_w^H$ . Then it is straightforward to show that

$$Q_0(i) = \bar{\mathbf{h}}_0(i)^H (\bar{\mathbf{H}}_0(i) \bar{\mathbf{H}}_0(i)^H + \rho^{-1} \mathbf{I})^{-1} \bar{\mathbf{h}}_0(i), \quad (\text{B.33})$$

where,  $\bar{\mathbf{h}}_0(i) := \boldsymbol{\Lambda}_w^{-\frac{1}{2}} \mathbf{U}_w^H \hat{\mathbf{h}}_{0,0}(i)$ , and  $\bar{\mathbf{H}}_0(i) := \boldsymbol{\Lambda}_w^{-\frac{1}{2}} \mathbf{U}_w^H \hat{\mathbf{H}}_0(i)$ . We find that  $\bar{\mathbf{H}}_0(i)$  is full row-rank w.p.1 since  $\boldsymbol{\Lambda}_w^{-\frac{1}{2}} \mathbf{U}_w^H$  is invertible. Let the eigen-decomposition of  $\bar{\mathbf{H}}_0(i) \bar{\mathbf{H}}_0(i)^H = \bar{\mathbf{U}}_0(i) \bar{\boldsymbol{\Lambda}}_0(i) \bar{\mathbf{U}}_0(i)^H$ , where  $\bar{\boldsymbol{\Lambda}}_0(i) = \text{diag}(\bar{\lambda}_0^{\max}(i), \dots, \bar{\lambda}_0^{\min}(i))$  with  $\bar{\lambda}_0^{\max}(i) \geq \dots \geq \bar{\lambda}_0^{\min}(i)$ . Then note that  $\bar{\lambda}_0^{\min}(i) > 0$  w.p.1 since  $\bar{\mathbf{H}}_0(i)$  is full row-rank w.p.1. Then

$$\begin{aligned} Q_0(i) &= \bar{\mathbf{h}}_0(i)^H [\bar{\mathbf{U}}_0(i) (\bar{\boldsymbol{\Lambda}}_0 + \rho^{-1} \mathbf{I}) \bar{\mathbf{U}}_0(i)^H]^{-1} \bar{\mathbf{h}}_0(i) \\ &= \bar{\mathbf{h}}_0(i)^H \bar{\mathbf{U}}_0(i) \text{diag}((\bar{\lambda}_0^{\max}(i) + \rho^{-1})^{-1}, \dots, (\bar{\lambda}_0^{\min}(i) + \rho^{-1})^{-1}) \bar{\mathbf{U}}_0(i)^H \bar{\mathbf{h}}_0(i) \\ &\leq \frac{\|\bar{\mathbf{h}}_0(i)\|^2}{\bar{\lambda}_0^{\min}(i) + \rho^{-1}} \end{aligned} \quad (\text{B.34})$$

$$\leq \frac{\|\bar{\mathbf{h}}_0(i)\|^2}{\bar{\lambda}_0^{\min}(i)}. \quad (\text{B.35})$$

Now  $\mathbb{E}[\bar{\lambda}_0^{\min}(i)] = c_1 > 0$  since  $\bar{\lambda}_0^{\min}(i) > 0$  w.p.1. It is clear that  $Q_0(i)$  is bounded from above by a constant independent of SNR. Consequently,  $R''_{0,i,D_r}$  is also bounded from above by a constant independent of SNR. Then substituting (B.34) in (B.32), dividing both sides with  $\log \rho$  and taking limits as  $\rho \rightarrow \infty$ .

$$\lim_{\rho \rightarrow \infty} \frac{R''_{0,i,D_r}(\rho)}{\log \rho} \leq 0. \quad (\text{B.36})$$

But  $R''_{0,i,D_r}(\rho) \geq 0$  and hence

$$\lim_{\rho \rightarrow \infty} \frac{R''_{0,i,D_r}(\rho)}{\log \rho} \geq 0. \quad (\text{B.37})$$

Clearly, (B.36) and (B.37) imply that

$$\lim_{\rho \rightarrow \infty} \frac{R''_{0,i,D_r}(\rho)}{\log \rho} = 0. \quad (\text{B.38})$$

Next, we assume the best case scenario that  $\forall k \in \{1, \dots, N-1\}$  and  $i \in \mathcal{I}_d$  the rate allocated to the code whose symbol is transmitted on the  $k$ -th subcarrier of the  $i$ -th data multicarrier symbol obeys

$$\lim_{\rho \rightarrow \infty} \frac{R''_{k,i,D_r}(\rho)}{\log \rho} = 1. \quad (\text{B.39})$$

Considering this best case scenario under perfect CSI, combining (B.38) and (B.39), we see that the spectral efficiency of IC-LC-LSP is

$$\lim_{\rho \rightarrow \infty} \frac{R_{D_r}(\rho)}{\log \rho} = \lim_{\rho \rightarrow \infty} \frac{1}{Q_c N_s} \sum_{i \in \mathcal{I}_d} \sum_{k=0}^{N-1} \frac{R''_{k,i,D_r}(\rho)}{\log \rho} \quad (\text{B.40})$$

$$\leq \frac{Q_d(N-1)}{Q_c N_s} \quad (\text{B.41})$$

$$< \frac{Q_d N}{Q_c N_s}. \quad (\text{B.42})$$

This completes the proof.

