

Accounting for channel constraints in joint source-channel video coding schemes

Shuo Zheng

▶ **To cite this version:**

Shuo Zheng. Accounting for channel constraints in joint source-channel video coding schemes. Signal and Image Processing. Université Paris Saclay (COmUE), 2019. English. NNT : 2019SACLT005 . tel-02050971

HAL Id: tel-02050971
https://pastel.hal.science/tel-02050971

Submitted on 27 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

TELECOM ParisTech

NNT : 2019SACLT005

Thèse de doctorat

Prise en compte des contraintes de canal dans les schémas de codage vidéo conjoint du source-canal

Thèse de doctorat de l'Université Paris-Saclay
préparée à Télécom ParisTech

Ecole doctorale n°580 Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat : Traitement du Signal et des images

Thèse présentée et soutenue à Paris, le 5 février 2019, par

SHUO ZHENG

Composition du Jury :

Michèle WIGGER Professeur, Télécom ParisTech	Président
François-Xavier COUDOUX Professeur, Université Polytechnique Hauts-de-France	Rapporteur
Aline ROUMY Chargé de Recherche, INRIA Rennes	Rapporteur
Marc LENY Directeur de recherche, Ektacom	Examinateur
Jean-Marie GORCE Professeur, INSA Lyon	Examinateur
Marco CAGNAZZO Professeur,Télécom Paristech	Directeur de thèse
Michel KIEFFER Professeur, Université de Paris-Sud	Co-directeur de thèse

Accounting for Channel Constraints in Joint Source-Channel Video Coding Schemes

February 11, 2019

Résumé

Les schémas de Codage Vidéo Linéaire (CVL) inspirés de SoftCast ont émergé dans la dernière décennie comme une alternative aux schémas de codage vidéo classiques. Ces schémas de codage source-canal conjoint exploitent des résultats théoriques montrant qu'une transmission (quasi-) analogique est plus performante dans des situations de multicast que des schémas numériques lorsque les rapports signal-à-bruit des canaux (C-SNR) diffèrent d'un récepteur à l'autre. Dans ce contexte, les schémas de CVL permettent d'obtenir une qualité de vidéo décodée proportionnelle au C-SNR du récepteur.

Une première contribution de cette thèse concerne l'optimisation de la matrice de précodage de canal pour une transmission de type OFDM de flux générés par un CVL lorsque les contraintes de puissance diffèrent d'un sous-canal à l'autre. Ce type de contrainte apparait en sur des canaux DSL, ou dans des dispositifs de transmission sur courant porteur en ligne (CPL). Cette thèse propose une solution optimale à ce problème de type multi-level water filling et nécessitant la solution d'un problème de type Structured Hermitian Inverse Eigenvalue. Trois algorithmes sous-optimaux de complexité réduite sont également proposés. Des nombreux résultats de simulation montrent que les algorithmes sous-optimaux ont des performances très proches de l'optimum et réduisent significativement le temps de codage. Le calcul de la matrice de précodage dans une situation de multicast est également abordé.

Une seconde contribution principale consiste en la réduction de l'impact du bruit impulsif dans les CVL. Le problème de correction du bruit impulsif est formulé comme un problème d'estimation d'un vecteur creux. Un algorithme de type Fast Bayesian Matching Pursuit (FBMP) est adapté au contexte CVL. Cette approche nécessite de réserver des sous-canaux pour la correction du bruit impulsif, entrainant une diminution de la qualité vidéo en l'absence de bruit impulsif. Un modèle phénoménologique (MP) est proposé pour décrire l'erreur résiduelle après correction du bruit impulsif. Ce modèle permet de d'optimiser le nombre de sous-canaux

à réserver en fonction des caractéristiques du bruit impulsif. Les résultats de simulation montrent que le schéma proposé améliore considérablement les performances lorsque le flux CVL est transmis sur un canal sujet à du bruit impulsif.

Abstract

SoftCast based Linear Video Coding (LVC) schemes have been emerged in the last decade as a quasi analog joint-source-channel alternative to classical video coding schemes. Theoretical analyses have shown that analog coding is better than digital coding in a multicast scenario when the channel signal-to-noise ratios (C-SNR) differ among receivers. LVC schemes provide in such context a decoded video quality at different receivers proportional to their C-SNR.

This thesis considers first the channel precoding and decoding matrix design problem for LVC schemes under a per-subchannel power constraint. Such constraint is found, e.g., on Power Line Telecommunication (PLT) channels and is similar to per-antenna power constraints in multi-antenna transmission system. An optimal design approach is proposed, involving a multi-level water filling algorithm and the solution of a structured Hermitian Inverse Eigenvalue problem. Three lower-complexity alternative suboptimal algorithms are also proposed. Extensive experiments show that the suboptimal algorithms perform closely to the optimal one and can reduce significantly the complexity. The precoding matrix design in multicast situations also has been considered.

A second main contribution consists in an impulse noise mitigation approach for LVC schemes. Impulse noise identification and correction can be formulated as a sparse vector recovery problem. A Fast Bayesian Matching Pursuit (FBMP) algorithm is adapted to LVC schemes. Subchannels provisioning for impulse noise mitigation is necessary, leading to a nominal video quality decrease in absence of impulse noise. A phenomenological model (PM) is proposed to describe the impulse noise correction residual. Using the PM model, an algorithm to evaluate the optimal number of subchannels to provision is proposed. Simulation results show that the proposed algorithms significantly improve the video quality when transmitted over channels prone to impulse noise.

Contents

1 Introduction **8**

1.1 Context and motivation 8

1.2 Contributions 12

1.3 Organization of thesis 15

2 Related work and prior results **17**

2.1 SoftCast: A joint source-channel coding scheme 17

2.1.1 Information-theoretic foundations of SoftCast 17

2.1.2 SoftCast 24

2.2 Improvements of SoftCast 27

2.2.1 Dcast 27

2.2.2 WaveCast 30

2.2.3 WSVC 30

2.2.4 Energy distribution Modeling 33

2.2.5 ParCast+ 36

2.2.6 Application of Shannon-Kotel'nikov Mapping In LVC 37

2.2.7 Conclusion 38

3 Optimal Power Allocation **40**

3.1 Introduction 40

3.2 Precoding and decoding matrices 41

3.3 Total Power Constraint 43

3.3.1 Optimal decoding matrix 44

3.3.2 Optimal precoding matrix 44

3.3.3 A toy example . . . 50
3.4 Per Subchannel Power Constraints . . . 50
3.4.1 Evaluation of $\widetilde{m}$ . . . 53
3.4.2 When the conditions of Theorem 2 are satisfied . . . 54
3.4.3 When the conditions of Theorem 2 are not satisfied . . . 54
3.5 Precoding Matrix Design in multicast scenarii . . . 57
3.5.1 Multicast scenario with linearly degraded multicast channels . . . 58
3.5.1.1 Total Power Constraint . . . 58
3.5.1.2 Per-subchannel power constraint . . . 61
3.5.2 General multicast channels . . . 63
3.5.2.1 Eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ larger than one . . . 63
3.5.2.2 Eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are small . . . 64
3.6 Simulations . . . 64
3.6.1 Simulation conditions . . . 64
3.6.2 Metadata . . . 68
3.6.3 Total Power Constraint . . . 69
3.6.4 Per Subchannel Power Constraints . . . 71
3.6.5 Mismatch . . . 72
3.6.5.1 Total Power Constraint . . . 74
3.6.5.2 Per-subchannel Power Constraint . . . 75
3.7 Conclusions . . . 76

4 Sub-Optimal Power Allocation **79**

4.1 Simple Chunk Scaling . . . 79
4.2 Power Allocation with Inferred Split Position (PAISP) . . . 80
4.3 PAISP with Dichotomy . . . 84
4.4 Power Allocation with Local Power Adjustment . . . 85
4.5 Limits of PAISP and PALPA . . . 89
4.6 Simulation results . . . 91
4.6.1 Comparison of the power allocation methods . . . 91
4.6.2 Complexity comparison . . . 97

4.6.3 Mismatch . . . 99
4.7 Conclusions . . . 101

5 Impulse error mitigation for LVC schemes 103

5.1 Introduction and main contributions . . . 103
5.2 Related work . . . 104
5.3 Linear Video Coding and OFDM Transmission Scheme . . . 106
5.3.1 Joint source-channel coding . . . 106
5.3.2 Transmission . . . 107
5.3.3 Channel model . . . 108
5.3.4 Baseline receiver . . . 109
5.3.5 Power allocation and chunk selection . . . 109
5.4 Impulse Noise Correction . . . 110
5.5 Sub-channel provisioning for impulse noise mitigation . . . 112
5.5.1 Residual noise after impulse noise mitigation . . . 112
5.5.2 Estimation of σ_{r}^2 . . . 116
5.5.3 Optimization of sub-channel provisioning . . . 120
5.6 Simulation . . . 122
5.6.1 Compared LVC schemes . . . 124
5.6.2 Simulation parameters . . . 125
5.6.3 Simulation results . . . 126
5.6.3.1 Impact of r_{d} on the efficiency of impulse noise correction . . . 126
5.6.3.2 Optimal subchannel provisioning . . . 129
5.6.3.3 Analysis of the effect of mismatched channel conditions . . . 129
5.7 Conclusion . . . 131

6 Conclusions and Perspectives 133

6.1 Conclusions . . . 133
6.2 Perspectives . . . 134
6.2.1 Precoding matrix design for multicast . . . 134
6.2.2 Optimization of the amount of Metadata . . . 136

6.2.3 Application of Deep Learning to SoftCast schemes 136

Chapter 1

Introduction

1.1 Context and motivation

In the last decades, a huge research effort has been devoted to design video coding and transmission systems to get the best received video quality for a given amount of channel resources. This is of paramount importance for cellular broadcasting, where channel conditions may be varying with time and among receivers, for multimedia transmission in wireless networks, but is also for communication over wired channels, such as Digital Subscriber Line (DSL) or Power Line Telecommunication (PLT) channels. The first case is illustrated in Figure 1.1, where a multimedia transmission is performed via DVB-T (Digital Video Broadcasting - Terrestrial) to different users experiencing channels with different characteristics. The second case is depicted in Figure 1.2, where a multimedia server transmits data along the power line using HomePlug AV2 [YAA+13], to receivers which are at different locations.

Traditional solutions for these use-cases consist in using a non-scalable or a scalable video encoder [SMW07]. Nevertheless, the source encoding is performed without knowing the actual channel characteristics, and this may cause channel underused or digital cliff. This problem is illustrated in Figure 1.3. A single layer MPEG4 codec selects a bit rate for video compression equal to the channel transmission rate, depending on the modulation and channel FEC (Forward Error Correction) scheme. Assume that the chosen modulation and coding scheme (MCS) is 16 QAM and rate 1/2 convolutionnal code. If the channel quality (C-SNR) is less (for example at 10 dB) than the C-SNR for which the MCS is adapted, a cliff effect will appear. On the other hand, if the channel quality is improved, for example C-SNR at 15dB, but the bit rate

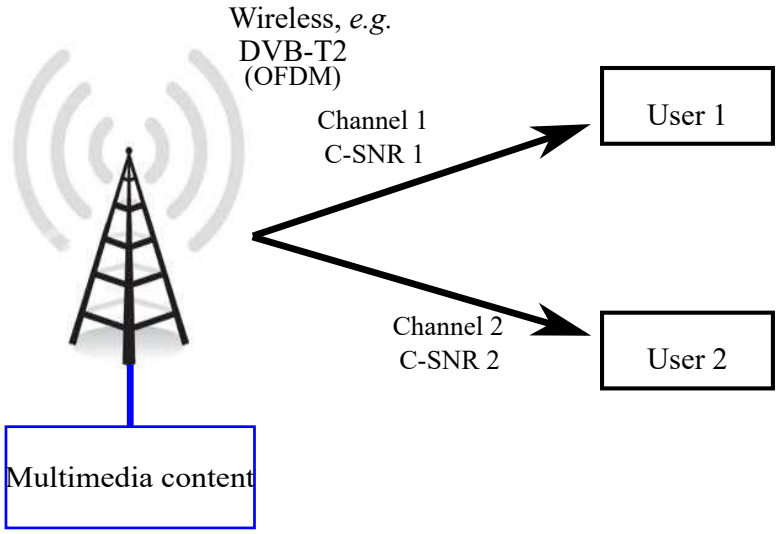


Figure 1.1: Multimedia broadcast to different users with different channel characteristics.

does not increase accordingly, then there is saturation problem. Hence in broadcast, choosing the bit rate that fits the receivers with the worst channel penalizes users with better channel conditions. Even though scalable coding facilitates transmission rate adaptation compared to a non-scalable scheme (such as H.264/AVC or HEVC [WSBL03, SOHW12]) at similar coding rate, the global coding efficiency of scalable schemes decreases with the number of scalability layers [WSO07].

Joint source-channel video coding (JSCVC) has the potential of dramatically improving the quality of the received video in such challenging conditions, as demonstrated by the breaking-through SoftCast video coding and transmission system [JK10a]. SoftCast is a JSCVC scheme that encodes the video content with linear-only operators (such as a full-frame DCT and scaling). For this reason, SoftCast-inspired schemes maybe referred to as Linear Video Coding (LVC) schemes. The original pixels are transformed into DCT coefficients, which are then grouped into chunks according to their variance. Chunks are then scaled and sent on the channel with an extremely dense modulation. In SoftCast, compression involves a full-GoP 3D-DCT and selection of transformed coefficients. Error protection is obtained by power allocation and resilience to packet losses is obtained by giving up temporal prediction and on the contrary ensuring that all packets contribute equally to the quality of the decoded video. Since in SoftCast, the transmitted symbols are linearly related to the original pixel values, the video quality at receiver scales linearly with the channel signal-noise-ratio (SNR). Therefore the cliff effect in broadcast is avoided. Detailed comparisons with H.264/AVC or SVC over 802.11 wireless

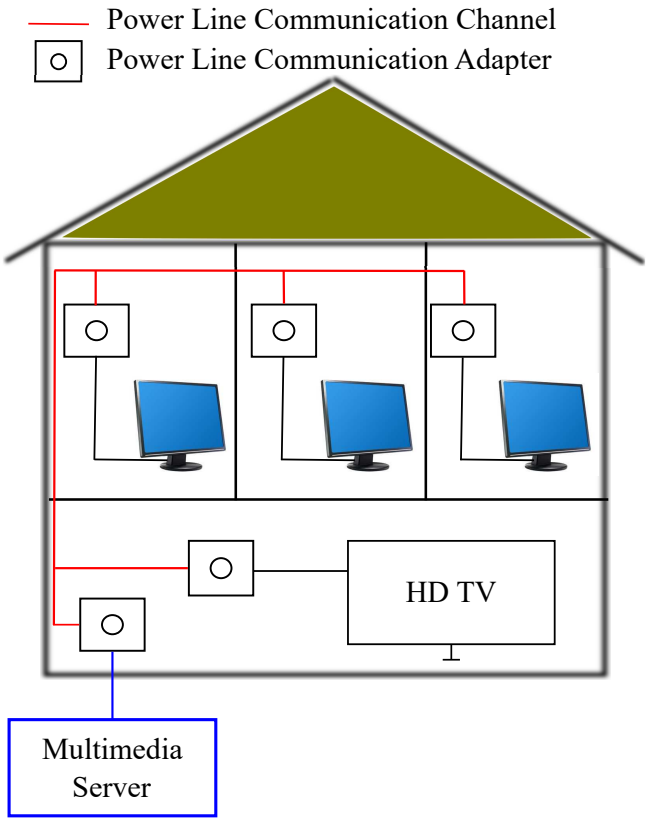


Figure 1.2: An illustration of HomePlug AV2

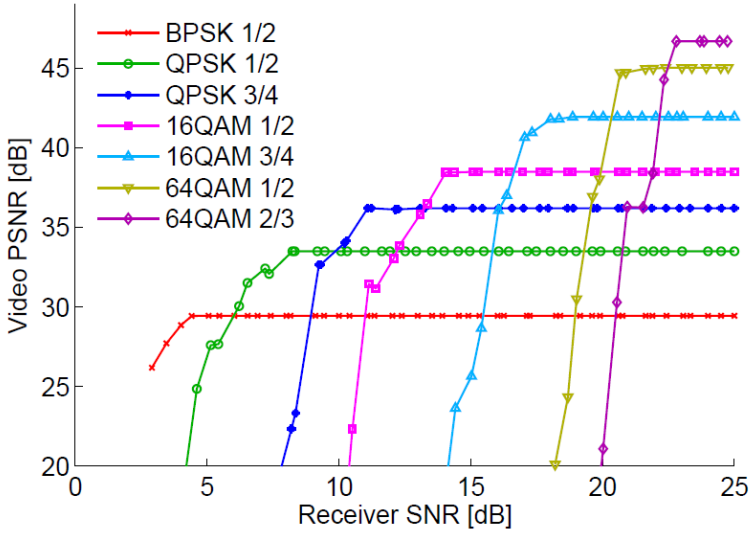


Figure 1.3: Video quality (in terms of PSNR) at receiver when a single-layer MPEG4 codec and different modulation and channel coding schemes are used. The figure comes from [JK10b].

networks show very clearly the advantages of SoftCast [JK10a]. The details of SoftCast will be presented in Section 2.1.

Since SoftCast and, more generally, LVC does not use entropy coding and temporal prediction, the efficiency in terms of pure source coding is less than that of classical video coders. Nevertheless, significant work has been done recently to improve the efficiency of LVC. A chunk shape and size optimization is proposed in [XWF+13]. The coding gain of the pixel-domain transform is analyzed in [XWX+16, XZW+17a]. Shannon-Kotel'nikov mappings are introduced [CK15, LLX+17] to reduce the number of dropped chunks under bandwidth constraint. Following the ideas of [PWS94, Sch95, SPA02, MP02, GARRM05, SPA06] hybrid digital-analog SoftCast-based architectures have been proposed in [CSY+13, SXM+14, YLL14, FWZA13, FXWZ12, FXZW15, ZFXZ13, ZLCW16, YLL15, LLZW18, ZWL+18]. On the other hand, the transmission channel characteristics may be considered to optimize SoftCast-based video transmission. The first papers considered a wideband AWGN channel [JSKG11]. Then, fading channels [HLL+17, ZLMW17, LHL+14b, ZLCW16], as well as MIMO channels [LHL+14a, LHL+14b] have then been considered. An adaptation of SoftCast with channel gains depending on the subchannel has been introduced in [HLL+17, ZWW+15, ZLMW17, LHL+14b, CSY+13] considering a total transmission power constraint. In that case, the chunks with the most energy are transmitted over the best channels after a proper scaling. More details about extensions of SoftCast are provided in Section 2.2.

Nevertheless, all of the previously mentioned papers consider a single constraint on the total transmission power when evaluating the optimal scaling factors for the chunks. For some channels, such as DSL or PLT [YAA+13], OFDM is employed and the power constraint depends on the subchannel, see for example, the constraint on the Power Spectrum Density (PSD) in PLT shown in Figure 1.4. Similarly, for multi-antenna transmission, each antenna may have its own power constraint [YL07]. In such situations, new power allocation schemes are needed. This is one of our contributions in this thesis. In Chapter (3), we propose an optimal precoding and decoding matrix design method for channels with per-subchannels power constraints. This methods involves multi-level water-filling [PLC04] and the solution of an inverse eigenvalue problem [ZZ95]. Then in Chapter (4), suboptimal power allocation methods will be presented, reduce significantly the execution time and have negligible performance loss compared to the

optimal design techniques of Chapter (3).

Another important issue for video transmission is the mitigation of impulse noise. Several communication channels may be prone to impulse noise, such as the Digital Subscriber Line (DSL) [Ned03] and the Power Line Telecommunications (PLT) channels [ZD02]. Impulse noise has a high magnitude (its power may be 50dB above that of the background noise), and when it is bursty, may corrupt the channel for more than 1 ms [ZD02]. If impulses are not corrected, the communication performance may be significantly degraded [ANQC14, LNE13], even if LVC schemes are more robust than classical video coding scheme to noise and channel mismatch [JK10b]. In Chapter 5, we have addressed this mitigation of impulse noise for SoftCast-based video transmission problem.

1.2 Contributions

In my thesis, we address two problems related to SoftCast-based video coding and transmission systems. The first is the power allocation under per-subchannel power constraints; the second is the optimal subchannel provisioning for impulse noise correction.

The first original contribution of this thesis is to optimize the power allocation for SoftCast-based video coding and transmission systems when the channel is made up of several parallel subchannels with different power constraints. The optimization consists in minimizing the receiver mean square error (MSE), and to do this, one has to find a *precoding* matrix that transforms the chunks' coefficients, modeled as independent Gaussian sources with different variances, so that they match the individual subchannel power constraints. One has also to determine an optimal *decoding* matrix at receiver. The optimization problem may be solved considering Karush-Kuhn-Tucker (KKT) conditions [BV04]. Nevertheless, this method leads to a system of nonlinear equations, which is difficult to solve directly as the number of subchannels increases.

A similar problem has been addressed in [YL07] in the context of downlink beamforming with per-antenna power constraints. Strong duality is used to transform this problem to an uplink beamforming problem (with signal to interference plus noise constraints) with uncertain noise. That solution is not directly amenable to our problem, since the downlink channel characteristics considered in [YL07] are independent of the beamforming vector to optimize. In

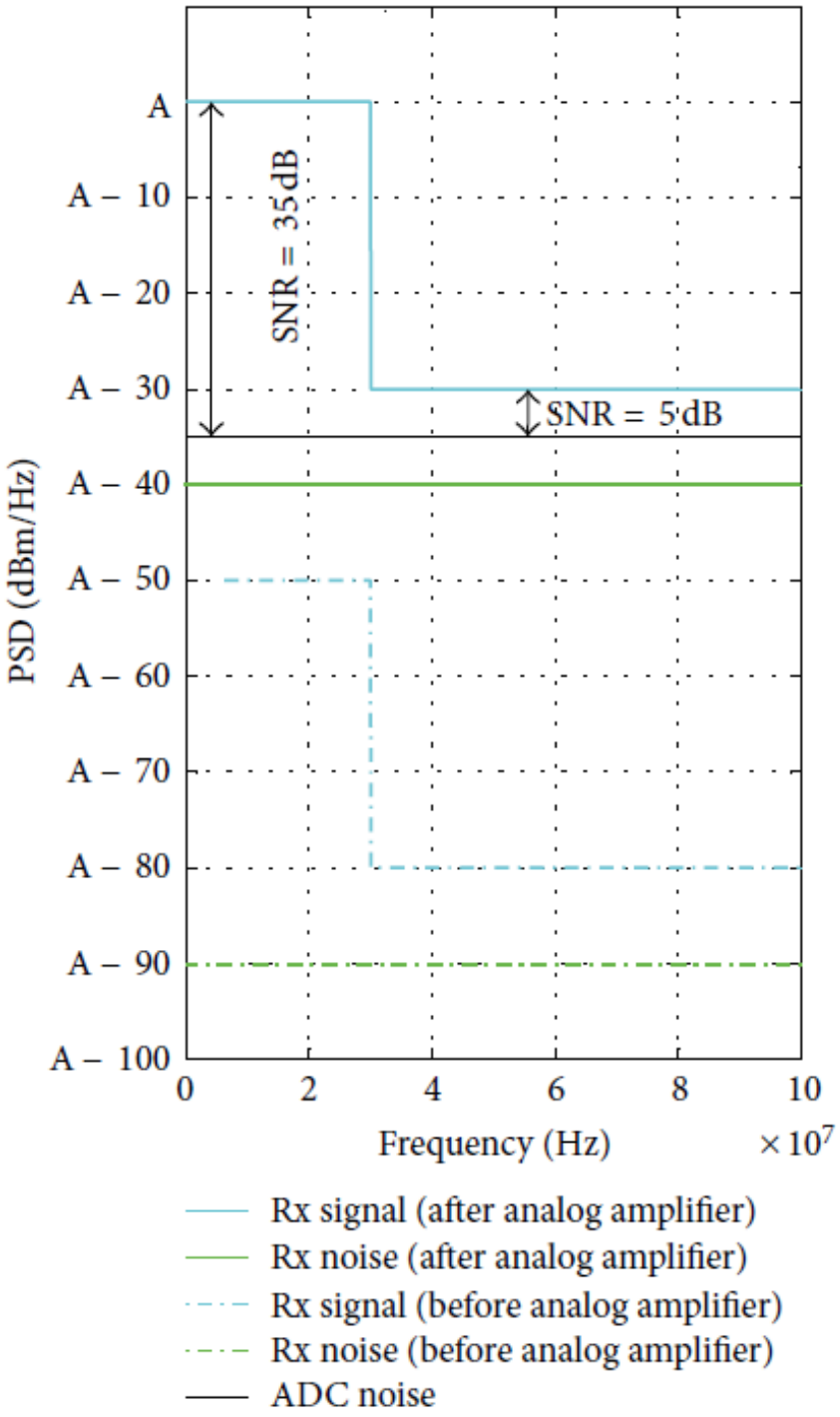


Figure 1.4: Power Spectrum Density (PSD) no power back-off in PLT. The figure is from [YAA+13].

our case the optimal decoding matrix depends on the precoding matrix.

In our work, the design of optimal precoding and decoding matrices with per-subchannel power constraints, after reformulation, will lead to an *inverse eigenvalue* problem. Such problem is found in several application contexts, see [Chu98] and the reference therein. One focuses on the specific class of Structured Hermitian Inverse Eigenvalue (SHIE) problem. This problem has been considered in [LP76] and encountered later in the context of CDMA [VA99] and MIMO communication [PLC04, PJ07]. The design proposed in [LP76] is optimal when a set of sufficient conditions on some problem parameters (in our case, the chunk variances, the subchannel power constraints, and the noise variances) are satisfied. When they are not satisfied, an heuristic approach has been proposed. Nevertheless, it is suboptimal and indeed no proof of optimality is provided in [LP76]. The multi-level water-filling approach proposed in [PLC04, PJ07], which aims to minimize the total transmission power with per-subchannel MSE constraints, allows one to find iteratively the optimal solution, but with a large computing cost.

The main contributions of the first part of our work consist in addressing the design of optimal precoding and decoding matrices with total and per-subchannel power constraints in the context of LVC. We provide an optimal solution and three lower-complexity alternative suboptimal solutions. For the optimal solution, the derivations of [LP76] are adapted, considering the majorization techniques used in the MIMO context by [PCL03] and the multi-level water-filling approach proposed in [PLC04, PJ07]. Inspired by the optimal approach, lower-complexity suboptimal design methods are proposed which are able to reduce significantly the design complexity. Moreover simulation results show that they have a very small performance degradation for most of the considered video sequences and are better than [LP76].

Moreover, we consider also the use case of point-to-multipoint video communication, which is a typical application of LVC schemes. In this case, the channel experienced by different users have different characteristics, in particular different noise levels. In such cases, the transmitter can only implement power allocation with respect to some target noise level, *e.g.,* the average noise level among users [YLL14, FXWZ12], introducing thus a mismatch between the actual channels' noise and the one used to design power allocation. We analyze the robustness of the proposed schemes to mismatched channel characteristics in Sections 3.5, 3.6.5, and 4.6.3. This is very important, since these results show the applicability of the optimal precoding matrix

design to the point-to-multipoint case.