## B.7 Proof for Theorem 5

For this proof, we consider the availability of (Genie aided) perfect channel state information (CSI) at the receiver. As before, we enforce this by setting the channel estimation error to zero. Recall that in Regime 2, the noise variance  $\sigma^2 \rightarrow 0$  for a fixed finite variance of the non-vanishing interference components. Recalling that

$\Sigma_{\mathbf{w}} > 0$ ,

$$\Sigma_{\mathbf{v}(i)} = \Sigma_{\mathbf{w}} + \Sigma_{\mathbf{w}_{\text{rici}(i)}} + \Sigma_{\mathbf{w}_{\text{isi}(i)}} \quad (\text{B.43})$$

$$\geq \Sigma_{\mathbf{w}_{\text{rici}(i)}} + \Sigma_{\mathbf{w}_{\text{isi}(i)}}. \quad (\text{B.44})$$

Consider Genie-aided global subcarrier processing (GGSP) where  $\mathbf{r}(i)$  is used to decode information on each subcarrier under perfect CSI. Let the rate allocated to the code with codewords transmitted on the  $(N-1)$ -th subcarrier of the  $i$ -th data multicarrier symbol, where  $i \in \mathcal{I}_d$ , be denoted by  $R_{(N-1),i}^{(\text{g})}$ . Note that all dominant ICI has been canceled since all previous subcarriers  $k \in \{0, 1, \dots, N-2\}$  have been decoded and are known to the receiver. Then the post combining SINR can be expressed as

$$\gamma_{N-1}(i)(\mathbf{z}_{N-1}(i)) = \frac{|\mathbf{z}_{N-1}(i)^H \hat{\mathbf{h}}_{N-1,N-1}(i)|^2}{\mathbf{z}_{N-1}(i)^H \Sigma_{\mathbf{v}(i)} \mathbf{z}_{N-1}(i)} \quad (\text{B.45})$$

$$\leq \frac{|\mathbf{z}_{N-1}(i)^H \hat{\mathbf{h}}_{N-1,N-1}(i)|^2}{\mathbf{z}_{N-1}(i)^H (\Sigma_{\mathbf{w}_{\text{rici}(i)}} + \Sigma_{\mathbf{w}_{\text{isi}(i)}}) \mathbf{z}_{N-1}(i)}. \quad (\text{B.46})$$

This shows  $\gamma_{N-1}(i)(\mathbf{z}_{N-1}(i))$  is bounded by a term independent of SNR. Clearly, then, the achievable rate  $R_{N-1,i}^{(\text{g})} = \mathbb{E}_{\hat{\mathbf{H}}(i)} \log(1 + \gamma_{N-1}(i))$  is also bounded by a term independent of SNR. The achievable rate on all other subcarriers is lower, since they are affected by the dominant ICI from subcarriers that are to be decoded in the future. Then  $\{R_{k,i}^{(\text{g})}\}_{k=0}^{N-2}$  are also bounded, and the achievable rate of the surrogate MCM system with GGSP is also bounded in Regime 2. Since the achievable rate does not grow with SNR, the spectral efficiency of the surrogate MCM system in Regime 2 with GGSP is zero. Then the spectral efficiency of LSP-constrained MCM reception with imperfect channel estimates in Regime 2 is also zero. This completes the proof.

## APPENDIX C

### ANALYSIS FOR CHANNEL RE-ESTIMATING RECEPTION

#### C.1 Proof for Lemma 2

We know that  $\text{rank}(\boldsymbol{\Sigma}_{\mathbf{h}}) = N_m$ , so that  $\boldsymbol{\Sigma}_{\mathbf{h}} = \mathbf{B}\boldsymbol{\Lambda}_h\mathbf{B}^H$  for some positive definite diagonal matrix  $\boldsymbol{\Lambda}_h \in \mathbb{C}^{N_m \times N_m}$ . We can then express  $\mathbf{h}(i)$  using the Karhunen-Loeve transform as

$$\mathbf{h}(i) = \mathbf{B}\boldsymbol{\lambda}(i), \quad (\text{C.1})$$

where  $\boldsymbol{\lambda}(i) \in \mathbb{C}^{N_m}$  is a zero-mean complex Gaussian random vector with covariance  $\boldsymbol{\Lambda}_h$ . Furthermore, there exists a unitary matrix  $\mathbf{U}(i)$  and positive semi-definite diagonal matrix  $\boldsymbol{\Lambda}_{N_p}(i)$  such that  $\mathbf{B}^H\mathbf{S}_{N_p}(i)^H\mathbf{S}_{N_p}(i)\mathbf{B} = \mathbf{U}(i)\boldsymbol{\Lambda}_{N_p}(i)\mathbf{U}(i)^H$ . Using  $\boldsymbol{\lambda}'(i) := \mathbf{U}(i)^H\boldsymbol{\lambda}(i)$ , the observations  $\mathbf{y}_{N_p}(i)$  can be expressed as

$$\mathbf{y}_{N_p}(i) = \mathbf{S}_{N_p}(i)\mathbf{h}(i) + \mathbf{w}_{N_p}(i) \quad (\text{C.2})$$

$$= \mathbf{S}_{N_p}(i)\mathbf{B}\boldsymbol{\lambda}(i) + \mathbf{w}_{N_p}(i) \quad (\text{C.3})$$

$$= \boldsymbol{\Lambda}_{N_p}^{\frac{1}{2}}(i)\boldsymbol{\lambda}'(i) + \mathbf{w}_{N_p}(i). \quad (\text{C.4})$$

We first show that the rank condition is a necessary condition. Realize from (C.2)-(C.4) that estimating  $\mathbf{h}(i)$  is equivalent to estimating  $\boldsymbol{\lambda}'(i)$ . Let  $\text{rank}(\mathbf{S}_{N_p}\mathbf{B}) = N'_m <$

$N_m$ . Then w.l.o.g. the first  $N'_m$  entries along the diagonal of  $\mathbf{\Lambda}_{N_p}(i)$  are positive and the rest are zero. Consequently, the MMSE estimates of the last  $N_m - N'_m$  components of  $\mathbf{\lambda}'(i)$  are identically zero and the estimation error for these  $N_m - N'_m$  components of  $\mathbf{\lambda}'(i)$  does not depend on the noise variance  $\sigma^2$ . Then there is no hope of finding a  $\sigma$ -invariant  $\mathbf{A}$  satisfying  $\Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} \leq \sigma^2 \mathbf{A}$  when  $\text{rank}(\mathbf{S}_{N_p} \mathbf{B}) = N'_m < N_m$ . This establishes that the rank condition is a necessary condition.

We now show that the rank condition is a sufficient condition. We first write the estimation error from (6.17) as

$$\begin{aligned}
\Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} &= \Sigma_{\mathbf{h}} - \Sigma_{\mathbf{h}} \mathbf{S}_{N_p}(i)^H \\
&\quad \times (\mathbf{S}_{N_p}(i) \Sigma_{\mathbf{h}} \mathbf{S}_{N_p}(i)^H + \sigma^2 \mathbf{I}_{N_p})^{-1} \mathbf{S}_{N_p}(i) \Sigma_{\mathbf{h}} \\
&= \mathbf{B} \left[ \mathbf{\Lambda}_h - \mathbf{\Lambda}_h \mathbf{B}^H \mathbf{S}_{N_p}(i)^H (\mathbf{S}_{N_p}(i) \mathbf{B} \mathbf{\Lambda}_h \mathbf{B}^H \mathbf{S}_{N_p}(i)^H \right. \\
&\quad \left. + \sigma^2 \mathbf{I}_{N_p})^{-1} \mathbf{S}_{N_p}(i) \mathbf{B} \mathbf{\Lambda}_h \right] \mathbf{B}^H \\
&= \mathbf{B} (\mathbf{\Lambda}_h^{-1} + \sigma^{-2} \mathbf{B}^H \mathbf{S}_{N_p}(i)^H \mathbf{S}_{N_p}(i) \mathbf{B})^{-1} \mathbf{B}^H. \tag{C.5}
\end{aligned}$$

The last step above is an application of the matrix inversion lemma [79]. Realize that  $\mathbf{B}^H \mathbf{S}_{N_p}(i)^H \mathbf{S}_{N_p}(i) \mathbf{B}$  is  $\sigma$ -invariant, positive definite and invertible if  $\text{rank}(\mathbf{S}_{N_p} \mathbf{B}) = N_m$ . We choose  $\mathbf{A}' = (\mathbf{B}^H \mathbf{S}_{N_p}(i)^H \mathbf{S}_{N_p}(i) \mathbf{B})^{-1}$ , and apply the matrix inversion lemma on (C.5) to obtain

$$\begin{aligned}
\Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} &= \\
&\quad \mathbf{B} (\sigma^2 \mathbf{A}' - \sigma^4 \mathbf{A}' (\mathbf{\Lambda}_h + \sigma^2 \mathbf{A}')^{-1} \mathbf{A}') \mathbf{B}^H. \tag{C.6}
\end{aligned}$$

Then for the choice  $\mathbf{A} = \mathbf{B} \mathbf{A}' \mathbf{B}^H$ , (C.6) shows that

$$\Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} - \sigma^2 \mathbf{A} \leq \mathbf{0}. \tag{C.7}$$

This shows that the rank condition is sufficient for  $k = N_p$ . It remains to be shown that the rank condition is sufficient for each  $k > N_p$ . Let  $\check{\mathbf{s}}_k(i)^H$  be the  $k^{\text{th}}$  row of  $\mathbf{S}(i)$ , so that  $\mathbf{S}_{k+1}(i) = [\mathbf{S}_k(i)^H \check{\mathbf{s}}_{k+1}(i)^H]^H$ . Then

$$\begin{aligned}
& \Sigma_{\tilde{\mathbf{h}}^{(k+1)}(i)|\mathbf{s}_{k+1}(i)} \\
&= \mathbf{B} \left( \Lambda_h^{-1} + \sigma^{-2} \mathbf{B}^H \mathbf{S}_{k+1}(i)^H \mathbf{S}_{k+1}(i) \mathbf{B} \right)^{-1} \mathbf{B}^H \\
&= \mathbf{B} \left[ \Lambda_h^{-1} + \sigma^{-2} \mathbf{B}^H \mathbf{S}_k(i)^H \mathbf{S}_k(i) \mathbf{B} \right. \\
&\quad \left. + \sigma^{-2} \mathbf{B}^H \check{\mathbf{s}}_{k+1}(i) \check{\mathbf{s}}_{k+1}(i)^H \mathbf{B} \right]^{-1} \mathbf{B}^H \\
&= \Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \\
&\quad - \sigma^2 \frac{\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \check{\mathbf{s}}_{k+1}(i) \check{\mathbf{s}}_{k+1}(i)^H \Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)}}{1 + \sigma^2 \check{\mathbf{s}}_{k+1}(i)^H \Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \check{\mathbf{s}}_{k+1}(i)}, \tag{C.8}
\end{aligned}$$

where (C.8) results from applying the matrix inversion lemma to the penultimate expression and then substituting (C.5) (with indices  $k$  instead of  $N_p$ ). This clearly implies that

$$\Sigma_{\tilde{\mathbf{h}}^{(k+1)}(i)|\mathbf{s}_{k+1}(i)} - \Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \leq \mathbf{0}. \tag{C.9}$$

We conclude from (C.7) and (C.9) that

$$\Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \leq \sigma^2 \mathbf{A}, \quad N_p \leq k \leq N_s - 1, \tag{C.10}$$

which establishes that the rank criterion is also a sufficient condition and completes the proof.