The second contribution of our work is considering the problem of impulse noise mitigation when video is encoded using an LVC scheme and transmitted using an Orthogonal Frequency-Division Multiplexing (OFDM) scheme for multi-carrier modulation over a wideband channel prone to impulse noise. In the time domain, the impulse noise is modeled as independent and identically distributed Bernoulli-Gaussian variables. A Fast Bayesian Matching Pursuit (FBMP) algorithm [SPZ08] is employed for impulse noise mitigation. This approach requires the provisioning of some OFDM subchannels to estimate the impulse noise locations and amplitudes. Provisioned subchannels cannot be used to transmit data and lead to a decrease of the video quality at receivers in absence of impulse noise. Using a phenomenological model (PM) of the residual noise variance after impulse mitigation in the subchannels, we have proposed an algorithm namely LVC with Optimal Subchannel Provisioning for Impulse noise Correction (LVC-OSP-IC), which is able to evaluate the optimal amount of subchannel to provision which minimizes the mean-square error of the decoded video at receivers. Simulation results show that the PM can accurately predict the number of subchannels to provision and that impulse noise mitigation can significantly improve the decoded video quality compared to a situation where all subchannels are used for data transmission.

1.3 Organization of thesis

The rest of the thesis is organized as follows.

Joint source-channel coding schemes, SoftCast and its developments are presented in Chapter 2. Our results on optimal precoding and decoding matrix design are presented in Chapter 3. We start in Section 3.3 by presenting the optimal power allocation when a total transmission power constraint is considered, which goes beyond the solution proposed in the original SoftCast, where a simplified precoding matrix design is considered. The design of the optimal precoding matrix under per-subchannel power constraints is presented in Section 3.4. The alternative low-complexity suboptimal algorithms named *Simple Chunk Scaling* (SCS), *Power Allocation with Inferred Split Position* (PAISP), *PAISP with Dichotomy*, and *Power Allocation with Local Power Adjustment* (PALPA) are presented in Chapter 4. The different solutions are compared in Section 4.6. The transmission of several videos over realistic PLT channel models

are considered and show the advantage of optimal and suboptimal precoding matrix design approaches. The robustness of the proposed schemes to variations of the channel characteristics is analyzed in Sections 3.6.5, 3.5, and 4.6.3.

Chapter 5, presents the mitigation of impulse noise for SoftCast-based video transmission. First, the application of FBMP for impulse noise mitigation is described in Section 5.4. Then Section 5.5 presents the method to compute the optimal number of subchannels to provision for impulse noise correction. Section 5.6 shows the simulation result, in which we can see that the mitigation of impulse noise for SoftCast-based video transmission has significantly improve the performance. Chapter 6 concludes the thesis and presents some research perspectives.

Chapter 2

Related work and prior results

SoftCast [JK10a] is a joint source-channel coding scheme that can resolve the unfairness problem encountered by conventional video codecs in the broadcast scenario. In the first section of this Chapter, we present the principle of SoftCast which is an analog coding [JSKG10a, Gob65] based joint source-channel video coding and video transmission scheme. In analog coding, the source component are directly mapped on the channel after multiplication with scaling factors. The advantage compared to digital coding scheme will be shown. Then in Section 2.2, we illustrate the extensions of SoftCast that improve the global performance.

2.1 SoftCast: A joint source-channel coding scheme

Source-channel separation theorem tells us that in point-to-point communication we can perform source coding and channel coding separately. However, in the broadcast scenario, a joint source-channel coding scheme can be better than separate coding in some cases [GRV03]. In Section 2.1.2 SoftCast [JK10a] an analog coding based joint source-channel video coding and video transmission scheme is presented.

2.1.1 Information-theoretic foundations of SoftCast

In this section, we will present the information-theoretic ideas which support SoftCast. At first the source-channel separation theorem is recalled. The definitions of rate R, channel capacity C and rate distortion function $R(D)$ are given as in [CT06].

Definition 1. Let $\mathcal{X}$ be a finite set of input channel symbols, $\mathcal{Y}$ be a finite set of output

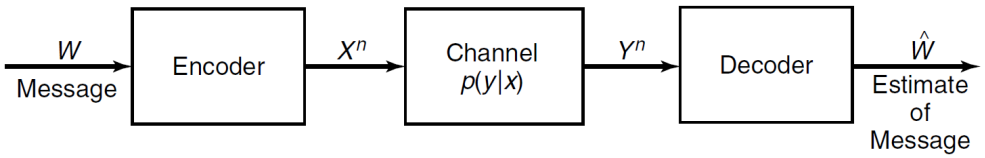


Figure 2.1: Communication channel. message W is drawn from index set $\{1, 2, \ldots, M\}$. The figure comes from [CT06].

channel symbols and $p(y|x)$ the channel transition probability, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. An (M, n) code for the channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ consists of

1) an index set $\{1, 2, \ldots, M\}$.

2) an encoding function X^n: $\{1, 2, \ldots, M\} \to \mathcal{X}^n$, yielding codewords $x^n(1), x^n(2), \ldots, x^n(M)$. The set of codwords is called *codebook*.

3) a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \ldots, M\},$$

which is deterministic rule that assigns a guess to each possible received vector.

The communication channel corresponding to these definitions is illustrated in Figure 2.1. For a discrete memoryless channel without feedback one has $p(y^n|x^n) = \prod_{i=1}^{n} p(y_i|x_i)$.

Definition 2. The rate R of an (M, n) code is

$$R = \frac{\log M}{n} \text{ bits per transmission.}$$

The choice of R will depend on the channel capacity, which is defined below.

Definition 3. The information channel capacity C of a discrete memoryless channel is

$$C = \max_{p(x)} I(X; Y),$$

where $I(X; Y)$ is mutual information and the maximum is taken over all possible input distributions $p(x)$.

The capacity of a Gaussian channel with power constraint P and noise variance N is [CT06, Theorem 9.1.1]

$$C(P) = \frac{1}{2}\log\left(1 + \frac{P}{N}\right) \text{ bits per transmission.} \tag{2.1}$$

Definition 4. The maximum probability of error $\lambda^{(n)}$ for an (M, n) code is defined as

$$\lambda^{(n)} = \max_{i \in \{1,2,\ldots M\}} \Pr\left(g\left(y^n\right) \neq i | X^n = x^n\left(i\right)\right).$$

The channel coding theorem [CT06, Theorem 7.7.1] shows that for a discrete memoryless channel, for every rate $R < C$, there exists a sequence of $\left(2^{nR}, n\right)$ codes with maximum probability of error going to zero.

Now let us introduce the rate-distortion code.

At first, assumes a source sequence $X_1, X_2, \ldots, X_n$ which are iid following $p\left(x\right)$, $x \in \mathcal{X}$ and $\mathcal{X}$ is a finite set. The source sequence X^n is encoded to an index $f_n\left(X^n\right) \in \left\{1, 2, \ldots, 2^{nR}\right\}$. The decoder estimates $\widehat{X}^n$ of X^n from this index and $\widehat{X}^n \in \widehat{\mathcal{X}}^n$.

Definition 5. A $\left(2^{nR}, n\right)$- rate distortion code consists of an encoding function

$$f_n : \mathcal{X}^n \to \left\{1, 2, \ldots, 2^{nR}\right\},$$

A decoding (reproduction) function,

$$g_n : \left\{1, 2, \ldots, 2^{nR}\right\} \to \widehat{\mathcal{X}}^n.$$

The distortion D associated with the $\left(2^{nR}, n\right)$ code is defined as

$$D = \mathrm{E} d\left(X^n, g_n\left(f_n\left(X^n\right)\right)\right)$$

where $d\left(x^n, \widehat{x}^n\right) = \frac{1}{n}\sum_{i=1}^{n} d\left(x_i, \widehat{x}_i\right)$ and d is a distortion measure and the expectation is with respect to the probability distribution on X

$$D = \sum_{x^n} p\left(x^n\right) d\left(x^n, g_n\left(f_n\left(x^n\right)\right)\right).$$

Definition 6. A rate distortion pair (R, D) is said to be achievable if there exists a sequence of $\left(2^{nR}, n\right)$-rate distortion codes (f_n, g_n) with $\lim_{n\to\infty} \sum_{x^n} p\left(x^n\right) d\left(x^n, g_n\left(f_n\left(x^n\right)\right)\right) \leqslant D$.

The rate distortion function $R\left(D\right)$ is the infimum of rates R such that (R, D) is achievable for a given distortion D. The distortion rate function $D\left(R\right)$ is the infimum of all distortion D

such that (R, D) is achievable for a given rate R.

The source-channel separation theorem with distortion [CT06, Theorem 10.4.1] reported below shows in which condition this distortion D can be achieved with a given channel capacity C.

Theorem 1. *(Source channel separation theorem with distortion) Let $V_1, V_2 \dots V_n$ be a finite alphabet iid source which is encoded as a sequence of n input symbols X^n of a discrete memoryless channel with capacity C. The output of channel Y^n is mapped onto the reconstruction alphabet $V^n = g(Y^n)$. Let $D = Ed\left(V^n, \widehat{V}^n\right) = \frac{1}{n}\sum_{i=1}^{n} Ed\left(V_i, \widehat{V}_i\right)$ be the average distortion achieved by this combined source and channel coding scheme. Then the distortion D is achievable if and only if the rate $R(D)$ is less than the channel capacity C*

$$R(D) < C. \tag{2.2}$$

This theorem enable us to design source encoder and channel encoder separately. The source encoder achieves the rate distortion by encoding the source sequence of length n into one of the $2^{nR(D)}$ messages. Then a channel encoder protects each one of these $2^{nR(D)}$ message from channel noises by encoding it into a sequence of n input symbols of channel with capacity C. If the distortion D can be achieved for a sufficiently large n, we must have $R(D) \leqslant C$. Source channel separation coding scheme works well as any joint source channel coding scheme in point-to-point communication provided the length of code n is infinity. In multiusers scenario where the channel capacity of each receiver is different, the rates for each receiver are shown below.

At first, let us consider a simple case with one sender and two receivers [CT06, Example 15.6.6]. The sender emits a sequence of iid Gaussian variables with variance P. The channel between the sender and the receivers are assumed to be with additive white Gaussian noise (AWGN) but the variances of noise experienced by the receivers are different. For example, the first receiver experiences a noise with a smaller variance $\sigma_1^2 < \sigma_2^2$. In this case, the encoder could encode with a common rate R_2 (coarse version) to both receivers, the receiver whith the better channel can receive refinement rate R_1 by using superposition coding [Cov72]. Then the

capacity region (R_1, R_2) where the probability of error could go to zero is

$$R_1 < \frac{1}{2}\log\left(1 + \frac{\alpha P}{\sigma_1^2}\right) \tag{2.3}$$

$$R_2 < \frac{1}{2}\log\left(1 + \frac{(1-\alpha)\,P}{\alpha P + \sigma_2^2}\right) \tag{2.4}$$

where $0 \leqslant \alpha \leqslant 1$.

To achieve this capacity region, the source has to use superposition coding. Receiver 2 which has the worst channel receives the coarse version of source symbol with distortion $D(R_2)$. Receiver 1 which has a better channel can receive the refinement version. On the other hand, the coarse version can also be decoded by Receiver 1, hence the distortion of Receiver 1 is $D(R_1 + R_2)$.

On the other hand, the condition of an optimal source-channel code [GRV03] is

$$R(D) = C(P), \tag{2.5}$$

where $C(P)$ represents the channel capacity which is a function of input cost (*e.g.* the power constraint P). In [GRV03], it has been shown that a joint source-channel coding scheme which sends directly this single iid Gaussian source with variance P over broadcast Gaussian channel performs well. This alternative solution is called as an uncoded scheme or it can be considerd as an analog coding scheme [Gob65, JSKG10b]. At receiver side, the two receivers which respectively have Gaussian channel noise with zero mean and variance σ_1^2 and σ_2^2 use a Linear Minimum Mean Square Error Estimator (LMMSE) to reconstruct the transmitted source. The reconstruction distortions D, which are measured by the Mean Square Error are respectively $\frac{P\sigma_1^2}{\sigma_1^2+P}$ and $\frac{P\sigma_2^2}{\sigma_2^2+P}$.

From the rate distortion function of Gaussian Source [CT06, Theorem 10.3.2],

$$R(D) = \frac{1}{2}\log\left(\frac{P}{D}\right), \tag{2.6}$$

it can be deduced that the $R(D)$ of Receiver 1 is

$$\begin{aligned} R(D) &= \frac{1}{2}\log\left(\frac{P}{\frac{p\sigma_1^2}{\sigma_1^2+P}}\right) \\ &= \frac{1}{2}\log\left(1+\frac{P}{\sigma_1^2}\right) \\ &= C(P). \end{aligned} \tag{2.7}$$

Eq. (2.7) shows that the rate for Receiver 1 is equal to the corresponding channel capacity. It is also the case for Receiver 2. Therefore analog coding in this situation is an optimal source-channel code. On the other hand, the distortions of superposition coding for the Gaussian source and mean square error distortion measure is the distortion function of Gaussian source [GRV03]. From (2.6), it can be deduced that the distortion rate function $D(R)$ of Gaussian source is

$$D(R) = P2^{-2R}. \tag{2.8}$$

Therefore from (2.3), (2.4), and (2.8), the distortion region of superposition coding can be deduced. In [GRV03], it has been shown that the analog joint source-channel coding scheme achieves a distortion pair point $\left(\frac{P\sigma_1^2}{\sigma_1^2+P}, \frac{P\sigma_2^2}{\sigma_2^2+P}\right)$ which is strictly outside of distortion region achieved by superposition coding (see Figure 2.2). This is for the single Gaussian source. The performance of this analog joint source channel coding with multi-variate Gaussian vector source is shown in below.

In [JSKG10b], the performance (the measure is mean square error) of analog coding based communication scheme and of digital communication is compared in point-to-point communication and in broadcast respectively with a multi-variate Gaussian vector source of dimension N. The covariance of this source is a diagonal matrix with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_N$, which are assumed in decreasing order. The source vector is transmitted over M AWGN channels with a specific SNR. In point-to-point communication, the optimum performance in terms of distortion of the analog coding based communication scheme and of the digital communication scheme are respectively D_{ana} and D_{dig}, which can be represented as

$$D_{\text{ana}} = \frac{\left(\sum_{i=1}^{k}\sqrt{\lambda_i}\right)^2}{M\text{SNR}+K} + \sum_{i=K+1}^{N}\lambda_i, \tag{2.9}$$

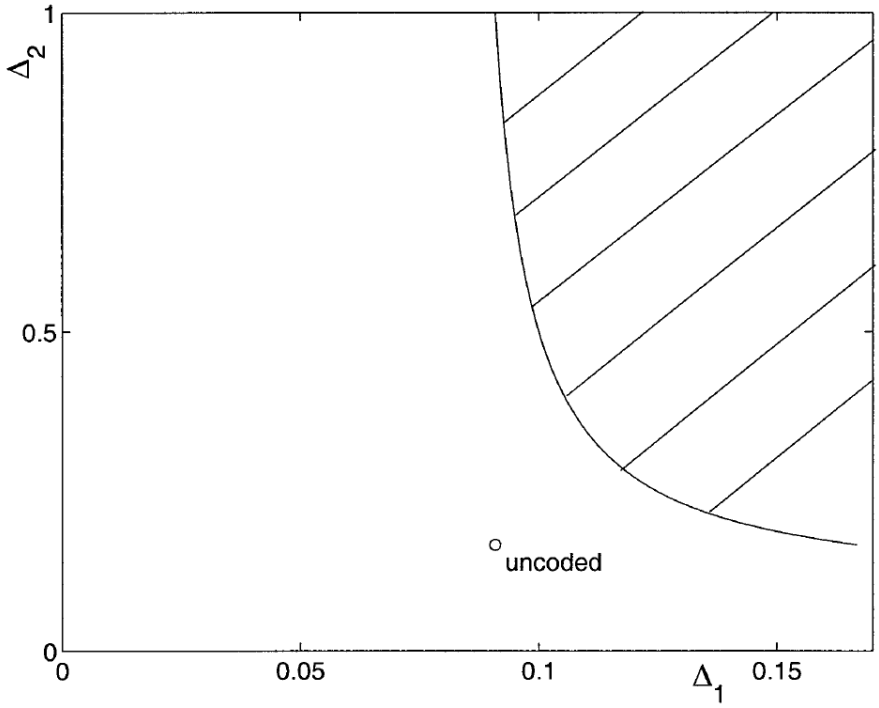


Figure 2.2: The achievable distortion. The horizontal axis Δ_1 and vertical axis Δ_2 represent respectively the distortion of Receiver 1 and Receiver 2. The circle represents the achieved distortion of analog coding, while the shallow region represents the achievable region of separation coding scheme. To have these distortions, the parameters are set with $P = 1$, $\sigma_1^2 = 0.1$, $\sigma_2^2 = 0.2$. The figure comes from [GRV03].

and

$$D_{\text{dig}} = K\left(\frac{\prod_{i=1}^{K}\lambda_i}{(\text{SNR}+1)^M}\right)^{1/K} + \sum_{i=K+1}^{N}\lambda_i, \tag{2.10}$$

where K is the number of transmitted source elements with analog communication or with digital communication, which depends on the channel SNR and bandwidth. (2.9) represents the distortion (MSE) of analog coding based communication under total power constraint. Since these are AWGN channels and the variance of noise is unity, the total power constraint can be represented as MSNR. The minimum MSE computation involves water-filling as shown in Section 3.3. Moreover, Eq. (2.10) represents the distortion (MSE) of N Gaussian random variables which are transmitted over M AWGN channels with a specific SNR by under digital communication scheme. To obtain (2.10), at first by using [CT06, Theorem 10.3.3], one gets

$$R(D) = \sum_{i=1}^{N}\frac{1}{2}\log\frac{\lambda_i}{D_i}, \tag{2.11}$$

where

$$D_i = \begin{cases} \gamma & \text{if } \gamma < \lambda_i \\ \lambda_i & \text{if } \gamma \geqslant \lambda_i \end{cases}, \tag{2.12}$$

where γ is chosen such that $\sum_{i=1}^{N} D_i = D$. Moreover, since the channel capacity of M AWGN channels (see (2.1)) is,

$$C\left(\mathrm{SNR}\right) = \frac{M}{2}\log\left(1+\mathrm{SNR}\right), \tag{2.13}$$

one gets 2.10 by setting (2.11) equal to 2.13.

In [JSKG10b], the ratio $\frac{D_{\mathrm{dig}}}{D_{\mathrm{ana}}}$ is compared in different scenarios. It shows that in point-to-point communication, the analog scheme is better than digital system for a very low SNR. In broadcast, analog scheme is nearly optimal as digital system when the dimension of source vector and dimension of channel is matched, otherwise the analog scheme is worse when the compression ($N > M$) or expansion of bandwidth ($N < M$). However the analog scheme is better when SNR of weak user is low. This advantage is helpful and has been used for video transmission in broadcast, in which the SNR is different among receivers. As we have seen in Figure 1.3, by using digital coding scheme for video compression in broadcast, it should decide the bit rate by considering the channel C-SNR of the worst channel and of the better channel. In this case, although the channel of one user has high C-SNR, the received video performance is not proportional to the channel quality. This is unfairness. Nevertheless an analog coding based video transmission scheme namely SoftCast [JK10a], in which the received video performance is linear with the C-SNR. It is shown in Section 2.1.2.

2.1.2 SoftCast

The architectures of SoftCast is shown in Figure 2.3. The input video signal undergoes a linear 3D-DCT, consisting of a full-frame 2D-DCT followed by a temporal 1D-DCT on a Group of Pictures (GoP) of n_{F} frames of $n_{\mathrm{R}} \times n_{\mathrm{C}}$ pixels. SoftCast works independently GoP by GoP. After a GoP has been transformed, the resulting coefficients are grouped into chunks. A chunk is a set of $n_{\mathrm{r}} \times n_{\mathrm{c}}$ spatial coefficients belonging to the same temporal subband (assuming they follow a similar distribution). The n_{Ck} chunks are sorted according to their variance λ_i , where $i = 1, \ldots, n_{\mathrm{ck}}$ and only the first ℓ of them may be sent, according to the bandwidth limitations

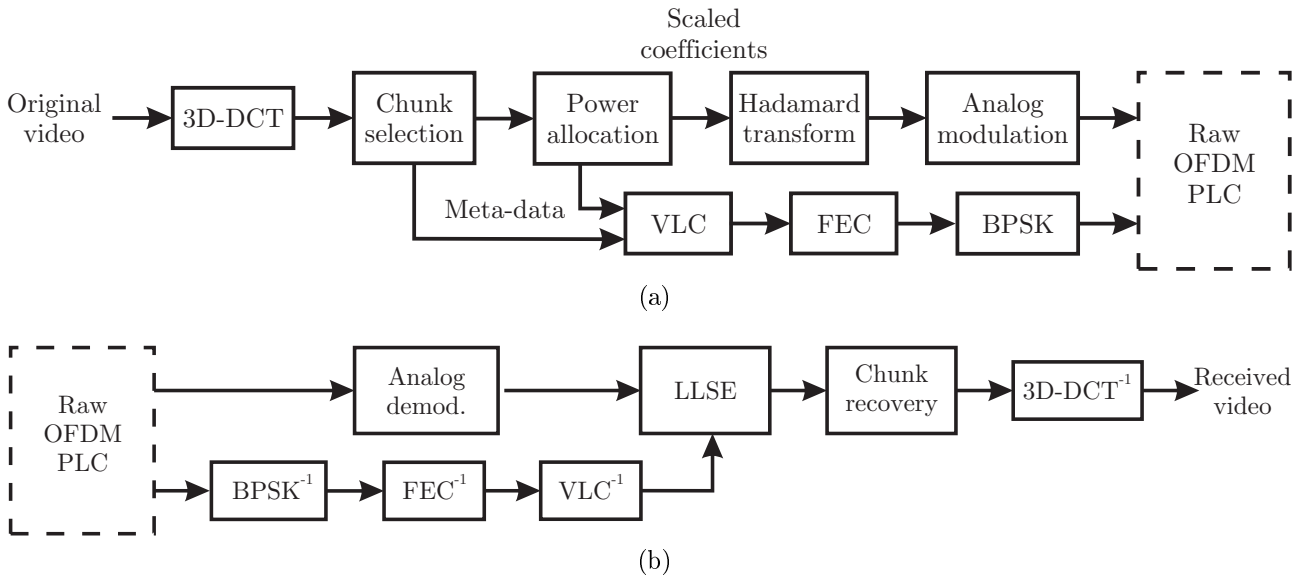


Figure 2.3: SoftCast transmitter (a) and receiver (b)

and to the power constraint of the channel. More details of chunk selection will be given in Section 3.3. The map of the selected chunks is robustly transmitted (*e.g.* using a strong FEC) as metadata on the channel: since the number of chunks is relatively small, the rate overhead is not a big issue.

The selected chunks are scaled by power allocation for error protection in order to minimize the reconstruction MSE at the decoder. In SoftCast, only the total power constraint P_{T} is considered. To compute the scaling factor g_i for each chunk, in SoftCast, it is assumed that the channel SNR is high encough and the channel is AWGN. Under these hypotheses, it is possible to find that

$$g_i = \lambda_i^{-\frac{1}{4}} \left(\sqrt{\frac{P_{\mathrm{T}}}{\sum_i^{n_{\mathrm{ck}}} \sqrt{\lambda_i}}} \right). \tag{2.14}$$

In order to increase the resilience to packet losses, SoftCast uses a Hadamard matrix to transform the chunks into equal-energy *slices*. The slices are then transmitted, *e.g.*, via OFDM. At the receiver, SoftCast uses the Linear Least Square Estimator (LLSE) to decode a sequence of received symbols. Thus, SoftCast uses linear transforms in compression, in error, and in loss protection. Combining with linear estimation at the receiver, all of these linear operations make that the quality of video in receiver scales linearly with the channel quality (C-SNR). The performance of SoftCast is shown in Fig 2.4. It shows that for conventional video coder, the video coder and channel coder should be been adjust when the SNR is changed, otherwise there is a cliff effect when the SNR decreases or there is a saturation problem when the SNR

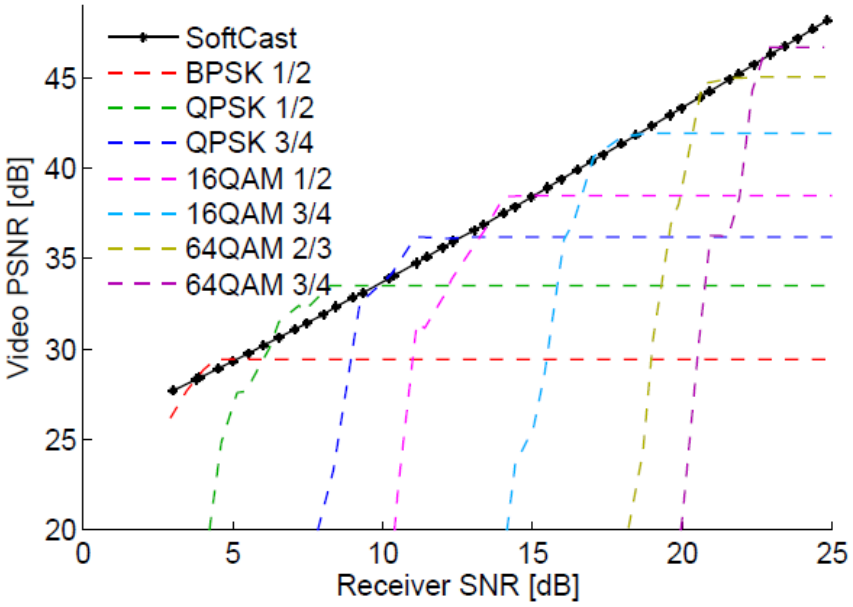


Figure 2.4: PSNR (dB) as a function of the channel SNR for SoftCast (in black) and for single-layer MPEG-4. Figure taken from [JK10b].

increases. However for SoftCast, the video's quality increases linearly evaluated with channel quality. This is important in broadcast scenario, the receiver who has high SNR channel can receive a high PSNR video, the receiver who has low SNR channel receives low PSNR video. It is unlike conventional video coder, which should choose an appropriate bit rate that could be transmitted over a low SNR channel, otherwise there is cliff effect. This is unfair for the receiver who has high SNR channel. SoftCast has resolved this fairness problem.

However SoftCast does not use entropy coding and motion estimation to reduce the redundancy information in the video and no quantization for compression, which will in return decrease the performance of SoftCast [JSKG10b]. To overcome the limitation of pure analog communication scheme and also keep the benefit of that, no cliff or threshold effect. A general hybrid digital analog (HDA) source and channel coding version is proposed in [MP02, SPA02, SPA06], which combines analog coding and digital coding. It provides a robust and graceful performance over a wide range channel SNR conditions. The one HDA architecture has been proposed in [PWS94, Sch95]. The SoftCast-based HDA will be presented in next section.

2.2 Improvements of SoftCast

Even though SoftCast offers a graceful video performance in broadcast scenario. However there is still a much room to improve SoftCast. For example, combines with digital coding scheme, *e.g.* quantization and motion estimation, to increase the performance; chunk size computation under power constraint and bandwidth constraint; the adaptation of SoftCast under more complex channel model and *etc.* In this section, we will present some important improvements of SoftCast.

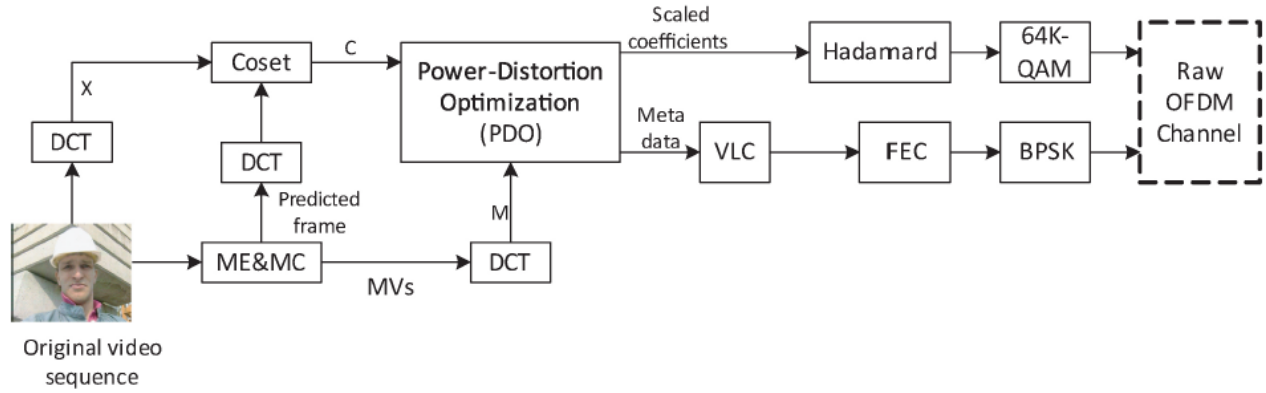
2.2.1 Dcast

One important development of SoftCast is Dcast [FWZA13]. Dcast is a distributed video coding scheme [GARRM05] and it also can be considered as a HDA scheme [SPA06]. The architecture of Dcast is shown in Figure 2.5. The key astute of Dcast is using side source information in encoding and decoding. The side information is computed by performing 2D-DCT over predicted frame, which is obtained by motion estimation and motion compsensation. In encoding (Figure 2.5a), the side information is used for Coset. In decoding (Figure 2.5b), after inverse of coset and inverse of DCT, the reconstructed pixels and predicted pixels are combined through LMMSE to reconstruct video. In this way, they improve the performance of SoftCast by 1.5dB at low channel SNR (see Figure 2.6). Similar work of Dcast can be found at [FWZ+12, ZFXZ13, FXZW15].

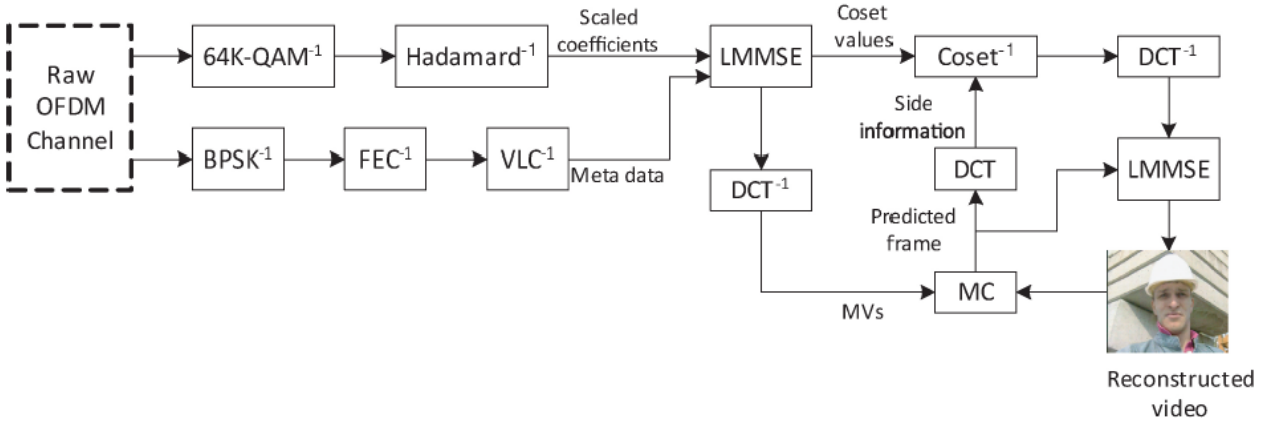
In the encoder of DCast (see Figure 2.5a), at first there is coset coding. Let X be an original 2D-DCT transformed video frame in a vector form. X_i is DCT coefficient in ith subband. For each X_i, Dcast has a uniform quantizer $Q_i()$ and get a residual value C_i

$$C_i = X_i - Q_i(X_i). \tag{2.15}$$

Then all C_i are transmitted after power allocation under total power constraint P_{coset} and modulation. In the receiver, $\widehat{C}_i$ is obtained after LMMSE decoding. The side information in receiver side is represented as S_is, where S_i is the predicted DCT coefficients of ith subband. In this case, the motion vectors which are estimated at encoder should be transmitted to receiver. The motion vectors are transformed by DCT and then scaled under total motion vector power



(a) Dcast encoder



(b) Dcast decoder

Figure 2.5: Dcast encoder and decoder. The figure comes from [FWZA13].

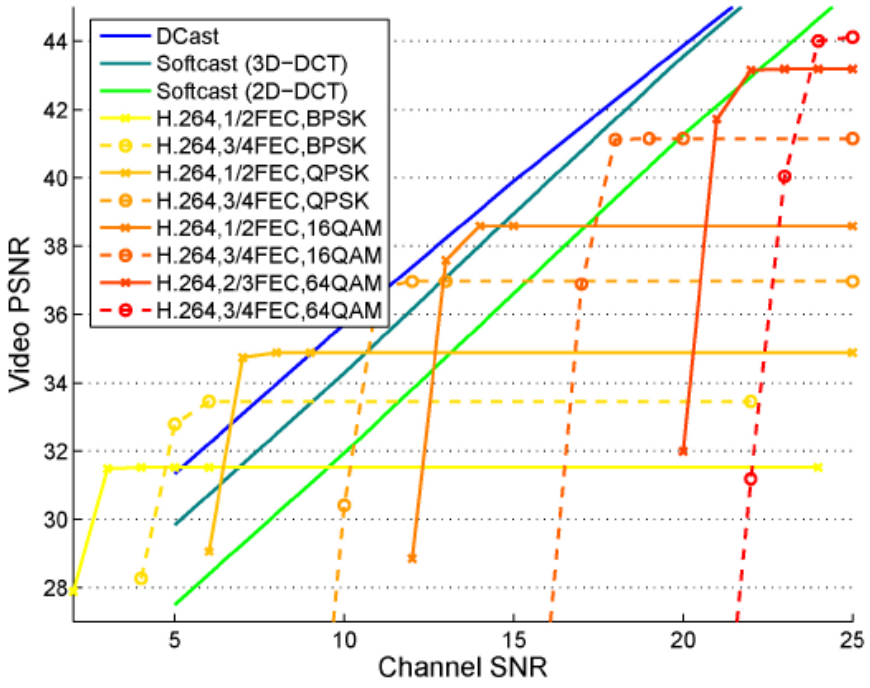


Figure 2.6: Comparison between Dcast, SoftCast, and H.264. DCast encoder is optimized for targer channel SNR of 5 dB. The figure comes from [FWZA13].

constraint P_{mv} . Since the motion vector could not be perfectly transmitted at receiver, the side information S_is also should be computed at the encoder to find the coset.

Next, the encoder designs $Q_i\left(\right)$ with a specific quantization step such that

$$Q_i\left(X_i\right) = Q_i\left(S_i - \widehat{C}_i\right) \tag{2.16}$$

with high probability. Therefore from (2.15) the reconstrcuted $\widehat{X}_i$ is

$$\widehat{X}_i = \widehat{C}_i + Q_i\left(S_i - \widehat{C}_i\right). \tag{2.17}$$

Moreover from (2.15),(2.16) and (2.17), it can deduce that the distortion D of X,

$$D = E\left[\left(X - \widehat{X}\right)^T \left(X - \widehat{X}\right)\right]$$

is close to the distortion D_{coset} of C,

$$D_{\text{coset}} = E\left[\left(C - \widehat{C}\right)^T \left(C - \widehat{C}\right)\right].$$

In Dcast, it is shown that D_{coset} is a function of P_{coset} and P_{mv}. Since the total power is $P_{\text{T}} = P_{\text{coset}} + P_{\text{mv}}$, minimizing D then becomes an optimization problem

$$\begin{aligned} &\min\ D_{\text{coset}} \\ &\text{s.t. } P_{\text{coset}} + P_{\text{mv}} = P_{\text{T}}. \end{aligned}$$

Since in Dcast D_{coset} is a convex function, the optimization problem can easily be solved by differentiation of D_{coset} with respect to variables P_{coset} and P_{coset} and set it to zero. Then the optimum power allocation pair $(P_{\text{coset}},\ P_{\text{mv}})$ is found and which will be used to scale the coset values and motion vector.

In Dcast, the side information is the predicted DCT coefficients. In [SPX+17], the side information is generated in a different way. A thumbnail of a image is decompressed and upsampled, and it is then used to retrieve correlated images from a database. Next a image is reconstructed by the retrieved images which will serve as a side information in decoding.

In decoding, using the local sparsity of residual image which is generated by subtracting the upsampled image from original image and exploiting the correlation between residual image and side information, the image is reconstructed.

2.2.2 WaveCast

Instead of using 1D-DCT to exploit the temporal correlation between frames, WaveCast [FXWZ12] uses motion compensated temporal filter (MCTF) [ST03, CCA+07, ACAB07] to reduce the inter frame redundancy. At low channel SNR, WaveCast increases the video PSNR by 2dB compared to SoftCast.

MCTF is a filter that uses motion trajectories in lifting-based transform performed over on a sequence of frames of video. For M-level MCTF, there are M output high-pass subbands and 1 low-pass subband. In WaveCast, after MCTF, 2D discrete wavelet transform is used to exploit the spatial redundancy of these $M + 1$output frames. Then the output wavelet coeffcients are scaled under total power constraint (See (2.14)).

2.2.3 WSVC

Wireless scalable video coding (WSVC) framework [YLL14] is a SoftCast-based hybrid digital-analog (HDA) coding scheme [PWS94, Sch95, MP02, SPA02, SPA06]. WSVC uses 2D-DWT (discrete wavelet transform) instead of using 2D-DCT, thus it has more spatial scalability. Moreover it achieves a PSNR gain up to 3.3dB over DCast (see Figure 2.7). In the following, WSVC is briefly presented.

The architecture of WSVC is shown in Figure 2.8. At the encoder (Figure 2.8a), at first each frame within a GOP is transformed by 2D-DWT to get four different subbands: LL, LH, HL, HH. The LL subband are then compressed by conventional video codec (*e.g.* H.264). Then the residual of LL subband which is the difference between the original and the reconstructed in H.264 and the other three high pass subbands LH, HL, HH are compressed by the SoftCast codec (temporal DCT and power allocation). The output stream of H.264 is considered as a base layer of video source, while the output of SoftCast codec is analog and is considered as an enhancement layer of video source. In the receiver (Figure 2.8b), the decoded base layer and the reconstructed LL subband residual allow to obtain a low resolution (LR) video. Finally,

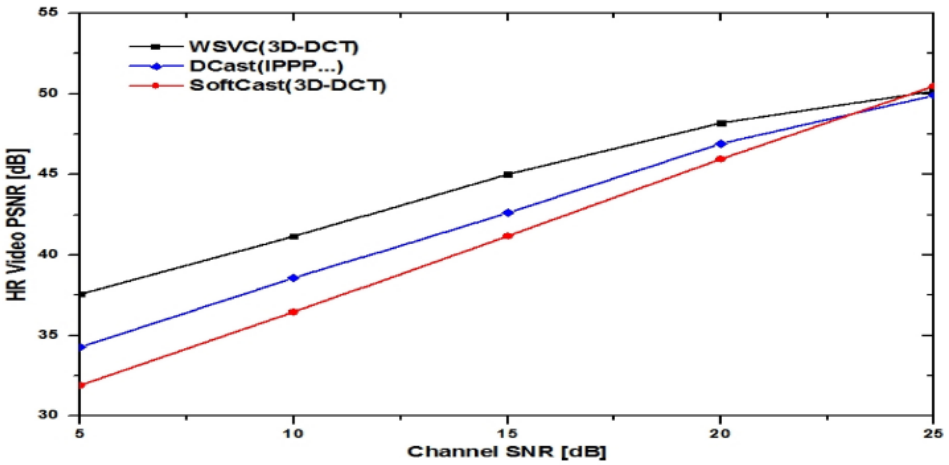


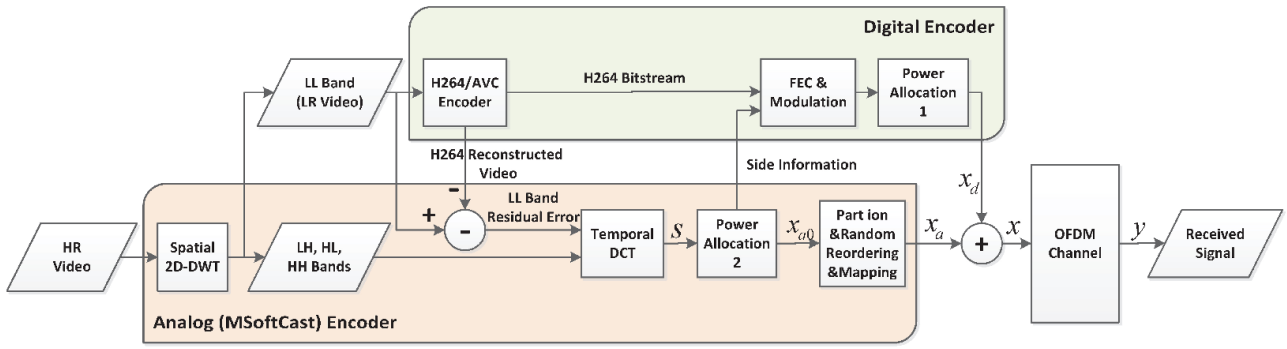
Figure 2.7: The performance (PSNR) of received HR (high resolution) video by using WSVC, Dcast (Section 2.2.1) and SoftCast for video sequence "Foreman". The target channel SNR for encoder design is $5 \sim 25$ dB. The bandwidth is 1.33MHz. The figure comes from [YLL14].

the reconstructed LR and the decoded enhancement layer source give the high resolution (HR) video. The procedure to transmit these two streams is shown in below.

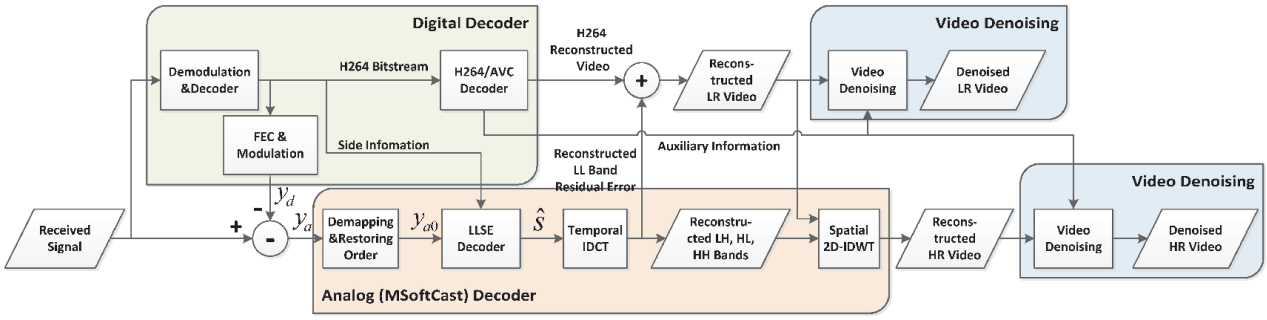
The digital stream at first is protected by forward error correction (FEC) code and then modulated by BPSK. Next each modulated component is allocated with average power P_{d}. For the enhancement layer, WSVC introduces a power allocation unit (PAU) whose role is similar to that of chunk in SoftCast. Let us assume that there are N_{p} PAUs in enhancement layer within a GoP, among which there are $N_{\mathrm{p}}^{\mathrm{LR}}$ PAUs coming from the residual of the LL subband. Then, the N_{p} PAUs are sorted in decreasing order of standard deviations σ_k, where $k = 1, \ldots, N_{\mathrm{P}}$. Moreover the average allocated power for PAU's component is $P_{\mathrm{a}}/2$, where the factor of 2 comes from the I/Q modulation as shown later. The components of each PAU are scaled with a scaling factor g_k under total power constraint $N_{\mathrm{p}} \frac{p_{\mathrm{a}}}{2}$. The forms of g_k are the same as in SoftCast (2.14), that is

$$g_k = \sqrt{\frac{N_{\mathrm{p}} \frac{p_{\mathrm{a}}}{2}}{\sigma_k \sum_{k=1}^{N_{\mathrm{P}}} \sigma_k}}. \tag{2.18}$$

At receiver side, it decodes the signal of base layer at first, and then it subtracts it from the received stream to get the enhancement layer. In this case the components of enhancement layer can be considered as noise when the base layer is decoded. In order to achieve a bit error rate (BER) in base layer which is not larger than a targer P_{E}^{T}, an approach is shown in below. In the I/Q modulation, the components of low variance PAUs are mapped on the I components



(a) the encoder part



(b) the decoder part

Figure 2.8: The framework of WSVC. The figure comes from [YLL14] .

and denoted as x_{a2} with average power P_{a2}, while the components of high variance PAUs are mapped on the Q components which are denoted as x_{a1} with average power P_{a1}. x_{a1} and x_{a2} compose x_a. Then FEC coded and BPSK modulated base layer components x_{d} are superposed with x_{a2} on I components. In this case, the transmitted signal is

$$x = x_a + x_d. \tag{2.19}$$

Assume the maximum variance of noise in the channel is N_m. The SNR of x_{d} should satisy a threshold $\gamma_o\left(P_{\mathrm{E}}^T\right)$ such that the target P_{E}^T could be achieved

$$\frac{P_{\mathrm{d}}}{P_{\mathrm{a2}} + N_m/2} \geqslant \gamma_o\left(P_{\mathrm{E}}^T\right). \tag{2.20}$$

Since the total average power is P_{T}, one gets

$$P_{\mathrm{d}} + P_{\mathrm{a}} \leqslant P_{\mathrm{T}}. \tag{2.21}$$

At the end, from (2.21), (2.20), $P_{\mathrm{a}}$and P_{d} can be easily computed

$$P_{\mathrm{a}} = \frac{(1+\mu)\left(1-\frac{\gamma_o\left(P_{\mathrm{E}}^{T}\right)}{2P_{\mathrm{T}}/N_m}\right)}{1+\mu+\gamma_o\left(P_{\mathrm{E}}^{T}\right)}P_{\mathrm{T}}$$
$$P_{\mathrm{d}} = P_{\mathrm{T}} - P_{\mathrm{a}}, \tag{2.22}$$

where $\mu = \frac{P_{a1}}{P_{a2}}$.

Now we have shown how does WSVC work. WSVC also has been applied in relay channel model [YLL15]. Many similar HDA schemes can be found at [FLWZ14, HLL$^+$15, ZLCW16, ZWL$^+$18, LLZW18]. In [ZWL$^+$18] and [LLZW18], an expression of overall distortion of HDA scheme is given, then using this expression to choose the parameters of system, for example the quantization step.

2.2.4 Energy distribution Modeling

In SoftCast [JK10a], after a GoP has been transformed by 3D-DCT, the resulting DCT coefficients are grouped into chunks. It is generally assumed that the coefficients within a chunk follow the same distribution and have the same variance. In this way, we only need to compute the scaling factor for each chunk rather than for each DCT coefficient. Moreover, only the variances of each chunk are transmitted as meta-data to receiver. Therefore the computation cost and overhead rate are reduced. However, the drawback is a reduced accuracy of the estimated variance of DCT components within each chunk, which can affect the overall performance [XZW$^+$17a]. To improve this estimation, in [XWF$^+$13, XZW$^+$17b], it has been proposed an adaptive chunk division scheme and a piecewise log linear model of energy distribution instead of using rectangular equal size chunk (See Figure 2.9). In this way, the exprimental result shows that it improves SoftCast by $3 \sim 5$dB and reduces the meta-data at the same time.

At first, the concept of transform gain [XWX$^+$16, XZW$^+$17a] is introduced. Let us consider a random vector $\mathbf{x} \in \mathbb{R}^N$, which can be the vector of all pixels or of all DCT coeffcients in a frame. For each component x_i with variance λ_i, a scaling factor g_i is computed from (2.14)

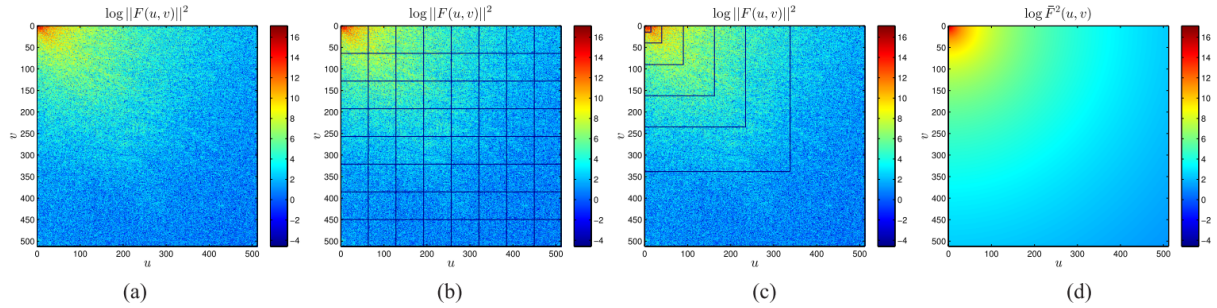


Figure 2.9: The chunk size adaptation. (a) The energy of DCT coefficients ($F(u,v)$, where (u,v) is the coordinate) in log domain; (b) equal size chunk in SoftCast; (c) adaptive chunk division; (d) curve-fitting based modeling scheme, $\bar{F}(u,v)$ is the estimated version of $F(u,v)$. The figure from [XZW+17b]

under total power constraint P_{T}, where $\sum_{i=1}^{N} E\left[g_i^2 x_i\right] \leqslant P_{\mathrm{T}}$, one gets

$$g_i = \lambda_i^{-\frac{1}{4}} \left(\sqrt{\frac{P_{\mathrm{T}}}{\sum_{i=1}^{N} \sqrt{\lambda_i}}} \right). \tag{2.23}$$

Since in [XWX+16, XZW+17a] it is assumed that the receiver does not know the variance of channel noise n_i, then the received compoent is simply decoded by inversing the scaling factor to get the reconstructed component

$$\widehat{x}_i = \frac{1}{g_i} \left(g_i x_i + n_i \right). \tag{2.24}$$

Then the distortion of x_i is

$$\begin{aligned} D_i &= E\left[\left(x_i - \widehat{x}_i \right)^2 \right] \\ &= \frac{\sigma_n^2}{g_i^2}, \end{aligned} \tag{2.25}$$

where σ_n^2 is variance of channel noise.

One gets the total distortion D_{t} as

$$\begin{aligned} D_{\mathrm{t}} &= \sum_{i=1}^{N} D_i \\ &= \sum_{i=1}^{N} \frac{\sigma_n^2}{g_i^2} \\ &= \frac{\sigma_n^2}{P_{\mathrm{T}}} \left(\sum_{i=1}^{N} \sqrt{\lambda_i} \right)^2, \end{aligned} \tag{2.26}$$

and the PSNR

$$\begin{aligned} \mathrm{PSNR}_{\mathrm{dB}} &= 10\log_{10} \left(\frac{255^2}{D_{\mathrm{t}}/N} \right) \\ &= c + \mathrm{CSNR}_{\mathrm{dB}} - 20\log_{10} \left(\frac{1}{N} \sum_{i=1}^{N} \sqrt{\lambda_i} \right), \end{aligned} \tag{2.27}$$

where $c = 10\log_{10}(255^2)$and $\mathrm{CSNR}_{\mathrm{dB}} = 10\log_{10}\left(\frac{P_{\mathrm{T}}}{N\sigma_N^2}\right)$.

From (2.27), it is observed that under total power constraint and a fixed channel noise variance, PSNR is increased when the term $\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lambda_i}$ is decreased. More precisely, the energy of frame$\sum_{i=1}^{N}\lambda_i$ does not change under orthogonal transform (*e.g.* DCT), but if the energy is concentrated in only a few components, then $\sum_{i=1}^{N}\sqrt{\lambda_i}$ is decreased. Thus, in [XWX+16] it is introduced the transform gain $G(X|\Gamma)$ for a transform $\Gamma : X(i) \to F(\mu)$, where $X(i)$ is an original component in a frame and $F(\mu)$ is component in the transformed frame

$$G(X|\Gamma) = \frac{\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lambda_{X,i}}}{\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lambda_{F,i}}}. \tag{2.28}$$

Therefore if $G(X|\Gamma)$ is large, the tranform is helpful to increase PSNR.

On the other hand, we have mentioned before that in practice it is impossible to compute scaling factor for each component. Only scaling factors for chunks are computed, moreover with the assumption that elements in a chunk have same variance. Let the estimated variance of component x_i be denoted as $\widetilde{\lambda}_i$, where $i = 1, \ldots, N$. Then (2.26) becomes

$$\widetilde{D}_{\mathrm{total}} = \frac{\sigma_n^2}{P_{\mathrm{T}}} \left(\sum_{i=1}^{N} \sqrt{\widetilde{\lambda}_i} \right) \left(\sum_{i=1}^{N} \frac{\lambda_i}{\sqrt{\widetilde{\lambda}_i}} \right). \tag{2.29}$$

[XZW+17a] shows that $\widetilde{D}_{\text{total}} \geqslant D_{\text{total}}$ and that the equality holds if and only if $\frac{\widetilde{\lambda}_1}{\lambda_1} = \frac{\widetilde{\lambda}_2}{\lambda_2} = \cdots = \frac{\widetilde{\lambda}_N}{\lambda_N}$, which means that the more accurate the variance estimations, the smaller the distortion.

From figure 2.9, it can be seen that the energy decreases along the distance $\rho = \sqrt{u^2 + v^2}$ from the upper left corner which is low frequency part, while the energy are almost same along the angle $\theta = \arctan\frac{u}{v}$. From these observations, two energy modeling scheme are proposed [XWF+13, XZW+17a]. In the first, the chunk size is adapted along the distance, in the second it is piecewise log-linear modeling along the distance. These algorithms also has been used in HDA scheme, for example [CSY+13, CXL+15] .

2.2.5 ParCast+

ParCast+ [LHL+14b] is a HDA scheme which considers the video transmission under fading channel in MIMO system. The encoding scheme is similar to WaveCast (Section 2.2.2), but the channel gains s_i^2 of each subchannel which are fed back by Channel Side Information (CSI) are taken account into the scaling factor computation under total power constraint. In ParCast+, the optimal decoding matrix is also not considered in the precoding matrix design, which is as same as in SoftCast[JK10a]. However under fading channel the scaling factors become [LHP+12]

$$g_i = (\lambda_i s_i)^{-\frac{1}{4}} \left(\sqrt{\frac{P_{\text{T}}}{\sum_i^{n_{\text{ck}}} \sqrt{\lambda_i s_i}}} \right).$$

Moreover, it proposes that a source component with high variance should be transmitted over a subchannel with higher channel gain, such that the reconstruction distortion can be reduced comparing to the other components and subchannels matching schemes. The framework of video coding and transmission in ParCast+ is shown in Figure 2.10. There are other papers which also work on fading channel [CSY+13, CLCW14b, CLCW14a, CXL+15, ZLCW16, HLL+17, ZLMW17], at which the video transmission under more complicated transmission conditions are considered, for example the channel state prediction and multicast for different users.