## C.2 Proof for Theorem 1

We first show that the spectral efficiency of the proposed communication strategy is *at least*  $\frac{N_s - N_m}{N}$  when the rank criterion holds. Recall that

$$R_k(\rho) \geq \mathbb{E} \left\{ \log(1 + \gamma^{(k)}(i)) \right\}. \tag{C.11}$$

Our approach will be to show that the required spectral efficiency can be achieved by using a sub-optimal zero forcing combiner  $\mathbf{c}^{(k)}(i) = \mathbf{e}_N^{(k)}$ , the  $k^{\text{th}}$  column of  $\mathbf{I}_N$ . Recall that we also used this combiner to split transmit power between pilot and data substreams. This choice of combiner implies that only observations  $\{y_k(i)\}_{i=0}^{N_b-1}$  are used in decoding the  $k^{\text{th}}$  data substream. Recall from the system model that  $y_k(i)$  is influenced by symbols  $\{s_m(i)\}_{m=k-N_h+1}^k$ , of which the symbols except  $s_k(i)$  are known (from previously decoded substreams or as pilots). There is no interference from yet-to-be-decoded substreams. It can be shown that the combiner output is

$$z_k(i) = [\hat{\mathbf{h}}_k^{(k)}(i)]_k s_k(i) + n_k(i) \quad (\text{C.12})$$

$$n_k(i) = \check{\mathbf{s}}_k(i)^H \tilde{\mathbf{h}}^{(k)}(i) + w_k(i), \quad (\text{C.13})$$

where  $\check{\mathbf{s}}_k(i)^H$  is the  $k^{\text{th}}$ -row of  $\mathbf{S}(i)$ . First, we bound the variance of the noise term in (C.13). In doing so, we can make use of Lemma 2 since we have assumed that a suitable pilot pattern that satisfies the rank criterion is used. Then we can write

$$\begin{aligned} \sigma_{n_k(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)}^2 &:= \text{E}[|n_k(i)|^2 | \mathbf{s}_k(i), \hat{\mathbf{h}}_k^{(k)}(i)] \text{ as} \\ \sigma_{n_k(i)|\mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)}^2 &= \text{E} \left[ \check{\mathbf{s}}_k(i)^H \Sigma_{\tilde{\mathbf{h}}^{(k)}(i)|\mathbf{s}_k(i)} \check{\mathbf{s}}_k(i) \right] + \sigma^2 \\ &\leq \text{E} \left[ \sigma^2 \check{\mathbf{s}}_k(i)^H \mathbf{A} \check{\mathbf{s}}_k(i) \right] + \sigma^2 \\ &\leq \sigma^2 \alpha_k, \end{aligned} \quad (\text{C.14})$$

for some positive semi-definite  $\sigma$ -invariant matrix  $\mathbf{A}$  and  $\sigma$ -invariant  $\alpha_k > 1$ . Then (C.14) can be used to express the SINR as

$$\gamma^{(k)}(i) \geq \frac{\left| [\hat{\mathbf{h}}_k^{(k)}(i)]_k \right|^2 \sigma_s^2}{\sigma^2 \alpha_k} \quad (\text{C.15})$$

$$= \frac{\left(1 - \frac{E_p}{E_{\text{tot}}}\right)}{(N_s - N_p) \alpha_k} \left| [\hat{\mathbf{h}}_k^{(k)}(i)]_k \right|^2 \rho. \quad (\text{C.16})$$

Then for the class of channels that enable pilots  $\mathbf{S}_{N_m}(i)$  to yield  $\text{rank}(\mathbf{S}_{N_m}(i)\mathbf{B}) = N_m$ , we first show that

$$\liminf_{\rho \rightarrow \infty} \frac{R_k(\rho)}{\log \rho} \geq 1 \quad \forall k \in \{N_p, \dots, N_s - 1\}. \quad (\text{C.17})$$

In this direction, we define  $\psi_k := [\hat{\mathbf{h}}_k^{(k)}(i)]_k$  and  $q_k := \frac{(1-E_p/E_{\text{tot}})}{(N_s-N_p)\alpha_k}$  to simplify the notation. With these definitions, we can say that

$$\frac{R_k(\rho)}{\log \rho} \geq \frac{\text{E} \log(1 + q_k |\psi_k|^2 \rho)}{\log \rho} \quad (\text{C.18})$$

$$\geq 1 + \frac{\text{E} \log(\rho^{-1} + q_k |\psi_k|^2)}{\log \rho}. \quad (\text{C.19})$$

Notice that the estimate  $\psi_k = [\hat{\mathbf{h}}_k^{(k)}(i)]_k$  is zero-mean Gaussian distributed. Then

$$\lim_{\rho \rightarrow \infty} \frac{\log(\rho^{-1} + q_k |\psi_k|^2)}{\log \rho} = \begin{cases} 0, & \psi_k \neq 0 \\ -1, & \psi_k = 0 \end{cases} \quad (\text{C.20})$$

$$= 0 \quad \text{w.p.1.} \quad (\text{C.21})$$

Taking limit infimum on both sides of (C.19), we see that

$$\begin{aligned} \liminf_{\rho \rightarrow \infty} \frac{R_k(\rho)}{\log \rho} &\geq 1 + \liminf_{\rho \rightarrow \infty} \frac{\text{E} \log(\rho^{-1} + q_k |\psi_k|^2)}{\log \rho} \\ &\geq 1 + \text{E} \left( \liminf_{\rho \rightarrow \infty} \frac{\log(\rho^{-1} + q_k |\psi_k|^2)}{\log \rho} \right) \\ &\geq 1. \end{aligned} \quad (\text{C.22})$$

In the above, the penultimate step is an application of Fatou's Lemma [78] and the last step applies (C.21). Using (C.22) and the fact that

$$R_{\text{tot}}(\rho) = \frac{1}{N} \sum_{k=N_m}^{N_s-1} R_k(\rho), \quad (\text{C.23})$$

we obtain

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} \geq \frac{N_s - N_m}{N}. \quad (\text{C.24})$$

Thus the proposed communication strategy attains a spectral efficiency of at least (C.24).