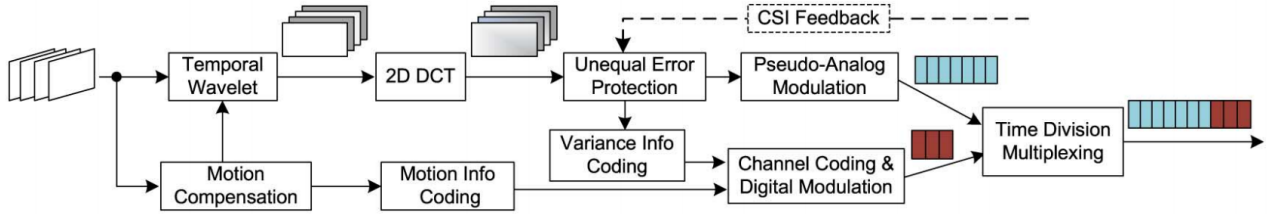


Figure 2.10: Framework of video coding and transmission in ParCast+

2.2.6 Application of Shannon-Kotel'nikov Mapping In LVC

In SoftCast [JK10a] the lower variance chunks should be discarded under bandwidth constraint. In this case, eventhough the channel quality (C-SNR) is increased, the performance could not be increased proportionally or it is saturated. [CK15] first introduces Shannon-Kotel'nikov (SK) Mapping [HFR09] in LVC.

SK mapping is helpful to reduce the number of discarded chunks under bandwidth constraint. For example, under $2:1$ SK mapping (bandwidth reduction), two iid source symbols are mapped onto a point of a parametric curve (double Archimedes' spiral). Therefore under bandwith constraint, in order to reduce the number of discarded chunks, a pair of chunks could be combined to a SK mapped chunks. However the distortion of reconstruction by using $2:1$ SK mapping has two contributions: one is the approximation of a couple of source points to one point of Archimedes' spiral; the other one is the channel noise which displaces the mapped point along the spiral arms. It is illustrated in Figure 2.11, in which Δ is the distance between spiral arms. Moreover Δ can be considered as quantization step and must be optimized given a channel state information (*e.g.* C-SNR).

The challenge of using SK mapping in LVC is the joint allocation of power and bandwidth to original chunks and SK mapped chunks. An illustration is shown in Figure 2.12. There are n_{T} chunks and n_{C} subchannels and $n_{\mathrm{T}} > n_{\mathrm{C}}$. The n_{SC} chunks among n_{T} are mapped on n_{SC} subchannels, the other $2n_{\mathrm{SK}}$ chunks are mapped on the remaining n_{SK} suchannels by $2:1$ SK mapping. The problem is how to compute n_{SK} and the power allocation between n_{SC} original chunks and $2n_{\mathrm{SK}}$ SK mapped chunks by given a channel condition.[CK15] resolved the power allocation problem under total power constraint by given a bandwidth allocation. Then [LLX+17] proposes a scheme to resolve the bandwidth allocation problem given a power allocation. By using this scheme, [LLX+17] proposes an algorithm called SK-Cast which is

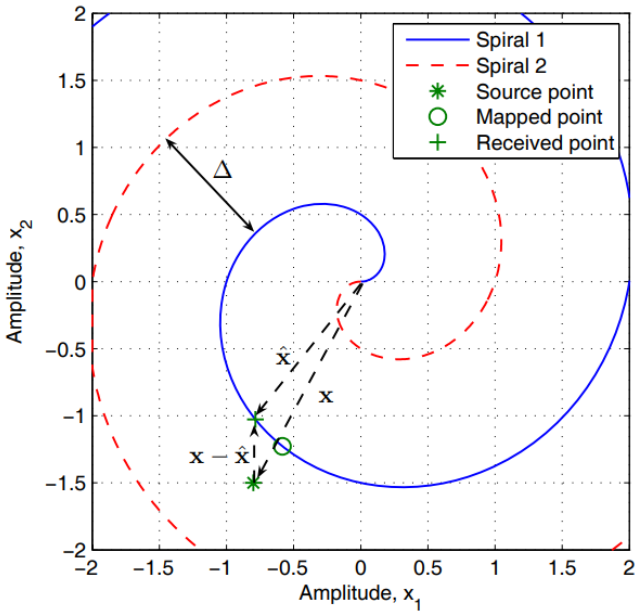


Figure 2.11: The components of distortion by using 2 : 1 Shannon-Kotel'nikov (bandwidth reduction) mapping on Archimedes' spiral. The figure comes from [HFR09].

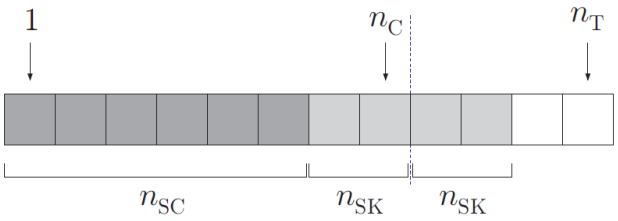


Figure 2.12: Bandwidth allocation of n_{SC} original chunks and $2n_{SK}$ SK-mapped chunks. The figure comes from [CK15].

a iterative way to allocate the power and bandwidth respectively to original chunks and SK mapped chunks. Moreover SK-Cast [LLX+17] is also a HDA scheme whose performance are better than those of WSVC (Section 2.2.3).

2.2.7 Conclusion

In this section, we have shown several important developments of SoftCast. By using HDA scheme to increase the performance (Dcast, WSVC, WaveCast). The improvement of the estimation accuracy of variances of DCT coefficients (Chunk Size Adaptation). To address the bandwidth constraint problem (Chunk Size Adaptation, SK-Cast) and the problem of video transmission under fading channel (ParCast+). There are the other interesting SoftCast-based video or image compression and transmission works. In [XLM+14, LXF+18], the transmission of image gradient is considered, which is relevant to the perceptual quality. Another perceptual

quality factor foveation point is taken account in [SYLL18]. The multiview and multiview with depth is considered in [CZX+17, FKAWO18a]. Convolutional neural network is also used at the decoder part of SoftCast to reduce the artifact [YFS18]. Moreover in [FKAWO18b, FKAWO18a] the Gaussian Markov random field (GMRF) is applied to reduce the metadata in SoftCast-based video transmission. All the papers that have mentioned here not only keeps the property of SoftCast that the video performance is linearly commensurate with C-SNR, but also improves the performance. However, only the total power transmission constraint and white Gaussian noise is considered in their problems. In other situations, the per-subchannel power constraint [YAA+13] and impulse noise [ZD02, Ned03] will be encountered. In Chapter 3 and Chapter 4, we will show our work to resolve the per-subchannel power constraint for SoftCast-based video transmission. Next in Chapter 5, a proposed impulse noise correction scheme for SoftCast based video transmission scheme will be presented.

Chapter 3

Optimal Power Allocation

3.1 Introduction

This chapter presents the joint design of precoding and decoding matrices that minimize the MSE in a SoftCast-based LVC and video transmission system (Figure 2.3) under per-subchannel power constraints. This extends results in [JK10b], where (*i*) the optimal decoding matrix is not considered for the design of the optimal precoding matrix, (*ii*) only a total power constraint is considered, (*iii*) precoding matrix design for the multiusers case is not provided.

First, Section 3.2 presents the transmission model, then Section 3.3 describes the classical minimum MSE solution under a total power constraint, proposed in [LP76], with an alternative proof involving majorization techniques advocated by [PCL03, PLC04, PJ07], where the source components were assumed all with unit variance. In our LVC case, we extend this result to source components with different variances. The solution of this first problem is then used in Section 3.4 to address the design of the precoding matrix minimizing the MSE under per-subchannel power constraints. In Section 3.6, the advantage of the proposed methods comparing respectively to [JK10b] under total power constraint and to [LP76] under per-subchannel power constraint is shown. Moreover in Section 3.6.5 the robustness of the proposed scheme to mismatched channel characteristics un has also been analyzed. We consider a multi-user scenario, where the transmitter uses a common precoding matrix for the transmission to different users.

Table 3.1 gathers the main notations used throughout this and next chapter. Random quantities are in bold, matrices in capital letters, vectors and scalars in small letters. $\mathbb{R}_+$ refers

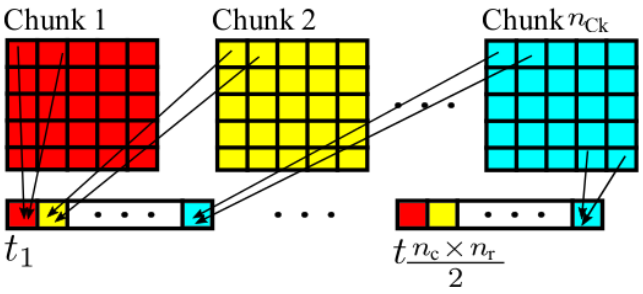


Figure 3.1: Vectorization of the chunks

to the set of non-negative real numbers and $\mathbb{R}_{++}$ to the set of positive real numbers.

3.2 Precoding and decoding matrices

For the precoding and decoding matrix design, one assumes that at the output of the 3D-DCT, the coefficients of similar variance are grouped into n_{Ck} chunks of the same size $n_{\mathrm{r}} \times n_{\mathrm{c}}$. Then a sequence of $n_{\mathrm{r}} \times n_{\mathrm{c}}$ vectors of dimension n_{Ck} is formed by selecting one coefficient per chunk for each vector, see Figure 3.1. These chunk vectors are assumed to be realizations of $n_{\mathrm{r}} \times n_{\mathrm{c}}$ independent and identically distributed zero-mean Gaussian random vectors $\mathbf{t}_i$, $i = 1 \ldots n_{\mathrm{r}} \times n_{\mathrm{c}}$ with covariance matrix $\Lambda = \mathrm{diag}(\lambda_1 \ldots \lambda_{n_{\mathrm{Ck}}})$. The matrix Λ is assumed to be diagonal, since $\mathbf{t}_i$ represents decorrelated 3D-DCT transformed pixels. In practice, the non-zero mean values of chunks are transmitted as metadata.

The chunk vectors $\mathbf{t}_i$ have to be transmitted over n_{SC} parallel AWGN subchannels with noise covariance matrix $N = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_{n_{\mathrm{SC}}}^2\right)$ and individual power constraints p_j, $j = 1, \ldots, n_{\mathrm{SC}}$. One has to find the optimal precoding and decoding matrices to minimise the MSE at receiver, while satisfying the per-subchannel power constraints. In what follows, the index i of $\mathbf{t}_i$ is omitted, since all vectors $\mathbf{t}_i$ have similar distribution and undergo the same processing. Moreover, without loss of generality, one assumes that the chunk indexing is such that $\lambda_1 \geqslant \cdots \geqslant \lambda_{n_{\mathrm{Ck}}}$.

The vector $\mathbf{t}$ is multiplied by a precoding transform matrix $G \in \mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}}$ to get

$$\mathbf{x} = G\mathbf{t}. \tag{3.1}$$

The received vector is

$$\mathbf{y} = G\mathbf{t} + \mathbf{v}, \tag{3.2}$$

Symbol	Set	Represents
n_{Ck}	$\mathbb{N}$	Dimension of source vector
n_{SC}	$\mathbb{N}$	Nb of parallel subchannels
G	$\mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}}$	Channel precoding matrix
H	$\mathbb{R}^{n_{\mathrm{Ck}} \times n_{\mathrm{SC}}}$	Decoding matrix
$n_{\mathrm{c}} \times n_{\mathrm{r}}$	$\mathbb{N}$	Chunk size
$\mathbf{t}$	$\mathbb{R}^{n_{\mathrm{Ck}}}$	Chunk vector
$\mathbf{x}$	$\mathbb{R}^{n_{\mathrm{SC}}}$	Transmitted vector
$\mathbf{y}$	$\mathbb{R}^{n_{\mathrm{SC}}}$	Received vector
$\mathbf{v}$	$\mathbb{R}^{n_{\mathrm{SC}}}$	Noise vector
σ_i^2	$\mathbb{R}_{++}$	variance noise of i-th subchannel
N	$\mathbb{R}_{++}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	Channel noise vector covariance
λ_i	$\mathbb{R}_{+}$	Variance of a chunk
Λ	$\mathbb{R}_{+}^{n_{\mathrm{Ck}} \times n_{\mathrm{Ck}}}$	diagonal source covariance matrix
p_{T}	$\mathbb{R}_{++}$	Total power constraint
ε	$\mathbb{R}_{+}$	mean-square reconstruction error
γ	$\mathbb{R}_{+}$	Lagrange multiplier
ℓ	$\mathbb{N}$	Nb of transmitted components of chunk vectors
p_i	$\mathbb{R}_{++}$	Power constraint in ith subchannel
s	$\mathbb{R}_{++}^{n_{\mathrm{SC}}}$	Vector of SNR constraints
s_i	$\mathbb{R}_{++}$	SNR constraint in i-th subchannel
S	$\mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	SNR constraints matrix
s_{eq}	$\mathbb{R}_{++}$	Total SNR constraint in equivalent channel
Z	$\mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	Orthogonal transform matrix
n_{SB}	$\mathbb{N}$	Nb of subblocks
α and β	$\mathbb{R}_{++}$	Parameters for PAISP
k	$\mathbb{N}$	Nb of receivers

Table 3.1: Main notations

where $\mathbf{v}$ is a vector of channel noise with $E(\mathbf{v}) = 0$ and $E\left(\mathbf{v}\mathbf{v}^T\right) = N$. To recover $\mathbf{t}$, $\mathbf{y}$ is multiplied by a decoding matrix $H \in \mathbb{R}^{n_{\mathrm{Ck}} \times n_{\mathrm{SC}}}$ to get

$$\widehat{\mathbf{t}} = H\mathbf{y}. \tag{3.3}$$

The mean-square reconstruction error is

$$\begin{aligned} \varepsilon &= \mathrm{tr}\left(E\left(\left(\mathbf{t}-\widehat{\mathbf{t}}\right)\left(\mathbf{t}-\widehat{\mathbf{t}}\right)^T\right)\right) \\ &= \mathrm{tr}\left(E\left((\mathbf{t}-H(G\mathbf{t}+\mathbf{v}))(\mathbf{t}-H(G\mathbf{t}+\mathbf{v}))^T\right)\right). \end{aligned} \tag{3.4}$$

Assuming that $\mathbf{t}$ and $\mathbf{v}$ are independent, $E\left(\mathbf{v}\mathbf{t}^T\right) = 0$ and ε becomes

$$\varepsilon = \mathrm{tr}\left(\Lambda - 2HG\Lambda + HG\Lambda G^T H^T + HNH^T\right). \tag{3.5}$$

3.3 Total Power Constraint

Before considering individual per-subchannel power constraints, we addresse the MSE minimization problem under a total power constraint. Without loss of generality, the noise variance indexing is such that $\sigma_1^2 \leqslant \cdots \leqslant \sigma_{n_{\mathrm{SC}}}^2$.

Assuming that a total transmission power constraint $p_{\mathrm{T}} = \sum_{i=1}^{n_{\mathrm{SC}}} p_i$ has to be satisfied, the channel input vector $\mathbf{x}$ has to be such that

$$E\left(\mathbf{x}^T\mathbf{x}\right) = \mathrm{tr}\left(E\left(\mathbf{x}\mathbf{x}^T\right)\right) \leqslant p_{\mathrm{T}}. \tag{3.6}$$

As a consequence, using (3.1), one gets

$$\mathrm{tr}(G\Lambda G^T) \leqslant p_{\mathrm{T}}. \tag{3.7}$$

One has thus to find

$$\begin{aligned} \left(\overline{G}, \overline{H}\right) = \arg\min_{G,H} \mathrm{tr}\left(\Lambda - 2HG\Lambda + HG\Lambda G^T H^T + HNH^T\right) \\ \text{s.t. } \mathrm{tr}(G\Lambda G^T) \leqslant p_{\mathrm{T}}. \end{aligned} \tag{3.8}$$

The Lagrangian function associated to (3.8) is

$$\begin{aligned}\mathcal{L}_{\mathrm{T}} &= \mathrm{tr}\left(\Lambda - 2HG\Lambda + HG\Lambda G^T H^T + HNH^T\right) \\ &\quad +\gamma\left(\mathrm{tr}(G\Lambda G^T) - p_{\mathrm{T}}\right), \end{aligned} \tag{3.9}$$

where $\gamma \geqslant 0$ is a Lagrange multiplier.

3.3.1 Optimal decoding matrix

For a given precoding matrix G, the optimal decoding matrix $\overline{H}$ is obtained by setting to 0 the partial derivative of $\mathcal{L}_{\mathrm{T}}$ with respect to H. One gets

$$\overline{H} = \Lambda G^T \left(G\Lambda G^T + N\right)^{-1}. \tag{3.10}$$

3.3.2 Optimal precoding matrix

From (3.10), one obtains

$$\overline{H}G\Lambda G^T + \overline{H}N = \Lambda G^T. \tag{3.11}$$

Right multiplying both sides of (3.11) by $\overline{H}^T$, one gets

$$\overline{H}G\Lambda G^T\overline{H}^T + \overline{H}N\overline{H}^T = \Lambda G^T\overline{H}^T. \tag{3.12}$$

Now using (3.12) in (3.5) leads to

$$\varepsilon = \mathrm{tr}\left(\Lambda - 2\overline{H}G\Lambda + \Lambda G^T\overline{H}^T\right). \tag{3.13}$$

Using the properties of the trace and (3.10) in (3.13), one gets an expression of ε that depends on G only

$$\begin{aligned} \varepsilon &= \operatorname{tr}\left(\Lambda - \overline{H} G \Lambda\right) \\ &= \operatorname{tr}\left(\Lambda - \Lambda G^T \left(G \Lambda G^T + N\right)^{-1} G \Lambda\right). \end{aligned} \tag{3.14}$$

One may rewrite (3.14) as

$$\varepsilon = \operatorname{tr}\left(\Lambda^{\frac{1}{2}} \left(I - \Lambda^{\frac{1}{2}} G^T \left(G \Lambda G^T + N\right)^{-1} G \Lambda^{\frac{1}{2}}\right) \Lambda^{\frac{1}{2}}\right) \tag{3.15}$$

The matrix inversion lemma leads to

$$I - \left(G\Lambda^{\frac{1}{2}}\right)^T \left(\left(G\Lambda^{\frac{1}{2}}\right)\left(G\Lambda^{\frac{1}{2}}\right)^T + N\right)^{-1} G\Lambda^{\frac{1}{2}} = \left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}$$

and (3.15) becomes

$$\begin{aligned} \varepsilon &= \operatorname{tr}\left(\Lambda^{\frac{1}{2}} \left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1} \Lambda^{\frac{1}{2}}\right) \\ &= \operatorname{tr}\left(\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1} \Lambda\right) \qquad (3.16) \\ &= f_\Lambda\left(\operatorname{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right)\right) \qquad (3.17) \end{aligned}$$

where

$$\Phi : \ A \in \mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}} \to \left(I + A^T N^{-1} A\right)^{-1} \in \mathbb{R}^{n_{\mathrm{Ck}} \times n_{\mathrm{Ck}}}$$

$$f_\Lambda : \mathbf{u} \in \mathbb{R}^{n_{\mathrm{Ck}}} \to (\lambda_1 u_1 + \cdots + \lambda_{n_{\mathrm{Ck}}} u_{n_{\mathrm{Ck}}}) \in \mathbb{R}_+.$$

In (3.17), the argument of f_Λ is the vector of the diagonal elements of $\Phi\left(G\Lambda^{\frac{1}{2}}\right)$. Introduce now the function $\underline{f}_\Lambda = f_\Lambda \circ \Pi_{\mathbf{u},\Lambda}$, where $\Pi_{\mathbf{u},\Lambda} : \mathbb{R}^{n_{\mathrm{Ck}}} \to \mathbb{R}^{n_{\mathrm{Ck}}}$ is the permutation that matches, for any $\mathbf{u} \in \mathbb{R}^{n_{\mathrm{Ck}}}$, the smallest u_i to the largest λ_i, the second smallest u_i to the second largest λ_i, *etc.* It has been shown in [PCL03, Appendix B] that given two vectors $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^n$,

their scalar product is minimized when the elements of $\mathbf{a}$ are sorted in increasing order and those of $\mathbf{b}$ are sorted in decreasing order. As a consequence, $\forall \mathbf{u} \in \mathbb{R}^{n_{\mathrm{Ck}}}$, $f_\Lambda(\mathbf{u}) \geqslant \underline{f}_\Lambda(\mathbf{u})$, with equality if $\Pi_{\mathbf{u},\Lambda}$ is the identity, *i.e.*, the values of $\mathbf{u}$ match those of Λ as described before.

The λ_is have been assumed sorted in decreasing order. Then, $\underline{f}_\Lambda$ is a Schur-concave function [MOA11, 3.A.4]. As a consequence, using [PCL03, Theorem 1], the matrix $G\Lambda^{\frac{1}{2}}$ that minimizes $\underline{f}_\Lambda\left(\mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right)\right)$ has the following structure

$$G\Lambda^{\frac{1}{2}} = T\begin{bmatrix} \mathrm{diag}\left(g_1\lambda_1^{1/2}\dots g_\ell\lambda_\ell^{1/2}\right) & \mathbf{0}_{\ell\times(n_{\mathrm{Ck}}-\ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times\ell} & \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times(n_{\mathrm{Ck}}-\ell)} \end{bmatrix}, \tag{3.18}$$

where $\ell \leqslant \min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$, and the g_is are scaling factors. In (3.18), T is the matrix whose columns are the eigenvectors of N^{-1} (sorted in decreasing order of their associated eigenvalues). Here, as N^{-1} is diagonal and $\sigma_1^2 \leqslant \dots \leqslant \sigma_{n_{\mathrm{SC}}}^2$, T is simply the identity matrix.

If one introduces $m_i = g_i^2\lambda_i$, then m_i is the power allocated to the components t_i of chunk vector $\mathbf{t}$ and ℓ is the number of components actually transmitted. We show how to compute ℓ later on, depending on n_{SC} and on the power constraint. If $\ell < n_{\mathrm{Ck}}$, there are some null columns in G, meaning that some components cannot be transmitted. Likewise, if $\ell < n_{\mathrm{SC}}$, there are null rows in G, which corresponds to the fact that the optimal solution does not use some subchannels.

Now, using (3.18), (3.17) becomes

$$\begin{aligned} \underline{f}_\Lambda\left(\mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right)\right) &= \underline{f}_\Lambda\left(\mathrm{diag}\left(\left(I+\Lambda^{\frac{1}{2}}G^TN^{-1}G\Lambda^{\frac{1}{2}}\right)^{-1}\right)\right) &(3.19)\\ &= \sum_{i=\ell+1}^{n_{\mathrm{Ck}}}\lambda_i + \sum_{i=1}^{\ell}\frac{\lambda_i}{1+m_i/\sigma_i^2}, &(3.20)\end{aligned}$$

The final MSE consists of two contributions. The first term $\sum_{i=\ell+1}^{n_{\mathrm{Ck}}}\lambda_i$ represents the variances of the components of the chunk vector that have not been transmitted when $\ell < n_{\mathrm{Ck}}$. These components are the $n_{\mathrm{Ck}}-\ell$ with the smallest variances since the λ_is are sorted in decreasing order. The second term depends on the variances λ_i of the ℓ remaining components, the variances σ_i^2 of the subchannel noise components, and the allocated powers given by m_i.

Now the optimization problem (3.8) consists in finding an optimal power allocation vector, which can be formulated as

$$[\overline{m}_1 \dots \overline{m}_\ell] = \arg \min_{[m_1 \dots m_\ell] \in \mathbb{R}_+^\ell} \sum_{i=1}^{\ell} \frac{\lambda_i}{1 + m_i/\sigma_i^2} + \sum_{i=\ell+1}^{n_{\mathrm{Ck}}} \lambda_i$$
$$\text{s.t.} \sum_{i=1}^{\ell} m_i \leqslant p_{\mathrm{T}}$$

To solve this convex optimization problem, we introduce the Lagrangian function

$$L_{\mathrm{T}}(m_1 \dots m_\ell, \gamma) = \sum_{i=\ell+1}^{n_{\mathrm{Ck}}} \lambda_i + \sum_{i=1}^{\ell} \frac{\lambda_i}{1 + m_i/\sigma_i^2} + \gamma \left(\sum_{i=1}^{\ell} m_i - p_{\mathrm{T}} \right), \tag{3.21}$$

where γ is the Lagrange multiplier. Then, differentiating (3.21) with respect to m_i and setting to zero, one gets,

$$\gamma = \frac{\lambda_i/\sigma_i^2}{(1 + m_i/\sigma_i^2)^2}, \tag{3.22}$$

which can be written as

$$\sqrt{\gamma} \left(\sigma_i^2 + m_i \right) = \sqrt{\lambda_i \sigma_i^2}. \tag{3.23}$$

Summing (3.23) over $i \in 1, \dots, \ell$ and recalling that $\sum_{i=1}^{\ell} m_i = p_{\mathrm{T}}$, one gets

$$\sqrt{\gamma} = \frac{\sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2}}{p_{\mathrm{T}} + \sum_{i=1}^{\ell} \sigma_i^2}. \tag{3.24}$$

Finally, we can compute the power allocation for each component of the chunk vector from (3.23),

$$m_i = \sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 \tag{3.25}$$

where $1 \leqslant i \leqslant \ell \leqslant \min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$. Since one should have $m_i > 0$, one chooses ℓ as the largest

integer less than $\min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$ that satisfies

$$\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 > 0, \ i = 1, \ldots, \ell,$$

where $\sqrt{\gamma}$ is given by (3.24).

From (3.20) one gets the minimum value of the distortion

$$\underline{f}_{\Lambda}\left(\mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right)\right) = \sum_{i=\ell+1}^{n_{\mathrm{Ck}}} \lambda_i + \sqrt{\gamma} \sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2} \tag{3.26}$$

Finally, the non-zero diagonal elements of the precoding matrix G can be computed as

$$\forall i \in 1, \ldots, \ell, \quad g_i = \left(\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2\right)^{1/2} / \sqrt{\lambda_i}. \tag{3.27}$$

Moreover, one has

$$\begin{aligned}
\Phi\left(G\Lambda^{\frac{1}{2}}\right) &= \left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1} \\
&= \mathrm{diag}\left(\frac{1}{1 + m_1/\sigma_1^2}, \ldots, \frac{1}{1 + m_\ell/\sigma_\ell^2}, 1, \ldots, 1\right) \\
&= \mathrm{diag}\left(\sqrt{\frac{\gamma\sigma_1^2}{\lambda_1}}, \ldots, \sqrt{\frac{\gamma\sigma_\ell^2}{\lambda_\ell}}, 1, \ldots, 1\right).
\end{aligned}$$

The λ_is are decreasing and the σ_i^2s are increasing. The components of $\Phi\left(G\Lambda^{\frac{1}{2}}\right)$ are thus sorted in increasing order and

$$\Pi_{\mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right),\Lambda} \mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right) = \mathrm{diag}\left(\Phi\left(G\Lambda^{\frac{1}{2}}\right)\right).$$

As a consequence,

$$\overline{G} = \begin{bmatrix} \mathrm{diag}\,(g_1 \ldots g_\ell) & \mathbf{0}_{\ell \times (n_{\mathrm{Ck}} - \ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}} - \ell) \times \ell} & \mathbf{0}_{(n_{\mathrm{SC}} - \ell) \times (n_{\mathrm{Ck}} - \ell)} \end{bmatrix}$$

is also such that $\varepsilon = f_{\Lambda}\left(\mathrm{diag}\left(\Phi\left(\overline{G}\Lambda^{\frac{1}{2}}\right)\right)\right)$ is minimized.

In summary under total power constraint, the expression of the optimal precoding and decoding matrices are respectively

$$\overline{G} = \begin{bmatrix} \operatorname{diag}(g_1 \dots g_\ell) & \mathbf{0}_{\ell\times(n_{\mathrm{Ck}}-\ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times\ell} & \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times(n_{\mathrm{Ck}}-\ell)} \end{bmatrix}, \tag{3.28}$$

and

$$\overline{H} = \Lambda \overline{G}^T \left(\overline{G} \Lambda \overline{G}^T + N \right)^{-1}. \tag{3.29}$$

In (3.28), $\ell \leqslant \min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$ is the largest integer such that

$$\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 > 0, \ i = 1, \dots, \ell \tag{3.30}$$

with

$$\sqrt{\gamma} = \frac{\sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2}}{p_{\mathrm{T}} + \sum_{i=1}^{\ell} \sigma_i^2} \tag{3.31}$$

and

$$g_i = \left(\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 \right)^{1/2} / \sqrt{\lambda_i}, \ i = 1, \dots, \ell. \tag{3.32}$$

With this optimal precoding matrix G, the distortion 3.17 becomes,

$$\varepsilon = \frac{\left(\sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2} \right)^2}{p_{\mathrm{T}} + \sum_{i=1}^{\ell} \sigma_i^2} + \sum_{i=\ell+1}^{n_{\mathrm{C}}} \lambda_i. \tag{3.33}$$

Only the ℓ components of the chunk vector with the largest variances are transmitted on the subchannels with smallest noise variances, which is consistent with the results in [LP76]. In SoftCast original paper [JK10a] the optimal decoding matrix is not considered in the precoding matrix design. Or it can be said in the other way that it assumes the SNR of channel (C-SNR) $\frac{p_{\mathrm{T}}}{\sum_{i=1}^{n_{\mathrm{SC}}} \sigma_i^2}$ is high enough such that the N could be approximated as zero in (3.29).

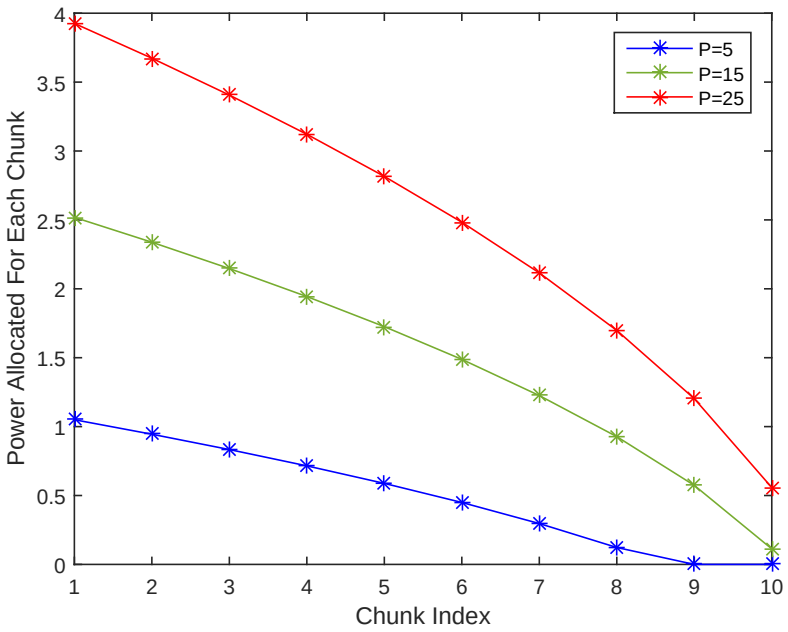


Figure 3.2: Optimal subchannel power allocation under total power constraint

3.3.3 A toy example

This example is adapted from [LP76]. We assume that there are total 10 independent channels and 10 chunks. The variance of each chunk is $(10, 9, \ldots, 1)$ and the covariance of channel noise is identity. The total power is $p_{\mathrm{T}} = 5,\ 15,\ 25$. The optimal power allocation for each chunk as shown in Figure 3.2,

We can see from Figure 3.2, when the total power allowed to be transmitted on the channel is not enough, for example $p_{\mathrm{T}} = 5$, the two chunks which have smallest variances will not be transmitted even if the bandwidth is enough.

3.4 Per Subchannel Power Constraints

In this section, we present the optimal precoding matrix design under per-subchannel power constraint. In the following, we assume now, again without loss of generality, that the subchannels are indexed by decreasing SNR: $\frac{p_1}{\sigma_1^2} \geqslant \frac{p_2}{\sigma_2^2} \geqslant \ldots \geqslant \frac{p_{n_{\mathrm{SC}}}}{\sigma_{n_{\mathrm{SC}}}^2}$. The power used for transmission on subchannel i is $\sum_{j=1}^{n_{\mathrm{Ck}}} g_{ij}^2 \lambda_j$, which corresponds to the ith diagonal element of $G\Lambda G^{\mathrm{T}}$. Therefore, the per-subchannel power constraints can be written as

$$\forall\ i \in \{1, \ldots, n_{\mathrm{SC}}\},\ \left(G\Lambda G^{\mathrm{T}}\right)_{i,i} \leqslant p_i. \tag{3.34}$$

The function to be minimized is the same as Eq. (3.5). The Lagrangian of this constrained optimization problem is thus

$$\begin{aligned}&\mathcal{L}(G,H,\gamma) = \\ &\quad \mathrm{tr}\left(\Lambda - 2HG\Lambda + HG\Lambda G^T H^T + HNH^T\right) \\ &\quad + \sum_{i=1}^{n_{\mathrm{SC}}} \gamma_i \left(\left(G\Lambda G^{\mathrm{T}}\right)_{i,i} - p_i\right),\end{aligned}$$

where $\gamma = (\gamma_1, \ldots, \gamma_{n_{\mathrm{SC}}})^{\mathrm{T}}$ is now a vector of Lagrange multipliers. For a given precoding matrix G, the optimum decoding matrix H is the same as in (3.29) and the objective function can again be expressed as

$$\varepsilon = \mathrm{tr}\left(\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right), \tag{3.35}$$

see (3.16). Now, introducing

$$G' = N^{-\frac{1}{2}}G, \tag{3.36}$$

Eq. (3.35) becomes

$$\begin{aligned}\varepsilon &= \mathrm{tr}\left(\left(I + \left(N^{\frac{1}{2}}G'\Lambda^{\frac{1}{2}}\right)^T N^{-1}\left(N^{\frac{1}{2}}G'\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) \\ &= \mathrm{tr}\left(\left(I + \left(G'\Lambda^{\frac{1}{2}}\right)^T \left(G'\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right),\end{aligned} \tag{3.37}$$

which has to be minimized with the constraints

$$\forall\, i \in \{1, \ldots, n_{\mathrm{SC}}\},\ \left(N^{\frac{1}{2}}G'\Lambda G'^{\mathrm{T}}N^{\frac{1}{2}}\right)_{i,i} = p_i. \tag{3.38}$$

This constraint may be rewritten as

$$G'\Lambda G'^{\mathrm{T}} = S, \tag{3.39}$$

with

$$S = \begin{pmatrix} p_1/\sigma_1^2 & * & * & * \\ * & p_2/\sigma_2^2 & * & * \\ * & * & \ddots & * \\ * & * & * & p_{n_{\mathrm{SC}}}/\sigma_{n_{\mathrm{SC}}}^2 \end{pmatrix}. \tag{3.40}$$

Thus, the per-subchannel power constraint enforces a structure to the matrix $G'\Lambda G'^{\mathrm{T}}$, namely it imposes that its diagonal elements are given by $s_i = \frac{p_i}{\sigma_i^2}$, while the off-diagonal elements, represented as $*$ can assume any real value.

Assume that some G' minimizing (3.37) with the constraint expressed by (3.39) has been found, then the optimal precoding matrix is $G = N^{\frac{1}{2}}G'$ and the corresponding H is found using (3.29). For this reason, one considers first the problem of finding the optimal precoding matrix G' with constraints on the signal-to-noise ratio (3.39) that minimizes (3.37). One can thus define an *equivalent channel* [LP76] with per-subchannel power constraints corresponding to the SNRs of the original subchannels and uncorrelated unit-variance noise components.

An important property of the equivalent channel is shown below.

Lemma 1. *[LP76] Consider a precoding matrix $\widetilde{G}$ leading to a given value ε of the distortion (3.37). For any $n_{SC} \times n_{SC}$ orthogonal matrix Z, the precoding matrix $G' = Z\widetilde{G}$ leads to the same distortion ε.*

As a consequence, one can consider the following approach (first introduced in [LP76]) to minimize (3.37) with the constraint (3.39). First, one searches a precoding matrix $\widetilde{G}$ that satisfies the total equivalent channel power constraint defined as the sum of the SNRs of all subchannels. This can be solved using the results of Section 3.3. Since the resulting precoding matrix does not necessarily satisfy the per-subchannel power constraints (3.39), one searches an orthogonal matrix Z such that $Z\widetilde{G}$ satisfies the per-subchannel power constraints (3.39). Sufficient conditions on the vector of eigenvalues $\widetilde{m} = (\widetilde{m}_1, \ldots, \widetilde{m}_{n_{\mathrm{SC}}})^T$ of $\widetilde{G}\Lambda\widetilde{G}^{\mathrm{T}}$ are provided in [MOA11, 9.B.2] to guarantee the existence of such matrix Z. Introducing the vector $s = \left(p_1/\sigma_1^2, \ldots, p_{n_{\mathrm{SC}}}/\sigma_{n_{\mathrm{SC}}}^2\right)^T$, the conditions are expressed in the following theorem.

Theorem 2. *[MOA11, 9.B.2]* If the entries of s and $\widetilde{m}$, arranged in non-increasing order

$\widetilde{m}_1 \geq \cdots \geq \widetilde{m}_{n_{\mathrm{SC}}}$, $s_1 \geq \cdots \geq s_{n_{\mathrm{SC}}}$, satisfy

$$\sum_{i=1}^{k} s_i \leqslant \sum_{i=1}^{k} \widetilde{m}_i \tag{3.41}$$

for all $k = 1, 2, \ldots, n_{\mathrm{SC}} - 1$ and

$$\sum_{i=1}^{n_{\mathrm{SC}}} s_i = \sum_{i=1}^{n_{\mathrm{SC}}} \widetilde{m}_i \tag{3.42}$$

then there exists a Hermitian matrix with diagonal s and vector of eigenvalues $\widetilde{m}$.

In practice, one finds $\widetilde{G}$ and evaluates $\widetilde{m}$ as shown in Section 3.4.1. If the sufficient conditions of Theorem 2 are satisfied, there exists an orthogonal matrix Z such that the diagonal of $Z\widetilde{G}\Lambda\widetilde{G}^{\mathrm{T}}Z^T$ is s. Several techniques are available to obtain Z in this case [ZZ95, VA99]. If the sufficient conditions are not satisfied, a suboptimal numerical method to obtain Z has been proposed in [LP76]. An optimal alternative approach is proposed in [PLC04] for the dual problem of power minimization under a per-channel MSE constraint and adapted in our context in Section 3.4.3.

3.4.1 Evaluation of $\widetilde{m}$

To find $\widetilde{G}$ and the related $\widetilde{m}$, consider the minimization of (3.37) with the total SNR constraint

$$s_{\mathrm{eq}} = \sum_{i=1}^{n_{\mathrm{SC}}} p_i/\sigma_i^2. \tag{3.43}$$

From the result of Section 3.3, since the equivalent channel has uncorrelated unit-variance noise components, the solution of this problem is

$$\widetilde{G} = \begin{pmatrix} \mathrm{diag}(\widetilde{g}_1, \ldots, \widetilde{g}_\ell) & 0 \\ 0 & 0 \end{pmatrix} \tag{3.44}$$

where $\widetilde{G} \in \mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}}$ and $\ell \leqslant \min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$ is the largest integer satisfying

$$\lambda_\ell \;\geqslant\; \gamma, \;\; \sqrt{\gamma} = \frac{\sum_{i=1}^{\ell} \sqrt{\lambda_i}}{s_{\mathrm{eq}} + \ell}, \;\; \text{and} \;\; \widetilde{g}_i = \sqrt{\frac{\sqrt{\lambda_i/\gamma} - 1}{\lambda_i}}. \tag{3.45}$$

As a consequence, $\widetilde{G}\Lambda\widetilde{G}^{\mathrm{T}}$ is a diagonal matrix, with eigenvalues (and vector of diagonal elements) $\widetilde{m}$ with entries given by

$$\widetilde{m}_i = \widetilde{g}_i^2 \lambda_i = \sqrt{\frac{\lambda_i}{\gamma}} - 1. \tag{3.46}$$

3.4.2 When the conditions of Theorem 2 are satisfied

In this case, Z can be computed using the approach described in [ZZ95, VA99]. Combining (3.44) and (3.36), the optimal precoding matrix for the initial problem becomes

$$G = N^{\frac{1}{2}} Z \widetilde{G}. \tag{3.47}$$

3.4.3 When the conditions of Theorem 2 are not satisfied

In that case, the multi-level water-filling approach proposed in [PLC04, Section VI] is used to split the vector of variances and the vector of SNR constraints into subvectors on which the conditions of Theorem 2 are tested again. If they are not satisfied the subvectors are split again in a recursive way. A solution necessarily exists since these conditions are satisfied when the size of the subvectors is 1.

We describe the optimal power allocation procedure (called *OptimalPrecoding*) in Algorithm 3.1. Its inputs are the vectors $\lambda = (\lambda_1 \dots \lambda_{n_{\mathrm{SC}}})$ of subband variances and $s = (s_1 \dots s_{n_{\mathrm{SC}}})$ of SNR constraints[1], both with components sorted in decreasing order. Its output is the optimal precoding matrix G'. We assume that the following four algorithms are available. *OptTotalPower* computes the optimal precoding matrix (3.44) and power allocation (3.46) under total power constraint. *CheckSuffCond* verifies whether the sufficient conditions (3.41) in Theorem 2 are satisfied. If this is not the case, it returns the largest index k such that $\sum_{i=1}^{k} s_i > \sum_{i=1}^{k} \widetilde{m}_i$. *SHIE* (Structured Hermitian Inverse Eigenvalue) computes the orthogonal transform matrix Z.

At the output of Algorithm 3.1, the optimal precoding matrix is block diagonal and consists

[1] To simplify presentation, one assumes here that $n_{\mathrm{Ck}} = n_{\mathrm{SC}}$. If this is not the case, one may zero-pad the vector of subband variances (when $n_{\mathrm{Ck}} < n_{\mathrm{SC}}$) or drop $n_{\mathrm{Ck}} - n_{\mathrm{SC}}$ components of low variance (when $n_{\mathrm{Ck}} > n_{\mathrm{SC}}$).

Algorithm 3.1 G' = OptimalPrecoding(λ,s)

```
1   i = 1 % Initial number of subblocks
2   G' = [] % Initialize G' as an empty matrix
3   do
4       if i = 1
5        k_(i) = 1, τ_(i) = n_SC
6       else
7        k_(i) = τ_(i−1) + 1 % Split position
8         τ_(i) = n_SC
9       end
10      do
11          % Find largest subvectors that
12          % satisfy Conditions (3.41) and (3.42)
13          λ_(i) = (λ_{k_(i)}, ..., λ_{τ_(i)}), s_(i) = (s_{k_(i)}, ..., s_{τ_(i)})
14          (G̃_(i), m̃_(i)) = OptTotalPower(λ_(i), s_(i))
15          (v, τ_(i)) = CheckSuffCond(m̃_(i),s_(i))
16      while v is false
17      Z_(i) = SHIE(m̃_(i), s_(i))
18      G'_(i) = Z_(i) G̃_(i)
19      i = i + 1 % Increase number of subblocks
20  while τ_(i−1) < n_SC
21  n_SB = i − 1 % Final number of subblocks
```

of n_{SB} submatrices

$$G' = \begin{pmatrix} Z_{(1)}\widetilde{G}_{(1)} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & Z_{(i)}\widetilde{G}_{(i)} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & Z_{(n_{\mathrm{SB}})}\widetilde{G}_{(n_{\mathrm{SB}})} \end{pmatrix}. \tag{3.48}$$

In the loop 10-16, *OptimalPrecoding* tries to find the largest subvectors $\lambda_{(i)}$ and $s_{(i)}$ such that the sufficient conditions (3.41) and (3.42) are satisfied with the matrix $\widetilde{G}_{(i)}$ designed in such a way that the total power constraint on these subvectors is satisfied. The transform matrix $Z_{(i)}$ is then evaluated.

OptimalPrecoding is a multi-level water-filling algorithm, in which the inverse γ^{-1} of the Lagrange multiplier γ in (3.45) represents the water level. Consider a transform subblock $Z_{(i)}\widetilde{G}_{(i)}$ associated to the subvectors $\lambda_{(i)}$ and $s_{(i)}$ for which the conditions of Theorem 2 are satisfied. In [PLC04, Appendix D], it is shown that the water level of an upper level subblock

(corresponding to large chunk variance and large channel SNR) is not less than that of a lower level subblock (corresponding to smaller chunk variance and channel SNR).

Contrary to the total power constrained case, the optimal precoding matrix is in general block-diagonal, see (3.48). Consequently, chunks may be mixed together and transmitted over several subchannels.

The *OptimalPrecoding* algorithm may be relatively complex, since the search for the transform subblock $Z_{(i)}\widetilde{G}_{(i)}$ always starts with the subvectors $\left(\lambda_{k_{(i)}}, \ldots, \lambda_{n_{\mathrm{sc}}}\right)$ and $\left(s_{k_{(i)}}, \ldots, s_{n_{\mathrm{sc}}}\right)$ with $k_{(i)} = \tau_{(i-1)} + 1$. In what follows, $k_{(i)}$ is called the i-th split position. The size of these vectors is progressively reduced until (3.41) and (3.42) are satisfied. In the worst case, this may require $n_{\mathrm{sc}} - \tau_{(i-1)}$ iterations, and as many evaluations of the corresponding optimal precoding matrix under total power constraint. As a consequence, the complexity to find all the split positions in the worst case (when λ and s are split into n_{SC} components) is $O\left(n_{\mathrm{SC}}^3\right)$, see [PLC04, AppendixD].

In the proposed scheme, the optimal scaling matrix replaces the power allocation and the Hadamard transform performed by SoftCast. Once the design has been performed, and since in most of the cases, the optimal precoding matrix G' is block diagonal, the overhead related to the multiplication by G' of each chunk vector is comparable to that of a scaling followed by an Hadamard transform and remains limited. The optimal precoding and decoding matrix design requires the knowledge of chunk variances, which are available at transmitter, and need to be sent to receivers as metadata, see Section 3.6.2. The characteristics of each subchannel need also to be known at transmitter. This information may be fed back by the receivers. In case of transmission to several receivers, as is the typically case in LVC schemes, the precoding matrix design in transmitter will be shown in Section 3.5. The precoding matrix will then be mismatched with the channels of most receivers. Nevertheless, as will be shown in Section 3.6.5, provided that each receiver adopts the decoding matrix adapted to the precoding matrix and to its actual channel conditions (which can be estimated e.g. using the pilot carriers of the OFDM scheme), the performance loss compared to a perfectly matched situation is rather small.

3.5 Precoding Matrix Design in multicast scenarii

In this section, we consider the precoding matrix design problem for a multiuser scenario, where a single SoftCast encoded stream is transmitted to k different users, each of which experiences different channel conditions. This type of problem has been considered, *e.g.*, in [KR13] in the context of relay-assisted multicast. A min-max problem formulation is considered, where the aim is to design the precoding and decoding matrices so as to minimize the worst MSE among receivers. Here, our aim is to minimize the average MSE among receivers.

More precisely, the transmitter sends some SoftCast encoded stream in n_{SC} subchannels. For receiver i, the covariance matrix of the noise is referred to as N_i. As in the single-user case, N_i is assumed to be diagonal for $i \in \{1, \ldots, k\}$. These channel models are called *general multicast channels* in what follows. Moreover we also introduce *linearly degraded multicast channels*, in which one has

$$N_i = \alpha_i N_{\mathrm{ref}} \tag{3.49}$$

where N_{ref} is a diagonal matrix with diagonal elements $\left(\sigma^2_{\mathrm{ref},1}, \ldots, \sigma^2_{\mathrm{ref},n_{\mathrm{SC}}}\right)$ and the α_is are positive coefficients. In this case, the noise variance of the j-th subchannel of user i is $\alpha_i \sigma^2_{\mathrm{ref},j}$.

The precoding matrix G is the same for all users. Assuming that each receiver knows G, Λ, and N_i, it may use the optimal decoding matrix obtained from (3.10), $\bar{H}_i = \Lambda G^T \left(G\Lambda G^T + N_i\right)^{-1}$. Then, from (3.16), one gets the average distortion among receivers as

$$\begin{aligned}
\varepsilon_{\mathrm{T}} &= \tfrac{1}{k} \sum_{i=1}^{k} \varepsilon_i \\
&= \tfrac{1}{k} \sum_{i=1}^{k} \mathrm{tr}\left(\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1} \Lambda\right) \\
&= \tfrac{1}{k} \mathrm{tr}\left(\left(\sum_{i=1}^{k} \left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\right) \Lambda\right).
\end{aligned} \tag{3.50}$$

The problem considered now is to design G so as to minimize (3.50), assuming that all N_i are known at transmitter side (they may be fed back by the receivers, when k is not too large). First, the precoding matrix design for linearly degraded multicast channels is considered in Section 3.5.1. Then the case of general multicast channels is considered in Section 3.5.1.2. In what follows, the chunk indexing is such that $\lambda_1 \geqslant \cdots \geqslant \lambda_{n_{\mathrm{Ck}}}$.

3.5.1 Multicast scenario with linearly degraded multicast channels

In this section, we address the precoding matrix design problem under total power constraint and per-subchannel power constraint in the case of the linearly degraded multicast channels.

3.5.1.1 Total Power Constraint

Our problem is to find a matrix G that minimizes

$$\varepsilon_{\mathrm{T}} = f_{\Lambda}\left(\mathrm{diag}\left(\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)\right)\right) \tag{3.51}$$

where

$$\Phi_{\mathrm{T}}:\ A \in \mathbb{R}^{n_{\mathrm{SC}}\times n_{\mathrm{Ck}}} \rightarrow \sum_{i=1}^{k}\left(I + A^{T}N_i^{-1}A\right)^{-1} \in \mathbb{R}^{n_{\mathrm{Ck}}\times n_{\mathrm{Ck}}} \tag{3.52}$$

$$f_{\Lambda} : \mathbf{u} \in \mathbb{R}^{n_{\mathrm{Ck}}} \rightarrow \frac{1}{k}(\lambda_1 u_1 + \cdots + \lambda_{n_{\mathrm{Ck}}} u_{n_{\mathrm{Ck}}}) \in \mathbb{R}_{+}.$$

with the total power constraint or with the per-subchannel power constraint.

As in the single-user case, since the λ_is are assumed in decreasing order, f_{Λ} is minimized when the components of its argument $\mathbf{u}$, which are the diagonal elements of $\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)$ in (3.51), are in increasing order. Then, in this situation f_{Λ} is a Schur-concave function (see Section 3.3.2). Let us now introduce the vector t of the eigenvalues of $\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)$ in increasing order, then from [MOA11, 3.A.1] and [MOA11, 9.B.1], one gets

$$f_{\Lambda}(t) \leqslant f_{\Lambda}\left(\mathrm{diag}\left(\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)\right)\right) = \varepsilon_{\mathrm{T}},$$

where the lower bound can be achieved if the argument of f_{Λ} is the vector of the eigenvalues of $\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)$, or if $\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)$ is a diagonal matrix and with diagonal elements in increasing order. Since

$$\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right) = \sum_{i=1}^{K}\left(I + \frac{1}{\alpha_i}\left(G\Lambda^{\frac{1}{2}}\right)^{T}N_{\mathrm{ref}}^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1},$$

a sufficient condition for $\Phi_{\mathrm{T}}\left(G\Lambda^{\frac{1}{2}}\right)$ to be diagonal with elements in increasing order is that $\left(G\Lambda^{\frac{1}{2}}\right)^{T}N_{\mathrm{ref}}^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ is a diagonal matrix with diagonal elements are in decreasing order.

In this case, from [PCL03, lemma12], under total power constraint (3.7), one deduces that

the optimal structure of G is

$$G = \begin{bmatrix} \mathrm{diag}\,(g_1 \dots g_\ell) & \mathbf{0}_{\ell\times(n_{\mathrm{Ck}}-\ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times\ell} & \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times(n_{\mathrm{Ck}}-\ell)} \end{bmatrix}. \tag{3.53}$$

where $\ell \leqslant \min(n_{\mathrm{SC}}, n_{\mathrm{Ck}})$.

Now, by using (3.53), (3.50) can be written as

$$\frac{1}{k}\mathrm{tr}\left(\sum_{i=1}^{k}\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^{T} N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) = \sum_{j=\ell+1}^{n_{\mathrm{Ck}}}\lambda_j + \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{\lambda_j}{1+m_j\sigma_{i,j}^{-2}}. \tag{3.54}$$

where $\sigma_{i,j}^2$ represents the variance of noise at jth subchannel of ith receiver, $m_j = g_j^2\lambda_j$ which represents the allocated power in the jth chunk and $j = 1, \dots, \ell$.

Accounting for the total power constraint $\sum_{j=1}^{\ell} m_j \leqslant p_{\mathrm{T}}$, one may introduce the Lagrangian associated to (3.51)

$$L_{\mathrm{T}} = \sum_{j=\ell+1}^{n_{\mathrm{Ck}}}\lambda_i + \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{\lambda_j}{1+m_j\sigma_{i,j}^{-2}} + \gamma\left(\sum_{j=1}^{\ell} m_j - p_{\mathrm{T}}\right), \tag{3.55}$$

where $\gamma \geqslant 0$ is the Lagrange multiplier.

Let us differentiate (3.54) respect to m_j and set it equal to zero, we get

$$\frac{1}{k}\sum_{i=1}^{k}\frac{\lambda_j\sigma_{i,j}^{-2}}{\left(1+m_j\sigma_{i,j}^{-2}\right)^2} = \gamma. \tag{3.56}$$

For some receivers, $m_j\sigma_{i,j}^{-2} > 1$ (SNR larger than one for the considered sub-channel) and for some others $m_j\sigma_{i,j}^{-2} < 1$ (SNR smaller than one). Let $\mathcal{K}_j^+$ and $\mathcal{K}_j^-$ the set of receivers for which $m_j\sigma_{i,j}^{-2} > 1$ and $m_j\sigma_{i,j}^{-2} < 1$, respectively. Then one may rewrite (3.56) approximately as

$$\frac{1}{k}\left(\frac{\lambda_j}{m_j^2}\sum_{i\in\mathcal{K}_j^+}\sigma_{i,j}^2 + \sum_{i\in\mathcal{K}_j^-}\lambda_j\sigma_{i,j}^{-2}\right) = \gamma, \tag{3.57}$$

from which one deduces

$$\frac{\lambda_j}{m_j^2}\sum_{i\in\mathcal{K}_j^+}\sigma_{i,j}^2 = k\gamma - \sum_{i\in\mathcal{K}_j^-}\lambda_j\sigma_{i,j}^{-2}.$$

For all subchannel indexes j such that $k\gamma - \sum_{i\in\mathcal{K}_i^-} \lambda_j \sigma_{i,j}^{-2} > 0$, one gets

$$m_j = \sqrt{\frac{\lambda_j \sum_{i\in\mathcal{K}_j^+} \sigma_{i,j}^2}{k\gamma - \sum_{i\in\mathcal{K}_j^-} \lambda_j \sigma_{i,j}^{-2}}} \tag{3.58}$$

for the others, one should take $m_j = 0$. When there are too many receivers for which the jth subchannel is poor, no power is allocated to that subchannel. Finally, γ is chosen such that

$$\sum_{j=1}^{\ell} m_j = p_{\mathrm{T}}.$$

One obtains a relatively complex water-filling problem where ℓ and γ have to be adjusted so as to minimize ε_{T} and $k\gamma - \sum_{i\in\mathcal{K}_i^-} \lambda_j \sigma_{i,j}^{-2} > 0$. A simplified solution is shown below.