On the other hand, consider that perfect CSI is available at the receiver through a genie. In this situation, well known results in [51, 76] dictate that the spectral efficiency of a communication strategy that transmits  $(N_s - N_m)$  data substreams over  $N$  channel-uses cannot exceed  $\frac{N_s - N_m}{N}$  even with optimal joint decoding. The proposed communication strategy has poorer performance than the genie aided strategy since it uses imperfect CSI, and can only have a poorer spectral efficiency. This observation leads us to conclude that

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} = \frac{N_s - N_m}{N}. \quad (\text{C.25})$$

This completes the proof.

### C.3 Proof for Proposition 1

We need to demonstrate that  $\text{rank}(\mathbf{S}_{N_p}(i)\mathbf{B}) = N_m = N_h(2D + 1)$ . Recall that the pilot pattern used is

$$s_k(i) = \sqrt{\frac{N_h E_p}{N_m}} \delta_{\langle k \rangle_{N_h}}, \quad 0 \leq k < N_m. \quad (\text{C.26})$$

Recalling the structure of  $\mathbf{S}(i)$  and the fact that, for the CE-BEM channel,  $\mathbf{B} = \mathbf{I}_{N_h} \otimes \mathbf{F}$ , where the  $N_s \times (2D + 1)$  matrix  $\mathbf{F}$  is defined element-wise as  $[\mathbf{F}]_{m_1, m_2} = \frac{1}{\sqrt{N_s}} e^{j \frac{2\pi}{N_s} m_1 (m_2 - D)}$ . Under these conditions, it is straightforward to show that

$$\mathbf{S}_{N_p}(i)\mathbf{B} = \sqrt{\frac{N_h E_p}{N_m}} \mathbf{P} (\mathbf{I}_{N_h} \otimes \mathbf{M}), \quad (\text{C.27})$$

In (C.27), the  $(2D + 1)N_h \times (2D + 1)N_h$  row-permutation matrix  $\mathbf{P}$  is defined element-wise by  $[\mathbf{P}]_{m_1, m_2(2D+1)+m_3} = [\mathbf{I}_{(2D+1)N_h}]_{m_1, m_3 N_h + m_2}$  where  $0 \leq m_1 < (2D + 1)N_h$ ,

$0 \leq m_2 < N_h$  and  $0 \leq m_3 < (2D + 1)$ . Furthermore, the  $(2D + 1) \times (2D + 1)$  complex matrix  $\mathbf{M}$  is defined element-wise as  $[\mathbf{M}]_{m_1, m_2} = \frac{1}{\sqrt{N_s}} \exp(\frac{j2\pi N_h m_1 m_2}{N_s})$ . Then,

$$\text{rank}(\mathbf{S}_{N_p}(i)\mathbf{B}) = \text{rank}(\mathbf{P}(\mathbf{I}_{N_h} \otimes \mathbf{M})) \quad (\text{C.28})$$

$$= \text{rank}(\mathbf{I}_{N_h} \otimes \mathbf{M}) \quad (\text{C.29})$$

$$= N_h \text{rank}(\mathbf{M}) \quad (\text{C.30})$$

$$= N_h(2D + 1) = N_m. \quad (\text{C.31})$$

In the above, (C.28) is obtained using (C.27), (C.29) is a result of  $\mathbf{P}$  being, by definition, a permutation of the columns of  $\mathbf{I}_{(2D+1)N_h}$ , (C.30) is a standard result for block-diagonal matrices and the final step is a result of  $\mathbf{M}$  being a full rank Vandermonde matrix.

Then applying Lemma 2 and Theorem 6 we see that when the proposed communication strategy is used for single carrier transmission over doubly selective fading channels,

$$\lim_{\rho \rightarrow \infty} \frac{R_{\text{tot}}(\rho)}{\log \rho} = \frac{N_s - N_m}{N}. \quad (\text{C.32})$$

This concludes the proof.

## C.4 Proof for Proposition 2

Recall that the post-combining SINR for the  $N_p^{\text{th}}$  substream is

$$\gamma^{(N_p)}(i) = \frac{|\mathbf{c}^{(N_p)}(i)^H \hat{\mathbf{h}}_{N_p}^{(N_p)}(i)|^2 \sigma_s^2}{\mathbf{c}^{(N_p)}(i)^H \boldsymbol{\Sigma}_{\mathbf{v}^{(N_p)}(i) | \mathbf{s}_k(i), \hat{\mathbf{h}}^{(k)}(i)} \mathbf{c}^{(N_p)}(i)}. \quad (\text{C.33})$$