Assuming high SNR after power allocation for all subchannels and receiver, *i.e.*, $\frac{m_j}{\sigma_{i,j}^2} \gg 1$, (3.56) becomes

$$\frac{1}{k}\frac{\lambda_j}{m_j^2}\sum_{i=1}^{k} \sigma_{i,j}^2 = \gamma,$$

and one gets

$$m_j = \frac{\sqrt{\lambda_j \frac{1}{k}\sum_{i=1}^{k} \sigma_{i,j}^2}}{\sqrt{\gamma}}. \tag{3.59}$$

Moreover since $\sum_{j=1}^{\ell} m_j = p_{\mathrm{T}}$, one gets

$$\sqrt{\gamma} = \frac{\sum_{j=1}^{\ell} \sqrt{\lambda_j \frac{1}{k}\sum_{i=1}^{k} \sigma_{i,j}^2}}{p_{\mathrm{T}}}, \tag{3.60}$$

and

$$m_j = \frac{\sqrt{\lambda_j \frac{1}{k}\sum_{i=1}^{k} \sigma_{i,j}^2}}{\sum_{j=1}^{\ell} \sqrt{\lambda_j \frac{1}{k}\sum_{i=1}^{k} \sigma_{i,j}^2}} p_{\mathrm{T}},$$

and the scaling factor g_j in (3.53) can be computed as

$$g_j = \sqrt{\frac{m_j}{\lambda_j}}. \tag{3.61}$$

This analytical solution requires that the SNR of all subchannels is high enough after power allocation, which in turns requires a large p_{T}. From (3.59), we can see that the variances of the noise to be considered in the precoding matrix design is the average of the variances of the

noise for each receivers. This is no more the case when p_{T} is not large enough, and one has to resort to a numerical solution of the water-filling problem.

3.5.1.2 Per-subchannel power constraint

The procedure to compute the precoding matrix in that case is similar to that in Section 3.4. We assume that the subchannels are indexed such that $p_1/\sigma^2_{\mathrm{ref},1} \geqslant \cdots \geqslant p_{n_{\mathrm{SC}}}/\sigma^2_{\mathrm{ref},n_{\mathrm{SC}}}$.

One assumes first that the eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are larger than one for $i = 1, \ldots, k$. Then 3.50 becomes

$$\begin{aligned}
\varepsilon_{\mathrm{T}} &= \frac{1}{k}\sum_{i=1}^{k} \mathrm{tr}\left(\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) \\
&\approx \frac{1}{k}\mathrm{tr}\left(\sum_{i=1}^{k}\left(\left(\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right)\right) \\
&= \frac{1}{k}\mathrm{tr}\left(\left(\sum_{i=1}^{K}\left(\frac{1}{\alpha_i}\left(G\Lambda^{\frac{1}{2}}\right)^T N_{\mathrm{ref}}^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\right)\Lambda\right) \\
&= \frac{1}{k}\mathrm{tr}\left(\left(\sum_{i=1}^{K}\alpha_i\right)\left(\left(G\Lambda^{\frac{1}{2}}\right)^T N_{\mathrm{ref}}^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) \\
&= \frac{1}{k}\mathrm{tr}\left(\left(\left(G\Lambda^{\frac{1}{2}}\right)^T\left(\left(\sum_{i=1}^{K}\alpha_i\right)N_{\mathrm{ref}}\right)^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) \\
&= \frac{1}{k}\mathrm{tr}\left(\left(\left(G\Lambda^{\frac{1}{2}}\right)^T\left(\sum_{i=1}^{K}N_i\right)^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right). \qquad (3.62)
\end{aligned}$$

Now, introducing

$$N_o = \sum_{i=1}^{K} N_i/k,$$

(3.62) becomes

$$\varepsilon_{\mathrm{T}} = \mathrm{tr}\left(\left(\left(G\Lambda^{\frac{1}{2}}\right)^T N_o^{-1}\left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right). \qquad (3.63)$$

Here again, one has to consider the average channel among users. Then introducing the *equivalent channel* as in Section 3.4, $G'_o = N_o^{-\frac{1}{2}}G$, (3.63) becomes

$$\varepsilon_{\mathrm{T}} = \mathrm{tr}\left(\left(\left(G'_o\Lambda^{\frac{1}{2}}\right)^T\left(G'_o\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right),$$

and the per-subchannel power constraint (3.34) becomes

$$\forall\ i \in \{1, \ldots, n_{\mathrm{SC}}\},\ \left(N_o^{\frac{1}{2}} G'_o \Lambda G'^{\mathrm{T}}_o N_o^{\frac{1}{2}}\right)_{i,i} = p_i. \tag{3.64}$$

This constraint may be rewritten as

$$G'_o \Lambda G'^{\mathrm{T}}_o = S_o, \tag{3.65}$$

with

$$S_o = \begin{pmatrix} p_1/\sigma^2_{o,1} & * & * & * \\ * & p_2/\sigma^2_{o,2} & * & * \\ * & * & \ddots & * \\ * & * & * & p_{n_{\mathrm{SC}}}/\sigma^2_{o,n_{\mathrm{SC}}} \end{pmatrix}. \tag{3.66}$$

The following computation procedure to find G is the same as that in Section 3.4. At first the precoding matrix $\widetilde{G}_o$ under total power constraint which is $\sum_{i=1}^{n_{\mathrm{Sc}}} p_i/\sigma^2_{o,n_{\mathrm{SC}}}$ and unit variance noise is computed using (3.61). Then, an orthogonal matrix Z_o has to be found that satisfies the per-subchannel power constraint (3.66). At the end $G = N_o^{\frac{1}{2}} Z_o \widetilde{G}_o$.

The second situation is when the eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)$ are smaller than one for $i = 1, \ldots, k$. Then 3.50 becomes

$$\begin{aligned} \varepsilon_{\mathrm{T}} &= \frac{1}{k} \sum_{i=1}^{k} \mathrm{tr}\left(\left(I + \left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right)^{-1} \Lambda\right) \\ &\approx \frac{1}{k} \mathrm{tr}\left(\sum_{i=1}^{k} \left(\left(I - \left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right) \Lambda\right)\right) \\ &= \frac{1}{k} \mathrm{tr}\left(\left(kI - \left(G\Lambda^{\frac{1}{2}}\right)^T \sum_{i=1}^{k} N_i^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right) \Lambda\right) \\ &= \mathrm{tr}\left(\left(I - \left(G\Lambda^{\frac{1}{2}}\right)^T N_{\mathrm{L}}^{-1} \left(G\Lambda^{\frac{1}{2}}\right)\right) \Lambda\right), \end{aligned} \tag{3.67}$$

where

$$N_{\mathrm{L}}^{-1} = \frac{1}{k} \sum_{i=1}^{k} N_i^{-1}$$

with diagonal elements denoted as $\sigma^{-2}_{\mathrm{L},i}$, $i = 1, \ldots, n_{\mathrm{SC}}$. We can notice that in this situation, we can get (3.67) in the case of the general multicast channel model, presented in Section 3.5.2.2.

In what follows, the procedure to compute G is same as in the first situation. At first we compute a precoding matrix $\widetilde{G}_{\mathrm{L}} = N_{\mathrm{L}}^{-\frac{1}{2}} G$ in equivalent channel, at which subchannels are indexed such that $p_1\sigma_{\mathrm{L},1}^{-2} \geqslant \cdots \geqslant p_{n_{\mathrm{SC}}}\sigma_{\mathrm{L},n_{\mathrm{SC}}}^{-2}$. For the equivalent channel, ε_{T} becomes

$$\varepsilon_{\mathrm{T}} = \mathrm{tr}\left(\left(I - \left(\widetilde{G}_{\mathrm{L}}\Lambda^{\frac{1}{2}}\right)^{T}\left(\widetilde{G}_{\mathrm{L}}\Lambda^{\frac{1}{2}}\right)\right)\Lambda\right). \tag{3.68}$$

As in Section 3.5.1.1, the optimal structure of $\widetilde{G}_{\mathrm{L}}$ is

$$\widetilde{G}_{\mathrm{L}} = \begin{bmatrix} \mathrm{diag}\left(\widetilde{g}_1 \ldots \widetilde{g}_\ell\right) & \mathbf{0}_{\ell\times(n_{\mathrm{Ck}}-\ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times\ell} & \mathbf{0}_{(n_{\mathrm{SC}}-\ell)\times(n_{\mathrm{Ck}}-\ell)} \end{bmatrix}. \tag{3.69}$$

Now substituting (3.69) in (3.68) and denoting $\widetilde{g}_i^2\lambda_i$ as $\widetilde{m}_i$, one gets

$$\varepsilon_{\mathrm{T}} = \sum_{i=1}^{n_{\mathrm{Ck}}}\lambda_i - \sum_{i=1}^{\ell}\lambda_i\widetilde{m}_i. \tag{3.70}$$

Minimizing (3.70) under total power constraint $\sum_{i=1}^{n_{\mathrm{Ck}}}\widetilde{m}_i \leqslant \widetilde{p}_{\mathrm{T}}$ and $\widetilde{p}_{\mathrm{T}} = \sum_{i=1}^{n_{\mathrm{SC}}} p_i\sigma_{\mathrm{L},i}^{-2}$ is a Linear Programming problem. There is no simply analytical solution, but this problem can be solved numerically [BV04, Page 6]. Once we have found $\widetilde{m}_i$, we can compute $\widetilde{g}_i$.Then use the solution of Structured Hermitian Inverse Eigenvalue problem to find an orthogonal matrix Z_{L} to adapt the per-subchannel power constraints in the equivalent channel. At the end

$$G_{\mathrm{L}} = N_{\mathrm{L}}^{\frac{1}{2}} Z_{\mathrm{L}}\widetilde{G}_{\mathrm{L}}.$$

3.5.2 General multicast channels

In this section, the precoding matrix design for general multicast channels is considered. There are two situations considered. The first one is when the eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^{T} N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are larger than one. Then other one is that the eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^{T} N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are small.

3.5.2.1 Eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^{T} N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ larger than one

In this situation, under total power constraint, if $G\Lambda^{\frac{1}{2}}$ is an invertible matrix, one gets (3.62). Then G is computed from (3.61), but in this case, the subchannels have to be indexed in such a

way that $\sigma^2_{o,1} \leqslant \cdots \leqslant \sigma^2_{o,n_{\mathrm{SC}}}$. Next, under per-subchannel power constraints, when $n_{\mathrm{SC}} < n_{\mathrm{Ck}}$, the last $n_{\mathrm{Ck}} - n_{\mathrm{SC}}$ chunks are discarded. Moreover with the assumption that the eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are larger than one, from (3.61), no chunk will be discarded, therefore we can assume $G\Lambda^{\frac{1}{2}}$ as invertible and one gets 3.62. The following computation of G under per-subchannel power constraint is the same as in Section 3.5.1.2 and subchannels are indexed such that $p_1/\sigma^2_{o,1} \geqslant \cdots \geqslant p_{n_{\mathrm{SC}}}/\sigma^2_{o,n_{\mathrm{SC}}}$.

3.5.2.2 Eigenvalues of $\left(G\Lambda^{\frac{1}{2}}\right)^T N_i^{-1}\left(G\Lambda^{\frac{1}{2}}\right)$ are small

In this situation, we also can get (3.67). Then the computation of G under per-subchannel power constraint is the same as in Section 3.5.1.2. Under total power constraint, subchannels are indexed in such a way that $\sigma^{-2}_{\mathrm{L},1} \geqslant \cdots \geqslant \sigma^{-2}_{\mathrm{L},n_{\mathrm{SC}}}$. Then, from the computation in equivalent channel at Section 3.5.1.2, one can deduce that the objective function to be minimized in here is

$$\begin{aligned} \min_{m_i} \quad & \varepsilon_{\mathrm{T}} = \textstyle\sum_{i=1}^{n_{\mathrm{Ck}}} \lambda_i - \sum_{i=1}^{\ell} \lambda_i \sigma^{-2}_{\mathrm{L},i} m_i, \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{\ell} m_i \leqslant p_{\mathrm{T}} \end{aligned}$$

where $m_i = g_i^2 \lambda_i$.

3.6 Simulations

3.6.1 Simulation conditions

Table 3.2 summarizes the additional notations used in this section.

In the following simulations, one assumes that video has to be transmitted over an in-home power line channel to one or several receivers with different channel characteristics. The frequency range is from 1.8 MHz to 86.13 MHz, which is the same range considered by the HomePlug Alliance in the HomePlug AV2 specification [YAA+13]. The spacing between subchannels is $f_{\mathrm{SC}} = 24.414$ kHz and the maximum number of subchannels that may be used for data transmission is $\eta_{\mathrm{SC}} = 3217$. Not all subchannels are allowed for data transmission. In OFDM-based PLT systems like AV2, typically SNRs per subchannel are available. A realization of the individual subchannel SNRs is represented in Figure 3.3, which relates to a bad SISO link from ETSI STF 477 database. Assuming that each subchannel is corrupted by independent

Variable	Value	Signification
n_{F}	8	nb of frames per GoP
$n_{\mathrm{C}} \times n_{\mathrm{R}}$		Frame size
$n_{\mathrm{c}} \times n_{\mathrm{r}}$		Chunk size
n_{Ck}		nb of chunks in a GoP
f_{SC}	24.414 kHz	Spacing between subcarriers
η_{SC}	3217	Nb of available subchannels
β_{r}	30%	Nyquist filter roll-off
n_{VSC}		number of virtual subchannels
r_{SC}	37560	per-subchannel rate in symb/s
r_{Ck}		source chunk rate in chunk/s
ρ_{Ck}		per-subchannel chunk rate in chunk/s/subchannel
v_{Ck}		nb of chunks a subchannel can transmit per GoP
n_{gCk}		nb of group of chunks

Table 3.2: Additional notations

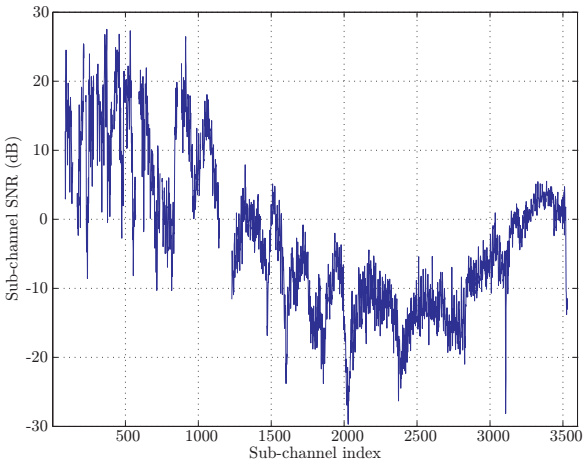


Figure 3.3: SNR as a function of the subchannel index for the considered PLT channel

white Gaussian noise sequences and considering the maximum per-subchannel transmission powers provided in [YAA+13], one may deduce the noise variance for each subchannel.

Considering a Homeplug AV2-type physical layer adapted to SoftCast, in which analog QAM and root-raised-cosine Nyquist filters with $\beta_{\mathrm{r}} = 30\,\%$ roll-off are used, one obtains a per-subchannel transmission rate

$$r_{\mathrm{SC}} = \frac{2 f_{\mathrm{SC}}}{1 + \beta_{\mathrm{r}}}, \tag{3.71}$$

which is here equal to $r_{\mathrm{SC}} = 37.56 \times 10^3$ real-valued symbols per second.

We consider the luminance component of a video source emitting r_{F} frames per second. The size of each frame is $n_{\mathrm{C}} \times n_{\mathrm{R}}$. To determine the way the chunks should be transmitted on the subchannel, one has to consider a given chunk size $n_{\mathrm{c}} \times n_{\mathrm{r}}$. With this choice, the video source

chunk rate is

$$r_{\mathrm{Ck}} = r_{\mathrm{F}} \frac{n_{\mathrm{R}} n_{\mathrm{C}}}{n_{\mathrm{r}} n_{\mathrm{c}}}, \tag{3.72}$$

and the per-subchannel chunk rate is obtained from (3.71) as

$$\rho_{\mathrm{Ck}} = \frac{r_{\mathrm{SC}}}{n_{\mathrm{r}} n_{\mathrm{c}}}. \tag{3.73}$$

Clearly, if

$$r_{\mathrm{Ck}} > \eta_{\mathrm{SC}} \rho_{\mathrm{Ck}},$$

a certain amount of chunks in each GoP can not be transmitted due to channel bandwidth constraints. Considering GoPs of constant size n_{F}, the number of chunks a subchannel can transmit for the duration of a GoP is

$$v_{\mathrm{Ck}} = \frac{n_{\mathrm{F}}}{r_{\mathrm{F}}} \frac{r_{\mathrm{SC}}}{n_{\mathrm{r}} n_{\mathrm{c}}}. \tag{3.74}$$

For the typical values of the parameters considered in these simulations, $v_{\mathrm{Ck}} > 1$, *i.e.*, several chunks may be transmitted on the same subchannel for the duration of a GoP.

To apply the precoding and decoding matrix design techniques, two approaches may be considered. The first is to consider v_{Ck} replicas of each subchannel, each replica (*virtual* subchannel) only being able to transmit a single chunk. With this approach, during the transmission of a GoP, one has thus

$$\begin{aligned} n_{\mathrm{VSC}} &= \eta_{\mathrm{SC}} v_{\mathrm{Ck}} \\ &= \eta_{\mathrm{SC}} \frac{n_{\mathrm{F}}}{r_{\mathrm{F}}} \frac{r_{\mathrm{SC}}}{n_{\mathrm{r}} n_{\mathrm{c}}} \end{aligned}$$

virtual subchannels available for the transmission of

$$n_{\mathrm{Ck}} = n_{\mathrm{F}} \frac{n_{\mathrm{R}} n_{\mathrm{C}}}{n_{\mathrm{r}} n_{\mathrm{c}}}$$

chunks. This approach is optimal, but leads to huge precoding and decoding matrices of $n_{\mathrm{VSC}} \times n_{\mathrm{Ck}}$ components. The alternative approach, adopted here, is to partition the n_{Ck} chunks

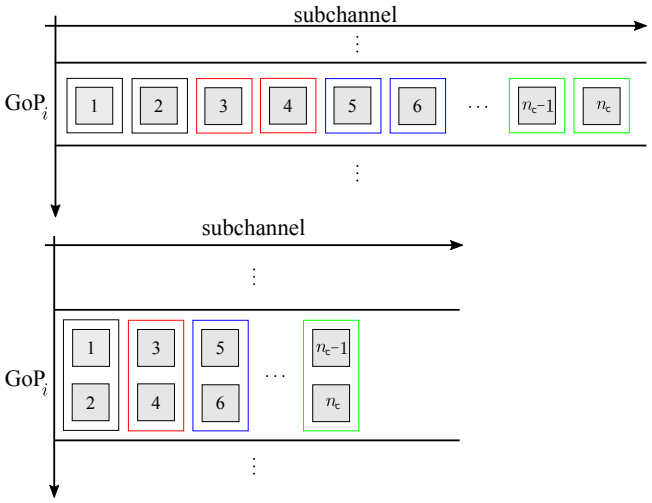


Figure 3.4: Organizations of the transmission of chunks of the i-th GoP: (top) each subchannel is duplicated into v_{Ck} virtual subchannels, each being able to transmit a single chunk per GoP; (bottom) chunks of similar variance are gather into groups of v_{Ck} chunks, each group of chunk being transmitted over a dedicated subchannel

in groups of v_{Ck} chunks of similar variance. There are thus

$$n_{\mathrm{gCk}} = \frac{n_{\mathrm{Ck}}}{v_{\mathrm{Ck}}}$$

groups of chunks. Then, v_{Ck} precoding (and decoding) matrices are designed considering the n_{gCk} chunks of same index in the groups of chunks. This second approach is suboptimal, but requires for each GoP the design of v_{Ck} smaller precoding and decoding matrices of size $\eta_{\mathrm{SC}} \times n_{\mathrm{gCk}}$. Figure 3.4 illustrates the two possible ways chunks may be transmitted over the available subchannels.

A set of video sequences with different characteristics in terms of spatial and temporal resolutions, and in terms of contents has been considered, namely the video sequences of classes B, C, D, E, and F used by the MPEG committee for the standardization of HEVC [OSS+12]. Their characteristics are given in Table 3.3. For the sake of simplicity, only the luminance component of these sequences has been considered. The precoding and decoding matrix design methods could be extended to color sequences using a proper weighting of the distortion of the chrominance components.

For each video sequence, the chunk size is chosen in such a way that n_{r} divides n_{R} and n_{c} divides n_{C}. When $r_{\mathrm{Ck}} \leqslant \eta_{\mathrm{SC}}\rho_{\mathrm{Ck}}$, only the best subchannels are selected. Moreover, for each GoP, $\left\lfloor \frac{n_{\mathrm{F}}}{r_{\mathrm{F}}} \frac{r_{\mathrm{SC}}}{n_{\mathrm{r}} n_{\mathrm{c}}} \right\rfloor$ chunks are transmitted on each subchannel. When $r_{\mathrm{Ck}} > \eta_{\mathrm{SC}}\rho_{\mathrm{Ck}}$, the chunks

Name	Frame rate [Hz]	Nb Frames	$n_{\mathrm{gCk}} \times v_{\mathrm{Ck}}$
Class B: Frames of 1920×1080 px, 13824 chunks of 40×30 px			
Kimonol	24	240	1393×10
BasketballDrive	50	500	2765×5
BQ Terrace	60	600	3217×4
Cactus	50	500	2765×5
ParkScene	24	240	1393×10
Class C: Frames of 832×480 px, 3328 chunks of 32×30 px			
PartyScene	50	500	555×6
BQMall	60	600	666×5
BasketballDrill	50	500	555×6
RaceHorses	30	300	333×10
Class D: Frames of 416×240 px, 832 chunks of 32×30 px			
BQSquare	60	600	167×5
RaceHorses	30	300	84×10
BlowingBubbles	50	500	139×6
BasketballPass	50	500	139×6
Class E: Frames of 1280×720 px, 6144 chunks of 40×30 px			
FourPeople	60	600	1536×4
Jonny	60	600	1536×4
KristenAndSara	60	600	1536×4
Class F: Frames of 1280×720 px, 6144 chunks of 40×30 px			
SlideShow	20	500	512×12

Table 3.3: Characteristics of the considered video sequences and corresponding chunk organization.

of least variance are dropped. The values of the parameters n_{gCk} and v_{Ck} are also provided in Table 3.3. In the simulations, always the best n_{gCk} subchannels are used.

3.6.2 Metadata

Metadata have to be transmitted without errors to the receiver so that it is able to decode the noisy precoded chunk vectors. A transmission of the precoding and decoding matrices should clearly be avoided, due to their size. Both matrices should be re-estimated at the receiver from metadata sent by the transmitter.

Considering total or per subchannel power constraints, the overhead due to metadata is similar to that of SoftCast, except for the information related to the channel characteristics. Considering a GoP of n_{F} frames containing $n_{\mathrm{Ck}} = n_{\mathrm{F}} \frac{n_{\mathrm{R}} n_{\mathrm{C}}}{n_{\mathrm{r}} n_{\mathrm{c}}}$ chunks, apart from the GoP size, frame size, and chunk size, a vector of n_{Ck} bits has to be sent first to indicate the transmitted chunks. Then, at most n_{Ck} chunk mean values and variances have to be sent. The channel

characteristics have to be known at the transmitter and at the receiver. In a point-to-point communication scenario, when they are fed back by the receiver to the transmitter, their retransmission as metadata is not required. In a multicast scenario, the transmitter may consider average channel characteristics from various receivers. These average characteristics have then to be sent as metadata to the receivers. This requires the transmission of n_{SC} noise variances. Per subchannel power constraints are usually fixed and are transmitted at most once during initialization of the communication.

Assume that the chunk mean values and variances as well as the channel characteristics are represented on 8 bits and that the metadata are channel coded with a rate 1/2 channel code. For Kimono1 with GoPs of 8 frames and the chunk characteristics in Table 3.3, one obtains a metadata rate of $13824 \times (1 + 8 + 8) \times 3 \times 2 = 1.41$ Mb/s. If the channel characteristics are refreshed at the GoP rate, one gets an additional metadata rate of $3217 \times 8 \times 3 \times 2 = 0.15$ Mb/s. Considering video transmission over PLT channels, the transmission of channel-coded metadata would require dedicated subchannels and a rate of about 1 % of the total rate available in the context of HomePlug AV2 [YAA+13].

3.6.3 Total Power Constraint

A first set of simulation is performed considering only a total power constraint. For each video, the organization of the chunk transmission is that described in Table 3.3. A unit variance noise is considered on each subchannel, while the total transmission power has been adjusted in such a way that the subchannel SNR is 15 dB (good transmission condition) or 5 dB (poor transmission condition).

The power allocation method described in Section 3.3 is compared to that presented in [JK10b]. Simulation results in terms of average PSNR of the received sequences are reported in Table 3.4.

On the good channel, the two allocation methods perform similarly. For the poor channel the proposed allocation method clearly outperforms that considered in [JK10b], as expected, since the latter assumes a relatively high channel SNR. This conclusion is confirmed by Figure 3.5, which shows the evolution of the PSNR with the channel SNR for the Kimonol sequence for both power allocation techniques. When the channel SNR is 0 dB, a gain of 0.93 dB in PSNR

Class	Name	PSNR (dB) Good channel		PSNR (dB) Poor channel		
		Proposed	SoftCast	Proposed	SoftCast	Gain
B	Kimonol	49.60	49.59	40.64	40.35	0.29
	BasketballDrive	45.50	45.50	36.60	36.31	0.29
	BQ Terrace	42.26	42.26	33.69	33.45	0.24
	Cactus	44.56	44.55	35.60	35.36	0.24
	Park Scene	45.97	45.97	37.00	36.75	0.25
Average PSNR Class B		45.58	45.57	36.71	36.44	0.26
C	Party Scene	40.18	40.18	31.22	31.01	0.21
	BQMall	42.90	42.89	33.95	33.66	0.29
	BaskeballDrill	44.55	44.53	35.57	35.28	0.29
	Race Horses	41.68	41.67	32.66	32.47	0.19
Average PSNR Class C		42.33	42.32	33.35	33.11	0.25
D	BQ Square	38.17	38.16	29.14	28.87	0.27
	Race Horses	41.08	41.07	32.04	31.80	0.24
	BlowingBubbles	41.08	41.07	32.11	31.85	0.26
	BasketballPass	42.73	42.72	33.76	33.51	0.25
Average PSNR Class D		40.77	40.76	31.76	31.51	0.26
E	FourPeople	48.61	48.61	39.58	39.2	0.38
	Jonny	50.07	50.07	41.06	40.69	0.37
	KristenAndSara	48.44	48.43	39.39	39.02	0.37
Average PSNR Class E		49.04	49.04	40.01	39.64	0.37
F	SlideShow	43.71	43.70	34.26	34.06	0.2

Table 3.4: Total Power Constraint

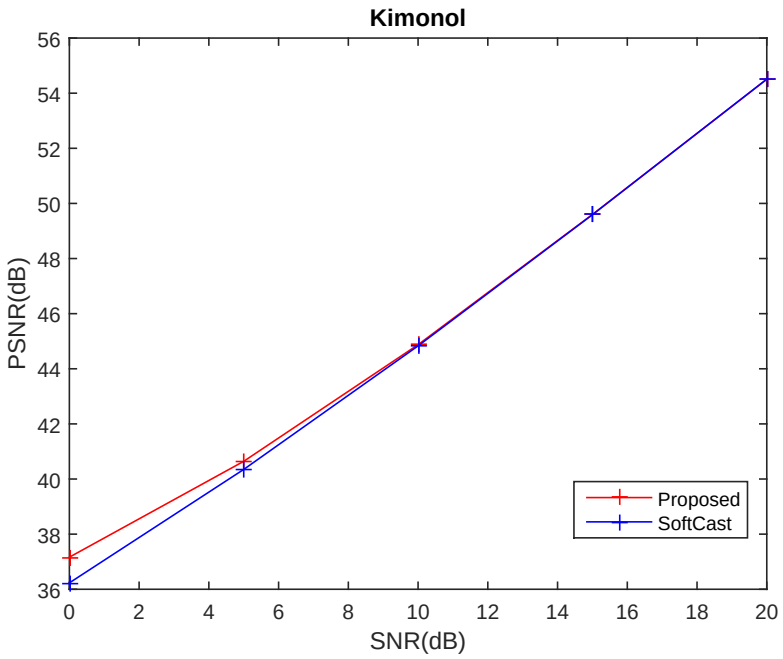


Figure 3.5: Evolution of the PSNR with the channel SNR for the Kimono1 sequence considering a total power constraint

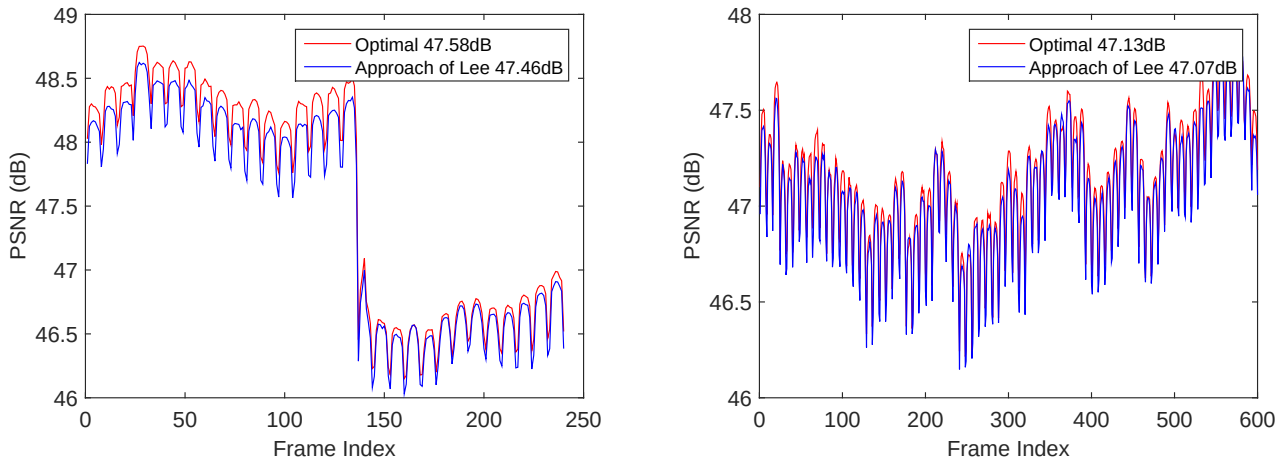
is observed.

3.6.4 Per Subchannel Power Constraints

Per-subchannel power constraints are now considered using the channel model described in Section 3.6.1. We compare the optimal allocation method proposed here with heuristic precoding matrix design approach in [LP76].

In Section 3.4, we have seen that an precoding matrix design with per-subchannel power constraints can be formulated as a design problem with an equivalent channel with per-subchannel SNR constraints and unit subchannel noise variances. As a consequence, for the simulations, one assumes again unit noise variance on all subchannels and adjust the transmission power of chunks on each subchannel to have subchannel SNR matching those described in Figure 3.3.

Figure 3.6 illustrates the PSNR of frames of video sequences Kimono1 and Fourpeople when considering the heuristic precoding matrix design approach proposed by Lee in [LP76] and optimal allocation method proposed here. An average gain of 0.12 dB and 0.06 dB are respectively obtained with the proposed optimal design. Similar gains are observed with the other sequences.



(a) Kimonol (b) Fourpeople

Figure 3.6: PSNR comparison for the frames of Kimono1 and Fourpeople, considering per-subcahnnel power constrained precoding matrix design with the approach of [LP76] and the proposed approach.

3.6.5 Mismatch

In this section, one considers the impact of a channel mismatch on the proposed design technique under total power constraint and per subchannel power constraints. This represents scenarii such as a transmission to receivers with different channel characteristics or a precoding matrix design with outdated information on the channel characteristics.

To illustrate the effect of channel mismatch, one assumes that the total power constraint or per-subchannel power constraints are fixed (provided by the standard, *e.g.* PLT channel, Figure 3.7), but that the channel noise or subchannel noises used for the precoding matrix design are not equal to actual channel noise or subchannel noises of receiver. Let N_{D} be the diagonal noise covariance matrix used by the transmitter for the precoding matrix design and let N_{A} be the covariance matrix of the actual noise affecting the subchannels. In case of channel mismatch, one has $N_{\mathrm{D}} \neq N_{\mathrm{A}}$. Both N_{D} and N_{A} are assumed to be perfectly known by the receiver, but the transmitter is only assumed to know N_{D}. This is realistic in a point-to-point scenario when the receiver feeds back channel state information to the transmitter. In a point-to-multipoint scenario, different N_{A}s are experienced by each receiver, and the transmitter has to select some average or worst-case channel characteristic N_{D}, which has to be transmitted to the receivers as meta-information.

Denote as G_{D} the optimal precoding matrix evaluated using (3.47) when considering N_{D}.

If the noise covariance matrix is N_{A}, one may still use the mismatched decoding matrix from (3.29) $H_{\mathrm{D}} = \Lambda G_{\mathrm{D}}^T \left(G_{\mathrm{D}} \Lambda G_{\mathrm{D}}^T + N_{\mathrm{D}}\right)^{-1}$ to get

$$\begin{aligned}
\varepsilon_1 &= \mathrm{tr}\left(\Lambda - 2H_{\mathrm{D}}G_{\mathrm{D}}\Lambda + H_{\mathrm{D}}G_{\mathrm{D}}\Lambda G_{\mathrm{D}}^T H_{\mathrm{D}}^T + H_{\mathrm{D}}N_{\mathrm{A}}H_{\mathrm{D}}^T\right) \\
&= \mathrm{tr}\left(\Lambda - H_{\mathrm{D}}G_{\mathrm{D}}\Lambda + H_{\mathrm{D}}\left(N_{\mathrm{A}} - N_{\mathrm{D}}\right)H_{\mathrm{D}}^T\right) \\
&= \mathrm{tr}\left(\left(I + \left(G_{\mathrm{D}}\Lambda^{\frac{1}{2}}\right)^T N_{\mathrm{D}}^{-1}\left(G_{\mathrm{D}}\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right) + \mathrm{tr}\left(H_{\mathrm{D}}^T H_{\mathrm{D}}\left(N_{\mathrm{A}} - N_{\mathrm{D}}\right)\right). \qquad (3.75)
\end{aligned}$$

The first term in (3.75) corresponds to the MSE obtained without mismatch. The term $\mathrm{tr}\left(H_{\mathrm{D}}^T H_{\mathrm{D}}\left(N_{\mathrm{A}} - N_{\mathrm{D}}\right)\right)$ may be positive or negative. When N_{A} is "smaller" than N_{D}, *i.e.*, the subchannels are less noisy than expected, the MSE ε_1 will be smaller than expected during the precoding matrix design. When N_{A} is "larger" than N_{D}, the channel is worse than expected, and the MSE is larger than expected. In both cases, the MSE variation is commensurate with the difference between N_{A} and N_{D}.

Alternatively, one may consider the decoding matrix adapted to N_{A} which expression is deduced from (3.10) as

$$H_{\mathrm{A}} = \Lambda G_{\mathrm{D}}^T \left(G_{\mathrm{D}} \Lambda G_{\mathrm{D}}^T + N_{\mathrm{A}}\right)^{-1} \qquad (3.76)$$

to get

$$\varepsilon_2 = \mathrm{tr}\left(\left(I + \left(G_{\mathrm{D}}\Lambda^{\frac{1}{2}}\right)^T N_{\mathrm{A}}^{-1}\left(G_{\mathrm{D}}\Lambda^{\frac{1}{2}}\right)\right)^{-1}\Lambda\right). \qquad (3.77)$$

The decoding matrix H_{A} is designed to minimize the reconstruction MSE considering that the precoding matrix is G_{D} and the channel noise covariance matrix is N_{A}. As a consequence, one has $\varepsilon_2 \leqslant \varepsilon_1$.

In the following simulations, the performance of video transmission under total power constraint and under per-subchannel power constraint for two receivers who have different channel conditions are tested. The channel for receiver1 consists in the n_{gCk} subchannels with the largest SNR of bad SISO link (see Section 3.6.1). The variance of noise in subchannel can be deduced from the power constraint and the SNR. The power constraint is the integral of the Power Spectrum Density (PSD) over the bandwidth (here is 24.414kHz in PLT). The PSD is -50 dBm/Hz from 1.8 to 30 MHz and -80 dBm/Hz from 30 MHz up to 100 MHz, which is shown in Figure 3.7. From [YAA+13], only carriers from 1.8 to 86.13 MHz are supported for

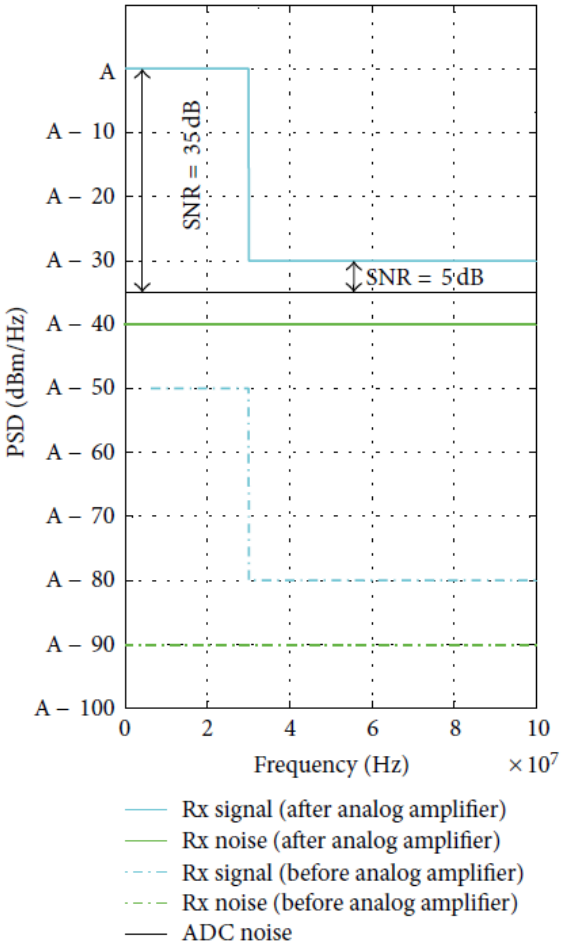


Figure 3.7: The figure from [YAA+13]. Power Spectrum Density (PSD) no power back-off in PLT

communication, but this issue is not considered in our simulation. The receiver2 has channel 2 where the variance of noise in each subchannel is 5 times larger than that of receiver1. The general multicast channels is considered in Section 3.6.5.2.

3.6.5.1 Total Power Constraint

In this section, we simulate video transmission under total power constraint for 2 receivers in linearly degraded multicast channels. The transmission of video sequence Kimonol and BQMall are tested. The total power constraint is the sum of per-subchannel power constraints. The precoding matrix design approach presented in Section 3.5.1.1 is applied at transmitter. For both receivers, the optimal decoding matrix (3.76) is used. Then the received video performance (PSNR) of different receivers are shown in Tab. 3.5, and are also compared to those in the case of point-to-point communication situation (no mismatch). The results show that the PNSR

Video	receiver 1 PSNR(dB)		receiver 2 PSNR(dB)	
	Point-to-Point	MultiCast	Point-to-Point	MultiCast
Kimonol	55.68	55.54	50.66	49.84
BQMall	47.66	47.63	41.64	41.31

Table 3.5: mismatch under total power constraint in linearly degraded multicast channel model

Video	receiver 1 PSNR(dB)		receiver 2 PSNR(dB)	
	Point-to-Point	MultiCast	Point-to-Point	MultiCast
Kimonol	47.58	47.38	42.16	42.15
BQMall	44.91	44.84	38.29	38.29

Table 3.6: Linearly degraded multicast channel per-subchannel power constraint

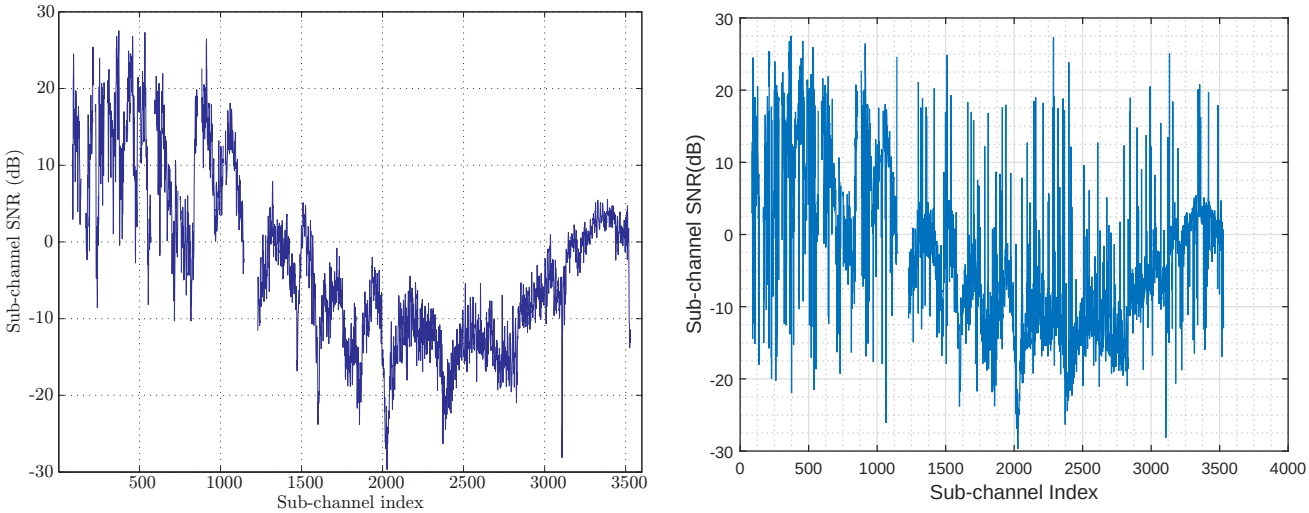
loss is rather small.

3.6.5.2 Per-subchannel Power Constraint

In this section, at first, the linearly degraded multicast channels is considered. The proposed method in Section 3.5.1.2 is considered here. The transmssion of the video sequences Kimonol and BQMall are considered. Then the PSNR results of simulation are compared to those under point-to-point communication as shown in Tab 3.6. We can see the PSNR loss in mismatch is negligible.

Then we test the robustness of the proposed optimal method under general multicast channels. we consider the video sequence BQMall. In multicast scenario, we assume receiver 1 has the channel as shown in Figure 3.8a and refereed to channel 1 and receiver 2 has channel 2, which is generated by flipping randomly the subchannels of channel 1. Several channels can be generated by flipping a fraction of the subchannel, see Figure 3.8b for an example.

We consider three methods to compute the precoding matrix. In the first one, the method proposed in Section 3.5.2.1 is considered, at which the per-subchannel power constraint in equivalent channel is $p_i/\sigma_{o,i}^2$. In the second one, we compute the average per-subchannel SNRs of channel1 and of channel2. Then the subchannels are ordered in decreasing order of the average SNRs. Next, by using the method proposed in Section 3.4, we compute the precoding matrix for the equivalent channel with average per-subchannel SNR constraints. At end, the precoding matrix of equivalent channel is multiplied by an average covariance of noise obtained as division of the per-subchannel power constraint by the average SNR to adapt the per-subchannel power constraint. The third scheme consist in using the SNR constraints of channel 1 as a reference



(a) PLT Channel 1. Generated from the data of [ETS15] (b) Channel 2 10% flip from Channel 1

Figure 3.8: SNR as a function of the subchannel index

to compute precoding matrix for equivalent channel and by using the method proposed in Section 3.4. For receivers, the optimal decoding matrix is applied at decoder. The simulation results of PSNR for different receivers under different random flipping probability are shown in Table. 4.5.

We can see the method proposed in Section 3.5.2.1 is better than the others, the performance only decreases 0.3dB each time the probability of flipping increases by 5%. Figure 3.9 shows that in the case of probability of flipping is 10% for receiver2's channel, the reconstructed first frames of BQMall for receiver 2 in multicast by considering different per-subchannel power constraints as a reference in equivalent channel. Full sequences are available at *https://drive.google.com/drive/folders/1umL5qeN35kT54JQhcu0-1SpANm8go5UE?usp=sharing*.

3.7 Conclusions

In the context of LVC, this chapter addresses the problem of optimal precoding and decoding matrix design when the video has to be transmitted over parallel additive white Gaussian noise (AWGN) channel with different characteristics. One has considered first that a total transmission power budget has to be allocated between the subchannels. Then, additional per-subchannel power constraints have been considered to address transmission contexts such as PLT channels or multi-antenna systems. At last, the transmission for several receivers who

receiver 1	receiver 2	Precoder	receiver1 PSNR(dB)	receiver2 PSNR(dB)
Chan.1	5% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.66	44.66
		Ave. SNR	41.76	41.78
		Chan.1	44.91	38.00
—	10% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.38	44.38
		Ave.SNR	39.80	39.71
		Chan.1	44.91	35.32
—	15% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.09	44.09
		Ave.SNR	38.29	38.29
		Chan.1	44.91	33.65
—	20% flip from Chan.1	$p_i/\sigma^2_{o,i}$	43.77	43.77
		Ave.SNR	37.16	37.17
		Chan.1	44.91	32.34

Table 3.7: Mismatch For video sequence BQMall in general multicast channel



(a) no mismatch, PSNR =44.42dB



(b) $p_i/\sigma^2_{o,i}$, PSNR =44.05dB



(c) Ave. SNR, PSNR = 40.61dB



(d) Chan. 1, PSNR = 36.40dB

Figure 3.9: Reconstructed first frames of BQMall for receiver 2 in multicast by considering different per-subchannel power constraints as a reference in equivalent channel. The probability of flipping is 10%. (a) no mismatch, (b) $p_i/\sigma^2_{o,i}$, (c) Average SNR, (d) Channel 1.

have different channel conditions is considered.

In the first case, small gains compared to the reference SoftCast allocation have been observed essentially at low channel SNRs. In the second case, we have considered the same problem as in [LP76], but we provide an optimal solution which is adapted from the proposed Multi-level water-filling approach to resolve total transmission power minimization with per-subchannel MSE constraints problem in [PLC04, PJ07]. In case of mismatch of the precoding matrix with the actual channel characteristics, the benefits of the LVC paradigm are also preserved in our proposed solution.

Chapter 4

Sub-Optimal Power Allocation Under Per-subchannel Power Constraint

In Chapter 3, an optimal power allocation approach under per-subchannel power constraint has been proposed, however it has a high complexity $O(n^3)$. In this chapter four alternative suboptimal precoding are considered in the case of per-subchannel power constraints. In all cases, the resulting precoding matrix G still satisfies (3.34) but may lead to a larger MSE. In Section 4.1 a simple power allocation method is provided. Next, in Section 4.2, by inferring the split positions, the computation cost may be significantly reduced with respect to optimal algorithm. Nevertheless, some parameters have to be suitably set up. Then the other two suboptimal power allocation schemes, which do not require any parameter tuning, are presented in Section 4.3 and 4.4. In Section 4.5, the limits of suboptimal methods will be shown. Finally in Section 4.6, the performance comparison and complexity comparison of proposed suboptimal methods with optimal method (Section 3.4) is shown.

4.1 Simple Chunk Scaling

In the first method called Simple Chunk Scaling (SCS), the chunk of largest variance is transmitted over the subchannel with the best SNR, the chunk with the second largest variance is sent over the subchannel with the second best SNR, *etc.*, Figure 4.1. To fit the per-subchannel power constraints, the coefficients of the i-th chunk are multiplied by $g_{\mathrm{SCS},i} = \sqrt{p_i/\lambda_i}$, $i = 1 \ldots n_{\mathrm{SC}}$. This allocation is clearly suboptimal but can be easily evaluated. It can be considered as the

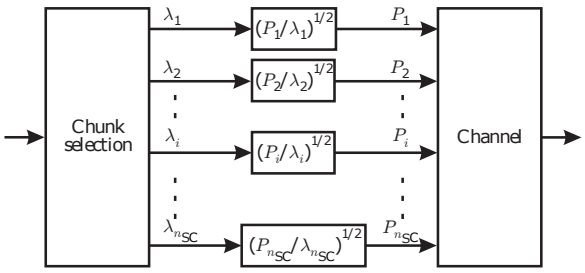


Figure 4.1: Simple chunk scaling

most straightforward and natural extension of Parcast [LHL+14b].

4.2 Power Allocation with Inferred Split Position (PAISP)

In this suboptimal approach, one tries to infer the split positions of Algorithm 3.1. For that purpose, one first analyzes the change of SNR constraints along the subchannels and the change of variance along the chunks. Consider the subvectors $\lambda_{(i)} = \left(\lambda_{k_{(i)}}, \ldots, \lambda_{\tau_{(i)}}\right)$ and $s_{(i)} = \left(s_{k_{(i)}}, \ldots, s_{\tau_{(i)}}\right)$ of λ and s of length $\mu(i) = \tau_{(i)} - k_{(i)} + 1$ introduced at Line 13 of Algorithm 3.1. From (3.46), the sufficient conditions (3.41) in Theorem 2 that are checked at Line 15 of Algorithm 3.1 can be rewritten as

$$\sum_{j=k_{(i)}}^{k} \left(\sqrt{\frac{\lambda_j}{\gamma}} - 1\right) \geqslant \sum_{j=k_{(i)}}^{k} s_j$$

$$\sum_{j=k_{(i)}}^{k} \sqrt{\lambda_j} \frac{\sum_{j=k_{(i)}}^{\tau_{(i)}} s_j + \ell}{\sum_{j=k_{(i)}}^{k_{(i)}+\ell-1} \sqrt{\lambda_j}} \geqslant \sum_{j=k_{(i)}}^{k} (s_j + 1)$$

or

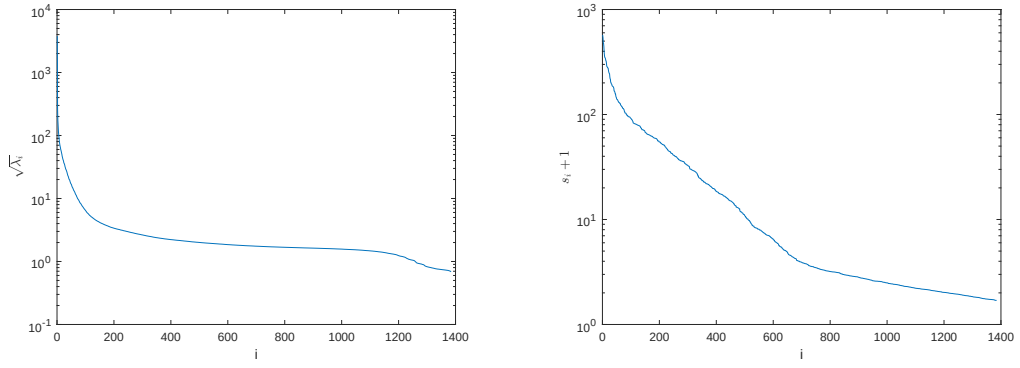
$$a_k \geqslant b_k \tag{4.1}$$

with

$$a_k = \frac{\sum_{j=k_{(i)}}^{k} \sqrt{\lambda_j}}{\sum_{j=k_{(i)}}^{k_{(i)}+\ell-1} \sqrt{\lambda_j}} \text{ and } b_k = \frac{\sum_{j=k_{(i)}}^{k} (s_j+1)}{\sum_{j=k_{(i)}}^{\tau_{(i)}} s_j+\ell} \tag{4.2}$$

for $k = k_{(i)}, \ldots, k_{(i)} + \ell - 2$, where $\ell \leqslant \mu(i)$ is the largest integer satisfying (3.45)[1]. Condition (3.42) corresponds to the total power constraint. It is satisfied by the design of $\widetilde{G}_{(s)}$.

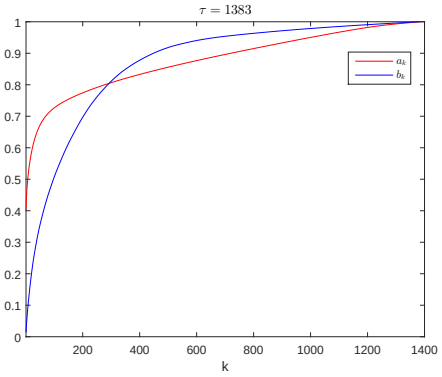
[1]The multi-level waterfilling is such that there may be components in the last subblock (lowest chunk variances) that are not transmitted and in that case $\ell \leqslant \mu(n_{\mathrm{SB}})$. For intermediate subblocks in the loop 10-16 of Algorithm 3.1, $\ell = \mu(i)$.



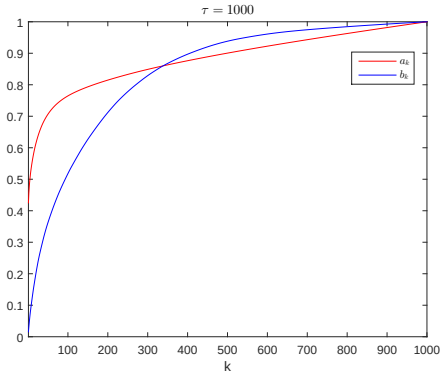
(a) (b)

Figure 4.2: (a) Standard deviations of a chunk vector of the first GoP of Kimonol (after reordering) and (b) Modified SNR constraints, *i.e.*, s_i+1, associated to the channel in Figure 3.3 (after reordering).

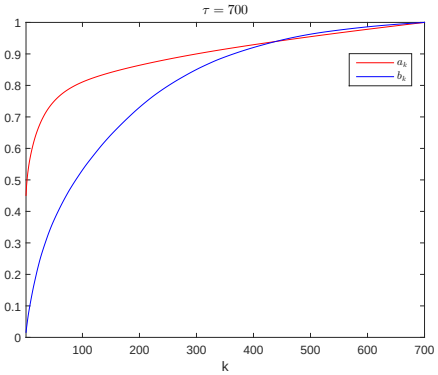
The conditions (4.1) are satisfied for example when the initial λ_is are very large compared to the other λ_is (the transform gain is large) while the SNRs s_i are more homogeneous. Consider, for example, Figures 4.2a and 4.2b which represent respectively the standard deviations of the components of a chunk vector (see Section 3.6.1) of the first GoP of the Kimonol video sequence and the corresponding modified subchannel SNR constraints, *i.e.*, s_i+1 (with $k_{(i)} = 1$, $\tau_{(i)} = n_{\mathrm{SC}} = 1383$). Figure 4.3a represents the evolution of a_k and b_k as functions of k. One observes that the conditions (4.1) are not satisfied for $k = 293$ up to $k = n_{\mathrm{SC}} - 1 = 1382$. This is consistent with the fact that on Figures 4.2a and 4.2b, the right part of the plot of the modified SNR constraint vector is less flat than the corresponding part of the plot of the vector of standard deviations. The optimal power allocation algorithm will then check the conditions (4.1) for the subvectors $\left(\lambda_1, \ldots, \lambda_{\tau_{(1)}}\right)$ and $\left(s_1, \ldots, s_{\tau_{(1)}}\right)$ with $\tau_{(1)} = n_{\mathrm{SC}} - 1$. Nevertheless, considering the plots in Figure 4.2, it is likely that the conditions (4.1) will again not be satisfied and are in fact likely to be satisfied for $\tau_{(1)}$ closer to 293 than to 1382. To get shorter subvectors on which the conditions (4.1) are more likely to be satisfied, instead of using $\tau_{(i)}$ obtained from *CheckSuffCond,* the idea of PAISP is to consider subvectors $(\lambda_1, \ldots, \lambda_\tau)$ and $(s_1, \ldots, s_\tau)$ with $\tau \leqslant \tau(i)$ to avoid several iterations in the loop 10-16 of Algorithm 3.1. However, this may result in a value of τ smaller than the optimal split position. Figure 4.3 represent the values of the vectors a and b obtained for several $\tau_{(i)}$s. The optimal $\tau_{(i)}$ is 535, larger than 293, but much less than $\tau_{(1)} = n_{\mathrm{SC}} - 1$ which would have been chosen in the next iteration of the loop 10-16 of Algorithm 3.1.