For the sub-optimal combiner  $\mathbf{c}^{(N_p)}(i) = \mathbf{e}_N^{(N_p)}$ , the  $N_p^{\text{th}}$  column of  $\mathbf{I}_N$ , only observations  $\{y_{N_p}(i)\}_{i=0}^{N_b-1}$  are used to decode the  $N_p^{\text{th}}$  substream. Recall from the system

model that  $y_{N_p}(i)$  is influenced by symbols  $\{s_m(i)\}_{m=N_p-N_h+1}^{N_p}$ , of which, all the symbols except  $s_{N_p}(i)$  are known as pilots. For the pilot pattern used (6.38),  $s_k(i) = 0$  for  $N_p - N_h + 1 \leq k \leq N_p - 1$ . As a result, only the estimation error from the estimate of  $h_{N_p,0}(i)$  affects  $y_{N_p}(i)$ . Under these circumstances, it is straightforward to show that

$$\gamma^{(N_p)}(i) \geq \frac{\hat{\sigma}_{N_p}^2 \sigma_s^2 \zeta}{\tilde{\sigma}_{N_p}^2 \sigma_s^2 + \sigma^2}. \quad (\text{C.34})$$

In (C.34),  $\zeta$  is a zero-mean complex Gaussian random variable with unit variance,  $\hat{\sigma}_{N_p}^2$  is the variance of the estimate of  $h_{N_p,0}(i)$ , and  $\tilde{\sigma}_{N_p}^2$  the variance of the corresponding estimation error. The variances  $\hat{\sigma}_{N_p}^2$  and  $\tilde{\sigma}_{N_p}^2$  can be calculated as per their definitions in (6.41) and (6.42). In doing so, the covariance matrices of the estimates used to decode the  $N_p^{\text{th}}$  substream and the corresponding estimation error for the pilot pattern in (6.38) and the CE-BEM channel is given by

$$\Sigma_{\hat{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} = \Sigma_{\mathbf{h}(i)} - \Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} \quad (\text{C.35})$$

$$\begin{aligned} \Sigma_{\tilde{\mathbf{h}}^{(N_p)}(i)|\mathbf{s}_{N_p}(i)} &= \mathbf{B} \left[ \frac{(2D+1)N_h}{N_s} \mathbf{I}_{N_m} + \frac{\alpha_p E_{\text{tot}}}{2D+1} \right. \\ &\quad \left. \times \mathbf{I}_{N_h} \otimes (\mathbf{M}^H \mathbf{M}) \right]^{-1} \mathbf{B}^H, \end{aligned} \quad (\text{C.36})$$

where the  $(2D+1) \times (2D+1)$  matrix  $\mathbf{M}$  is defined element-wise as  $[\mathbf{M}]_{m_1, m_2} = \frac{1}{\sqrt{N_s}} \exp(j2\pi \frac{N_h m_1 m_2}{N_s})$ . (See Appendix C.3 for details.) The choice  $\alpha_{p,*}$  that maximizes this lower-bound on the post-combining SINR and consequently the lower-bound on the achievable rate is

$$\alpha_{p,*} = \arg \max_{\alpha_p \in (0,1)} \frac{\hat{\sigma}_{N_p}^2 \sigma_s^2}{\tilde{\sigma}_{N_p}^2 \sigma_s^2 + \sigma^2}. \quad (\text{C.37})$$

This completes the proof.

## BIBLIOGRAPHY

- [1] G. L. Stüber, *Principles of Mobile Communication*. Kluwer Academic Publishers, 2nd ed., 2001.
- [2] M. Russell and G. L. Stüber, “Interchannel interference analysis of OFDM in a mobile environment,” in *Proc. IEEE Veh. Tech. Conf.*, vol. 2, pp. 820–824, 1995.
- [3] J. Li and M. Kavehrad, “Effects of time selective fading on OFDM for broadband mobile applications,” *IEEE Commun. Letters*, vol. 3, pp. 332–334, Dec. 1999.
- [4] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, “Intercarrier interference in MIMO OFDM,” *IEEE Trans. Signal Processing*, vol. 50, pp. 2451–2464, Oct. 2002.
- [5] K. Liu, T. Kadous, and A. M. Sayeed, “Orthogonal time-frequency signaling over doubly dispersive channels,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 2583–2603, Nov. 2004.
- [6] P. A. Bello, “Characterization of randomly time-variant linear channels,” *IEEE Trans. Commun. Syst.*, vol. CS-11, pp. 360–393, Dec. 1963.
- [7] M. K. Tsatsanis and G. B. Giannakis, “Modeling and equalization of rapidly fading channels,” *International Journal of Adaptive Control and Signal Processing*, vol. 10, pp. 159–176, Mar. 1996.
- [8] X. Ma and G. B. Giannakis, “Maximum-diversity transmissions over doubly-selective wireless channels,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 1832–1840, July 2003.
- [9] Y. Liang and V. Veeravalli, “Capacity of noncoherent time-selective Rayleigh-fading channels,” *IEEE Trans. Inform. Theory*, vol. 50, p. 30953110, Dec. 2004.
- [10] B. Le Floch, M. Alard, and C. Berrou, “Coded orthogonal frequency division multiplex,” *Proc. IEEE*, vol. 83, pp. 982–996, June 1995.
- [11] R. Haas and J.-C. Belfiore, “A time-frequency well-localized pulse for multiple carrier transmission,” *Wireless Personal Commun.*, vol. 5, pp. 1–18, 1997.

- [12] W. Kozek and A. F. Molisch, “Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels,” *IEEE J. Select. Areas In Commun.*, vol. 16, pp. 1579–1589, Oct. 1998.
- [13] D. Schaufhuber, G. Matz, and F. Hlawatsch, “Pulse-shaping OFDM/BFDM systems for time-varying channels: ISI/ICI analysis, optimal pulse design, and efficient implementation,” in *Proc. IEEE Symp. PIMRC*, pp. 1012–1016, Sep. 2002.
- [14] H. Bölcskei, “Orthogonal frequency division multiplexing based on offset QAM,” in *Advances in Gabor Theory* (H. G. Feichtinger and T. Strohmer, eds.), Birkhäuser, 2002.
- [15] T. Strohmer and S. Beaver, “Optimal OFDM design for time-frequency dispersive channels,” *IEEE Trans. Commun.*, vol. 51, pp. 1111–1122, July 2003.
- [16] K. Matheus and K.-D. Kammeyer, “Optimal design of a multicarrier systems with soft impulse shaping including equalization in time or frequency direction,” in *IEEE Globecom*, vol. 1, pp. 310–314, 1997.
- [17] K. Matheus, K. Knoche, M. Feuersänger, and K.-D. Kammeyer, “Two-dimensional (recursive) channel equalization for multicarrier systems with soft impulse shaping (MCSIS),” in *Proc. IEEE Global Telecommunications Conf.*, vol. 2, pp. 956–961, 1998.
- [18] A. Vahlin and N. Holte, “Optimal finite duration pulses for OFDM,” *IEEE Trans. Commun.*, vol. 44, pp. 10–14, Jan. 1996.
- [19] S. Slimane, “Peak-to-average power ratio reduction of OFDM signals using broadband pulse shaping,” in *Proc. IEEE Veh. Tech. Conf.*, 2002.
- [20] P. Schniter, “Low-complexity equalization of OFDM in doubly-selective channels,” *IEEE Trans. Signal Processing*, vol. 52, pp. 1002–1011, Apr. 2004.
- [21] T. Hunziker and D. Dahlhaus, “Iterative detection for multicarrier transmission employing time-frequency concentrated pulses,” *IEEE Trans. Commun.*, vol. 51, pp. 641–651, Apr. 2003.
- [22] S. Das and P. Schniter, “A new pulse shaped frequency division multiplexing technique for doubly dispersive channels,” in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2004.
- [23] S. Das and P. Schniter, “Max-SINR ISI/ICI-shaped multi-carrier communication over the doubly dispersive channel,” *IEEE Trans. Signal Processing*, 2007. Accepted for publication.

- [24] Z. Tang and G. Leus, “Low-complexity equalization for maximum diversity transmissions over doubly-selective channels,” *IEEE Trans. Signal Processing*, vol. 54, pp. 3642–3648, Sept. 2006.
- [25] S.-J. Hwang and P. Schniter, “Efficient sequence detection for multicarrier transmissions over doubly dispersive channels,” *EURASIP J. on Applied Sig. Proc.*, 2006. Special issue on reliable communications over rapidly time-varying channels. Article ID: 93638.
- [26] W. G. Jeon, K. H. Chang, and Y. S. Cho, “An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels,” *IEEE Trans. Commun.*, vol. 47, pp. 27–32, Jan. 1999.
- [27] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, “On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels,” *IEEE Trans. Commun.*, vol. 49, pp. 1375–1387, Aug. 2001.
- [28] X. Cai and G. B. Giannakis, “Bounding performance and suppressing inter-carrier interference in wireless mobile OFDM,” *IEEE Trans. Commun.*, vol. 51, pp. 2047–2056, Dec. 2003.
- [29] L. Rugini, P. Banelli, and G. Leus, “Simple equalization of time-varying channels for ofdm,” *IEEE Commun. Letters*, vol. 9, pp. 619–621, July 2005.
- [30] L. Rugini, P. Banelli, and G. Leus, “Block DFE and windowing for doppler-affected OFDM systems,” in *Proc. IEEE Workshop Signal Processing Advances in Wireless Commun.*, 2005.
- [31] P. Robertson and S. Kaiser, “The effects of Doppler spreads on OFDM(A) mobile radio systems,” in *Proc. IEEE Veh. Tech. Conf.*, vol. 1, pp. 329–333, 1999.
- [32] B. Stantchev and G. Fettweis, “Time-variant distortions in OFDM,” *IEEE Commun. Letters*, vol. 4, pp. 312–314, Oct. 2000.
- [33] Y. Li and L. J. Cimini, Jr., “Bounds on the interchannel interference of OFDM in time-varying impairments,” *IEEE Trans. Commun.*, vol. 49, pp. 401–404, Mar. 2001.
- [34] A. M. Tonello, “Performance limits for filtered multitone modulation in fading channels,” *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2121 – 2135, Sept. 2005.
- [35] L. Tong, B. Sadler, and M. Dong, “Pilot assisted wireless transmissions: General model, design criteria and signal processing,” *IEEE Signal Processing Mag.*, vol. 21, pp. 12–25, Nov 2004.

- [36] X. Ma, G. B. Giannakis, and S. Ohno, “Optimal training for block transmissions over doubly-selective wireless fading channels,” *IEEE Trans. Signal Processing*, vol. 51, pp. 1351–1366, May 2003.
- [37] L. Yang, X. Ma, and G. B. Giannakis, “Optimal training for mimo fading channels with time- and frequency- selectivity,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. 821–824, May 2004.
- [38] A. P. Kannu and P. Schniter, “Reduced-complexity decision-directed pilot-aided tracking of doubly-selective channels,” in *Proc. Conf. Inform. Science and Systems*, 2004.
- [39] P. Schniter, “On doubly dispersive channel estimation for pilot-aided pulse-shaped multicarrier modulation,” in *Proc. Conf. Inform. Science and Systems*, (Princeton, NJ), Mar. 2006.
- [40] M. Medard, “The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 933–946, May 2000.
- [41] L. Zheng and D. N. C. Tse, “Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.
- [42] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, April 2003.
- [43] S. Misra, A. Swami, and L. Tong, “Cutoff rate analysis of the Gauss-Markov fading channel with adaptive energy allocation,” in *Proc. IEEE Workshop Signal Processing Advances in Wireless Commun.*, 2003.
- [44] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, “On the capacity of frequency-selective channels in training-based transmission schemes,” *IEEE Trans. Signal Processing*, vol. 52, pp. 2572–2583, Sept 2004.
- [45] X. Ma, L. Yang, and G. B. Giannakis, “Optimal training for MIMO frequency-selective fading channels,” *IEEE Trans. Wireless Commun.*, vol. 4, p. 453466, Mar 2005.
- [46] A. P. Kannu and P. Schniter, “Capacity analysis of MMSE pilot patterns for doubly selective channels,” in *Proc. IEEE Workshop Signal Processing Advances in Wireless Commun.*, June 2005.

- [47] A. P. Kannu and P. Schniter, "On the spectral efficiency of noncoherent doubly selective channels," in *Proc. Allerton Conf. Commun., Control, and Computing*, Sept 2006.
- [48] A. P. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels," *IEEE Trans. Signal Processing*, 2006. Submitted.
- [49] R. H. Etkin and D. N. C. Tse, "Degrees of freedom in some underspread MIMO fading channels," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1576–1608, April 2006.
- [50] S. Das and P. Schniter, "Local subcarrier processing for SIMO multicarrier systems over doubly selective channels," *IEEE Trans. Signal Processing*, 2007. In preparation.
- [51] E. Telatar, "Capacity of multi-antenna gaussian channels," *Bell Labs. Tech. Rep.*, 1995.
- [52] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, 1992.
- [53] P. Schniter, "A new approach to multicarrier pulse design for doubly-dispersive channels," in *Proc. Allerton Conf. Commun., Control, and Computing*, Oct. 2003.
- [54] G. Leus, S. Zhou, and G. B. Giannakis, "Orthogonal multiple access over time- and frequency-selective channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1942–1950, August 2003.
- [55] S. Das and P. Schniter, "Design of multi-carrier modulation for doubly selective channels based on a complexity constrained achievable rate metric," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2006.
- [56] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar 1998.
- [57] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
- [58] T. K. Y. Lo, "Maximum rasion transmission," *IEEE Trans. Commun.*, vol. 47, pp. 1458–1461, Oct 1999.

- [59] H. El Gamal, G. Caire, and M. O. Damen, “Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels,” *IEEE Trans. Inform. Theory*, vol. 30, pp. 968–985, June 2004.
- [60] D. J. Love and R. W. Heath, Jr., “Equal gain transmission in multiple-input multiple-output wireless systems,” *IEEE Trans. Commun.*, vol. 51, pp. 1102–1110, Jul 2003.
- [61] H. Bölcskei, D. Gesbert, and A. J. Paulraj, “On the capacity of OFDM based spatial multiplexing systems,” *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.
- [62] H. Bölcskei, M. Borgmann, and A. J. Paulraj, “Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM,” *IEEE J. Select. Areas In Commun.*, vol. 21, pp. 427–439, Apr. 2002.
- [63] G. C. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE J. Select. Areas In Commun.*, vol. 46, pp. 357–366, March 1998.
- [64] S. Das and P. Schniter, “Beamforming and combining strategies for MIMO-OFDM over doubly selective channels,” in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2006.
- [65] C. Kominakis and R. D. Wesel, “Joint iterative channel estimation and decoding in flat correlated Rayleigh fading,” *IEEE J. Select. Areas In Commun.*, vol. 19, pp. 1706–1717, Sept. 2001.
- [66] M. C. Valenti and B. D. Woerner, “Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels,” *IEEE J. Select. Areas In Commun.*, vol. 19, pp. 1697–1705, Sept. 2001.
- [67] K. Fu and A. Anastasopoulos, “Analysis and design of LDPC codes for time-selective complex-fading channels,” *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1175–1185, May 2005.
- [68] A. Krishnamoorthy and A. Anastasopoulos, “Code and receiver design for the noncoherent fast-fading channel,” *IEEE J. Select. Areas In Commun.*, vol. 23, pp. 1769–1778, Sept. 2005.
- [69] T. Li and O. M. Collins, “A successive decoding strategy for channels with memory,” *IEEE Trans. Inform. Theory*, vol. 53, pp. 628–646, Feb. 2007.
- [70] S. Das and P. Schniter, “Noncoherent communication over the doubly selective channel via successive decoding and channel re-estimation,” in *Proc. Allerton Conf. Commun., Control, and Computing*, Oct. 2007.

- [71] A. P. Kannu and P. Schniter, “On the spectral efficiency of noncoherent doubly selective block-fading channels,” *IEEE Trans. Inform. Theory*, 2007. Submitted.
- [72] J. A. C. Bingham, “Multicarrier modulation for data transmission: An idea whose time has come,” *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
- [73] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.
- [74] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [75] T. A. Thomas, T. P. Krauss, and F. W. Vook, “CHAMPS: A near-ML joint doppler frequency / TOA search for channel characterization,” in *Proc. IEEE Veh. Tech. Conf.*, Fall 2003.
- [76] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [77] G. D. Forney, Jr., “On the role of MMSE estimation in approaching the information-theoretic limits of linear gaussian channels: Shannon meets Wiener,” in *Proc. Allerton Conf. Commun., Control, and Computing*, 2003.
- [78] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 3rd ed., 1976.
- [79] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY: Cambridge University Press, 1985.