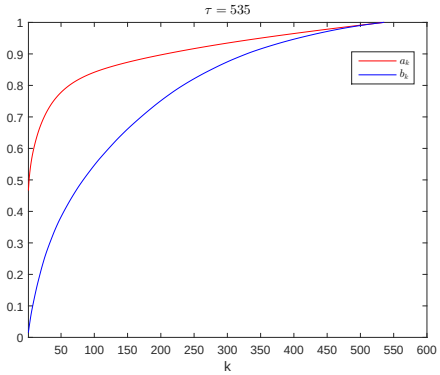
(a) $\tau = n_{\mathrm{SC}}$



(b) $\tau = 1000$



(c) $\tau = 700$



(d) $\tau = 535$

Figure 4.3: a_k and b_k for different values of τ

Algorithm 4.1 $G' = \text{PAISP}(\lambda, s, \alpha, \beta)$

```
1   μ = length(λ) % number of components of λ
2   [G̃, m̃] = OptTotalPower(λ, s)
3   [v, τ] = CheckSuffCond (m̃,s)
4   if v is true % Conditions (3.41) and (3.42) satisfied
5       Z= SHIE(m̃,s)
6       G' = ZG̃
7   else
8       if τ > αμ
9          τ = ⌈βμ⌉
10      end
11      λ(1) = (λ1 ... λτ), s(1) = (s1 ... sτ)
12      λ(2) = (λτ+1 ... λμ), s(2) = (sτ+1 ... sμ)
13      G'(1) =PAISP(λ(1), s(1),α, β)
14      G'(2) =PAISP(λ(2), s(2),α, β)
15  end
```

Power Allocation with Inferred Split Position (PAISP) algorithm is based on this observation and described in Algorithm 4.1. It takes initially $\lambda = (\lambda_1 \dots \lambda_{n_{\text{SC}}})$ and $s = (s_1 \dots s_{n_{\text{SC}}})$ of length $\mu = n_{\text{SC}}$ as inputs and two parameters $0 < \beta < \alpha < 1$. The way α and β have to be chosen is discussed in Section 4.6.2. When Conditions (3.41) and (3.42) are satisfied, the corresponding power allocation matrix is evaluated. If they are not satisfied, the largest index τ that violates Conditions (4.1) is evaluated. If $\tau > \alpha\mu$, then the chosen split index is reduced to $\tau = \beta\mu$. Else, if $\tau \leqslant \alpha\mu$, then τ is chosen as the split position. PAISP is then recursively called with inputs subvectors $\lambda_{(i)}$ and $s_{(i)}$ of length $\mu_{(i)}$, $i = 1, 2$ of λ and s. This avoids the repeated tests performed at Lines 13-15 of Algorithm 3.1.

To evaluate the complexity of PAISP to find all the split positions, the worst case is when $\tau = 1$ at Line 3 of Algorithm 4.1 at each recursion. In such case, there are n_{SC} recursions and in each recursion the complexity (mainly due to *OptTotalPower*) is proportional to the length μ of each subvector being considered. The total complexity is proportional to $\sum_{\mu=1}^{n_{\text{SC}}} \mu$ and hence is $O\left(n_{\text{SC}}^2\right)$.

The main drawback of PAISP is that α and β have to be properly tuned. The PAISP with Dicothomy in Section 4.3 and *Power Allocation with Local Power Adjustment* (PALPA) algorithm proposed in Section 4.4 do not need any parameter adjustment.

Algorithm 4.2 $G' = \text{PAISP}(\lambda, s)$

$\mu = \text{length}(\lambda)$ % number of components of λ
$\left[\widetilde{G},\, \widetilde{m}\right] = \text{OptTotalPower}(\lambda, s)$
$[v,\, \tau] = \text{CheckSuffCond}\ (\widetilde{m}, s)$
if v is true % Conditions (3.41) and (3.42) satisfied
  $Z = \text{SHIE}(\widetilde{m}, s)$
  $G' = Z\widetilde{G}$
else % Uses dichotomy find split position c
  $a = 1,\ b = \tau,\ c' = \frac{a+b}{2},\ \delta = b - a + 1$
  while $(\delta > 1)$
    $c = c'$
    $\lambda = (\lambda_1 \ldots \lambda_c),\ s = (s_1 \ldots s_c)$
    $\left[\widetilde{G},\, \widetilde{m}\right] = \text{OptTotalPower}(\lambda, s)$
    $[v,\, \tau] = \text{CheckSuffCond}\ (\widetilde{m}, s)$
    if v is true, $a = c,\ c' = \left\lfloor \frac{a+b}{2} \right\rfloor,\ \delta = |c - c'|$
    else $b = \tau,\ c' = \left\lfloor \frac{a+b}{2} \right\rfloor,\ \delta = |c - c'|$
    if $c' = 1,$ then $c = 1,\ \delta = 0$
  end
  $\lambda_{(1)} = (\lambda_1 \ldots \lambda_c),\ s_{(1)} = (s_1 \ldots s_c)$
  $\lambda_{(2)} = (\lambda_{c+1} \ldots \lambda_\mu),\ s_{(2)} = (s_{c+1} \ldots s_\mu)$
  $G'_{(1)} = \text{PAISP}(\lambda_{(1)}, s_{(1)})$
  $G'_{(2)} = \text{PAISP}(\lambda_{(2)}, s_{(2)})$
end

4.3 PAISP with Dichotomy

PAISP with Dicothomy takes initially $\lambda = (\lambda_1 \ldots \lambda_{n_{\text{SC}}})$ and $s = (s_1 \ldots s_{n_{\text{SC}}})$ of length $\mu = n_{\text{SC}}$ as inputs. At first the largest index τ that violates Conditions (3.41) is evaluated. Then PAISP searches the optimal split position at the interval $[a, b] = [1, \tau]$ by dichotomy. First, the midpoint $c = \frac{a+b}{2}$ is considered. If Conditions (3.41) are satisfied for $(\lambda_1 \ldots \lambda_c)$ and $(s_1 \ldots s_c)$, then PAISP updates $a = c$; Else the largest index τ that violates Conditions (3.41) is evaluated and PAISP updates $b = \tau$. These iterations are repeated until the difference between two successive midpoints is not larger than 1, see Algorithm 4.2.

To evaluate the complexity of PAISP to find all subvectors. The worst case is when $\tau = \mu - 1$ at each recursion of PAISP and when λ is split into n_{SC} components at the end. In such case, there are n_{SC} recursions and the total complexity is proportional to $\sum_{\mu=1}^{n_{\text{SC}}} \mu$ and hence is $O\left(n_{\text{SC}}^2\right)$.

4.4 Power Allocation with Local Power Adjustment

Power Allocation with Local Power Adjustment (PALPA) is an algorithm that takes initially $\lambda = (\lambda_1 \dots \lambda_{n_{\mathrm{SC}}})$ and $s = (s_1 \dots s_{n_{\mathrm{SC}}})$ as inputs. PALPA evaluates first the power allocated with a total power constraint. The resulting allocated power vector has entries $\widetilde{m}_i^{(0)}$, $i = 1, \dots, n_{\mathrm{SC}}$. If Conditions (3.41) and (3.42) are satisfied, the $\widetilde{m}_i^{(0)}$s are used to build the solution to the power allocation problem with per-subchannel power constraints using [ZZ95, VA99], see Section 3.4.2. Otherwise, let τ be the largest index for which Condition (3.41) is violated. As shown in what follows, the powers allocated to the chunk subvector $(\lambda_{\tau+1}, \dots, \lambda_{n_{\mathrm{SC}}})$ are then easily updated to match the total power constraint of the $n_{\mathrm{SC}} - \tau$ last subchannels, while satisfying Conditions (3.41). The corresponding part of the precoding matrix is then build, see Section 3.4.2. PALPA is then called iterative on $(\lambda_1 \dots \lambda_\tau)$ and $(s_1 \dots s_\tau)$ to build the remaining parts of the precoding matrix.

For the power allocation update, since τ is the largest index for which Condition (3.41) is violated, one has

$$\sum_{i=1}^{\tau} \widetilde{m}_i^{(0)} < \sum_{i=1}^{\tau} s_i \tag{4.3}$$

$$\sum_{i=1}^{k} \widetilde{m}_i^{(0)} \geqslant \sum_{i=1}^{k} s_i \tag{4.4}$$

for $k = \tau + 1, \dots, n_{\mathrm{SC}} - 1$ and

$$\sum_{i=1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} = \sum_{i=1}^{n_{\mathrm{SC}}} s_i. \tag{4.5}$$

From (4.3) and (4.5), one deduces that

$$\sum_{i=\tau+1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} > \sum_{i=\tau+1}^{n_{\mathrm{SC}}} s_i,$$

i.e., too much power has been allocated to the last $n_{\mathrm{SC}} - \tau$ components of λ. The total excess power is

$$\Delta^{(0)} = \sum_{i=\tau+1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} - \sum_{i=\tau+1}^{n_{\mathrm{SC}}} s_i. \tag{4.6}$$

The main idea of PALPA is to correct the values of $\widetilde{m}_i^{(0)}$, $i = \tau + 1, \dots, n_{\mathrm{SC}}$ to get $\widetilde{m}_i^{(1)}$, $i = \tau+1, \dots, n_{\mathrm{SC}}$ in such a way that the conditions (3.41) and (3.42) are valid for the subvectors

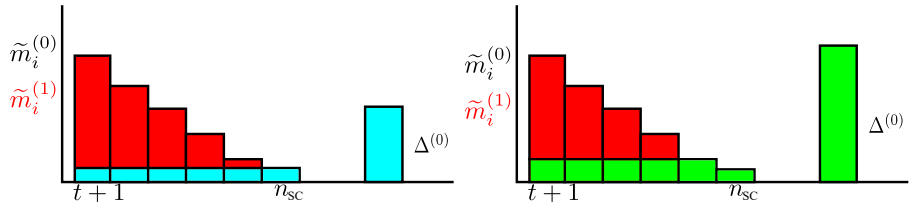


Figure 4.4: Initial ($\widetilde{m}_i^{(0)}$) and updated ($\widetilde{m}_i^{(1)}$) allocated powers when $\Delta^{(0)}$ is small (left) and when $\Delta^{(0)}$ is large (right)

$\left(\widetilde{m}_{\tau+1}^{(1)}, \ldots, \widetilde{m}_{n_{\rm SC}}^{(1)}\right)$ and $(s_{\tau+1}, \ldots, s_{n_{\rm SC}})$.

For that purpose, one evaluates first

$$\overline{\ell} = \max_{\tau+1 \leqslant \ell' \leqslant n_{\rm SC}} \ell'$$

such that for $i = \tau+1, \ldots, \ell'$,

$$\widetilde{m}_i^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\ell'+1}^{n_{\rm SC}} \widetilde{m}_j^{(0)}}{\ell' - \tau} \geqslant 0.$$

Since the $\widetilde{m}_i^{(0)}$s are decreasing, one may consider only the constraint

$$\widetilde{m}_{\ell'}^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\ell'+1}^{n_{\rm SC}} \widetilde{m}_j^{(0)}}{\ell' - \tau} \geqslant 0.$$

Then for $i = \tau+1, \ldots, n_{\rm SC}$, the updated allocated powers are

$$\widetilde{m}_i^{(1)} = \begin{cases} \widetilde{m}_i^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\overline{\ell}+1}^{n_{\rm SC}} \widetilde{m}_j^{(0)}}{\overline{\ell} - \tau} & \text{if } i \leqslant \overline{\ell} \\ 0 & \text{else.} \end{cases} \tag{4.7}$$

This correction corresponds to an increase of the water level, see Figure 4.4. It ensures that the source components with large variance are still allocated a larger power.

Proposition 1. *The power allocation ajustement performed by PALPA using (4.7) is such that for* $k = \tau+1, \ldots, n_{\rm SC}-1$,

$$\sum_{i=\tau+1}^{k} \widetilde{m}_i^{(1)} \geqslant \sum_{i=\tau+1}^{k} s_i$$

and

$$\sum_{i=\tau+1}^{n_{SC}} \widetilde{m}_i^{(1)} = \sum_{i=\tau+1}^{n_{SC}} s_i. \tag{4.8}$$

Proof. One first shows that the updated power allocation compensates the excess power, *i.e.*,

$$\Delta^{(0)} = \sum_{i=\tau+1}^{n_{\mathrm{SC}}} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)} \right).$$

Using (4.7), one has

$$\begin{aligned}
\sum_{i=\tau+1}^{n_{\mathrm{SC}}} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)} \right) &= \sum_{i=\tau+1}^{\bar{\ell}} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)} \right) + \sum_{i=\bar{\ell}+1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} \\
&= \sum_{i=\tau+1}^{\bar{\ell}} \left(\widetilde{m}_i^{(0)} - \left(\widetilde{m}_i^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\bar{\ell}+1}^{n_{\mathrm{SC}}} \widetilde{m}_j^{(0)}}{\bar{\ell} - \tau} \right) \right) + \sum_{i=\bar{\ell}+1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} \\
&= \sum_{i=\tau+1}^{\bar{\ell}} \frac{\Delta^{(0)} - \sum_{j=\bar{\ell}+1}^{n_{\mathrm{SC}}} \widetilde{m}_j^{(0)}}{\bar{\ell} - \tau} + \sum_{i=\bar{\ell}+1}^{n_{\mathrm{SC}}} \widetilde{m}_i^{(0)} \\
&= \Delta^{(0)}.
\end{aligned}$$

Then to show Proposition 1, one has $\forall k \in \{\tau+1, \ldots, n_{\mathrm{SC}} - 1\}$

$$\begin{aligned}
\sum_{i=1}^{\tau} \widetilde{m}_i^{(0)} + \sum_{i=\tau+1}^{k} \widetilde{m}_i^{(0)} &\geqslant \sum_{i=1}^{\tau} s_i + \sum_{i=\tau+1}^{k} s_i \\
\sum_{i=1}^{\tau} \widetilde{m}_i^{(0)} - \sum_{i=1}^{\tau} s_i &\geqslant \sum_{i=\tau+1}^{k} s_i - \sum_{i=\tau+1}^{k} \widetilde{m}_i^{(0)} \\
\sum_{i=\tau+1}^{k} \widetilde{m}_i^{(0)} - \sum_{i=\tau+1}^{k} s_i &\geqslant \Delta^{(0)}.
\end{aligned}$$

Now, one has

$$\begin{aligned}
\Delta^{(0)} &= \sum_{i=\tau+1}^{n_{\mathrm{SC}}} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)} \right) \\
&\geqslant \sum_{i=\tau+1}^{k} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)} \right),
\end{aligned}$$

Algorithm 4.3 $G' = \text{PALPA}\ (\lambda, s)$

$\mu = \text{length}(\lambda)$ % number of components of λ
$\left[\widetilde{G},\, \widetilde{m}\right] = \text{OptTotalPower}(\lambda, s)$
$[v, \tau] = \text{CheckSuffCond}\ (\widetilde{m}, s)$
if v is true % Conditions (3.41) and (3.42) satisfied
    $Z = \text{SHIE}(\widetilde{m}, s)$
    $G' = Z\widetilde{G}$
else
    $\Delta_{(2)} = \sum_{i=\tau+1}^{\mu} \widetilde{m}_i - \sum_{i=\tau+1}^{\mu} s_i$
    $\lambda_{(1)} = (\lambda_1 \dots \lambda_\tau),\ s_{(1)} = (s_1 \dots s_\tau)$
    $\widetilde{m}_{(2)}^{(0)} = (\widetilde{m}_{\tau+1} \dots \widetilde{m}_\mu)$
    $\lambda_{(2)} = (\lambda_{\tau+1} \dots \lambda_\mu),\ s_{(2)} = (s_{\tau+1} \dots s_\mu)$
    $G'_{(1)} = \text{PALPA}(\lambda_{(1)}, s_{(1)})$
    $G'_{(2)} = \text{LPA}(\widetilde{m}_{(2)}^{(0)}, \Delta_{(2)}, \lambda_{(2)}, s_{(2)})$
end

Algorithm 4.4 $G'_{(2)} = \text{LPA}(\widetilde{m}^{(0)}, \Delta^{(0)}, \lambda, s)$

$\mu' = \text{length}\left(\widetilde{m}^{(0)}\right)$
for $\ell' = \mu'$ down to 1
    if $\widetilde{m}_{\ell'}^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\ell'+1}^{\mu'} \widetilde{m}_j^{(0)}}{\ell'} \geqslant 0$ break;
end
for $i = 1$ to μ'
    if $i \leqslant \ell'$ then $\widetilde{m}_i^{(1)} = \widetilde{m}_i^{(0)} - \frac{\Delta^{(0)} - \sum_{j=\ell'+1}^{\mu'} \widetilde{m}_j^{(0)}}{\ell'}$
    else $\widetilde{m}_i^{(1)} = 0$
end
$Z_{(2)} = \text{SHIE}(\widetilde{m}^{(1)}, s)$
$G'_{(2)} = Z_{(2)}\text{diag}\left(\text{sqrt}\left(\widetilde{m}_1^{(1)}/\lambda_1, \dots, \widetilde{m}_{\mu'}^{(1)}/\lambda_{\mu'}\right)\right)$

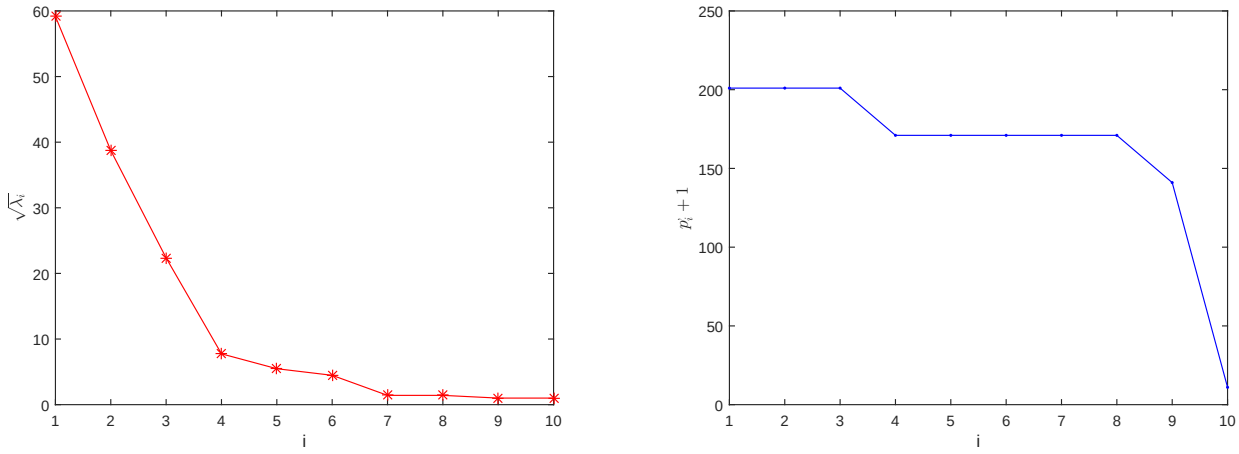
for all $k \in \{t+1, \dots, n_{\text{SC}} - 1\}$. Thus

$$\sum_{i=\tau+1}^{k} \widetilde{m}_i^{(0)} - \sum_{i=\tau+1}^{k} s_i \geqslant \sum_{i=\tau+1}^{k} \left(\widetilde{m}_i^{(0)} - \widetilde{m}_i^{(1)}\right)$$
$$\sum_{i=\tau+1}^{k} \widetilde{m}_i^{(1)} - \sum_{i=\tau+1}^{k} s_i \geqslant 0.$$

The proof of (4.8) follows the same lines. □

Algorithm 4.3 corresponds to the PALPA algorithm that calls the *Local Power Adjustment* (LPA) method (4.7) described in Algorithm 4.4. The latter evaluates also the precoding matrix $G'_{(2)}$ for the considered source subvector.

To evaluate the complexity of finding all the split positions with PALPA, the worst case is



(a) $\sqrt{\lambda_i}$ (b) $p_i + 1$

Figure 4.5: A toy example where the PAISP fails to estimate the splitting position.

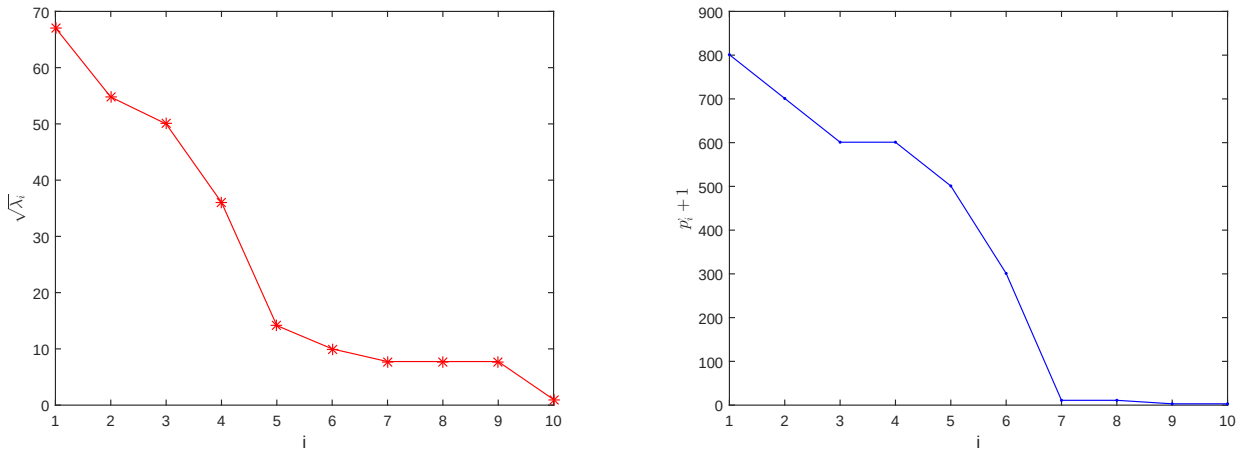
now obtained when $\tau = \mu - 1$ at Line 3 of Algorithm 4.3. At each recursion, the complexity is again mainly due to *OptTotalPower*, which is linear in the length μ of the vector to be processed. As a consequence, the total complexity is proportional to $\sum_{\mu=1}^{n_{\mathrm{SC}}} \mu$ and is again $O\left(n_{\mathrm{SC}}^2\right)$. From a practical point of view, it is better to modify Algorithm 4.3 in such a way to make it iterative rather than recursive. In our test, we use the iterative version of PALPA.

4.5 Limits of PAISP and PALPA

In this section, we consider two toy examples to show the drawbacks of PAISP (Section 4.2) and PALPA (Section 4.4).

For PAISP, intuition tells us that the optimal values of α and β depend on the variation of SNRs. If the variation of SNRs is fast at the end and slow at beginning (with respect to the standard deviations), then this method may fail to infer the best splitting position. For example, we consider the vectors of standard deivations and the vector of SNRs shown in Fig.4.5.

With these values of chunks standard deviations and SNR constraints, we ran the optimal power allocation algorithm and the two sub-optimal PAISP and PALPA. Then we compute the resulting MSE for the three of them. We foud that, while in this case PALPA will provide same mean square error as the optimal method, (namely, $\varepsilon = 12.66$), PAISP gives a higher distortion: we found $\varepsilon = 19.05$ with$\alpha = 0.8$ and $\beta = 0.5$ and $\varepsilon = 24.71$ with$\alpha = 0.5$ and $\beta = 0.3$. The



(a) $\sqrt{\lambda_i}$ (b) $p_i + 1$

Figure 4.6: A toy example where the PALPA works less better than optimal method. (4.6a) a vector of standard deviations; (4.6b) modified power constraint vector.

minimum MSE can be achieved with PAISP using $\alpha = 1$, but this simply means that PAISP works exactly as the optimal method.

In the case of PALPA, the allocation may fail when there are chunks with very low variances compared to the neighbor chunks' variance, and when the SNR constraints are flat around the split position. In this case, PALPA may allocate some power to chunks that would not be transmitted using the optimal allocation. An example is given in Fig.4.6.

As in the previous example, we ran the optimal and the suboptimal algorithms and computed the resulting distortion. We found that for the optimal method $\varepsilon = 39.27$, but for PALPA is $\varepsilon = 46.60$. We also observe that the optimal method would not allocate power to the last chunk, while PALPA does it. The power allocation m_i for each component under optimal method is : [1012.5 826.5 754.4 543.8 212.7 150.1 10 10 4 0]; under PALPA it is [1012.5 826.5 754.4 543.8 212.7 150.1 10 10 2 2]; under PAISP ($\alpha = 0.75$ $\beta = 0.5$) is [967.2 789.6 720.7 519.4 203.1 300 10 10 4 0]

As a conclusion of this section, we observe that in some special cases PAISP and PALPA methods may underperform with respect to the optimal one. The toy example given here help in understanding when this may happen. However, the point here is to understand how often these "pathological" cases may happen with real videos and how much they influence on video's quality. In order to answer to this question, we performed a comprehensive simulation campaign, which is detailed in the next section. However, we anticipate here the most important

results: with real video signal of resolutions ranging from 416×240 to 1920×1080, PAISP and PALPA methods achieved distortions that are very close to the optimal one. Moreover, their execution times are sensibly lower and they admit parallel implementations that could further speed them up.

4.6 Simulation results

In this section, we compare the performance of optimal allocation method (Section 3.4) with the four proposed suboptimal ones in terms of average PSNR on the same video sequences under per subchannel power constraints. The simulation conditions and metadata transmission are same as in Section 3.6.1 and Section 3.6.2. We also compare the optimal allocation with PAISP, PAISP with Dichotomy and PALPA in terms of complexity. At the end, in Section 4.6.3, the robustness of proposed suboptimal method is also tested.

4.6.1 Comparison of the power allocation methods

In Section 3.4, we have seen that precoding matrix design with per-subchannel power constraints can be formulated as a design problem with an equivalent channel with per-subchannel SNR constraints and unit subchannel noise variances. As a consequence, for the simulations, one assumes again unit noise variance on all subchannels and adjust the transmission power of chunks on each subchannel to have subchannel SNR matching those described in Figure 3.3.

The results of the simulation are shown in terms of average PSNR of the received sequences in Table 4.1. For PAISP and PALPA, the PSNR gap to optimality is never larger than 0.03 dB. In Figures 4.7, 4.8, 4.9 and 4.10, the evolution of the PSNR of some videos as a function of the frame index is shown. We observe that PAISP (with $\alpha = 0.75$ and $\beta = 0.5$), PAISP with Dichotomy and PALPA have very close PSNR performance. Only the results of PAISP are thus represented. All approaches clearly outperform the SCS allocation of Section (4.1), which can be considered as a natural extention of Parcast [LHL+14b].

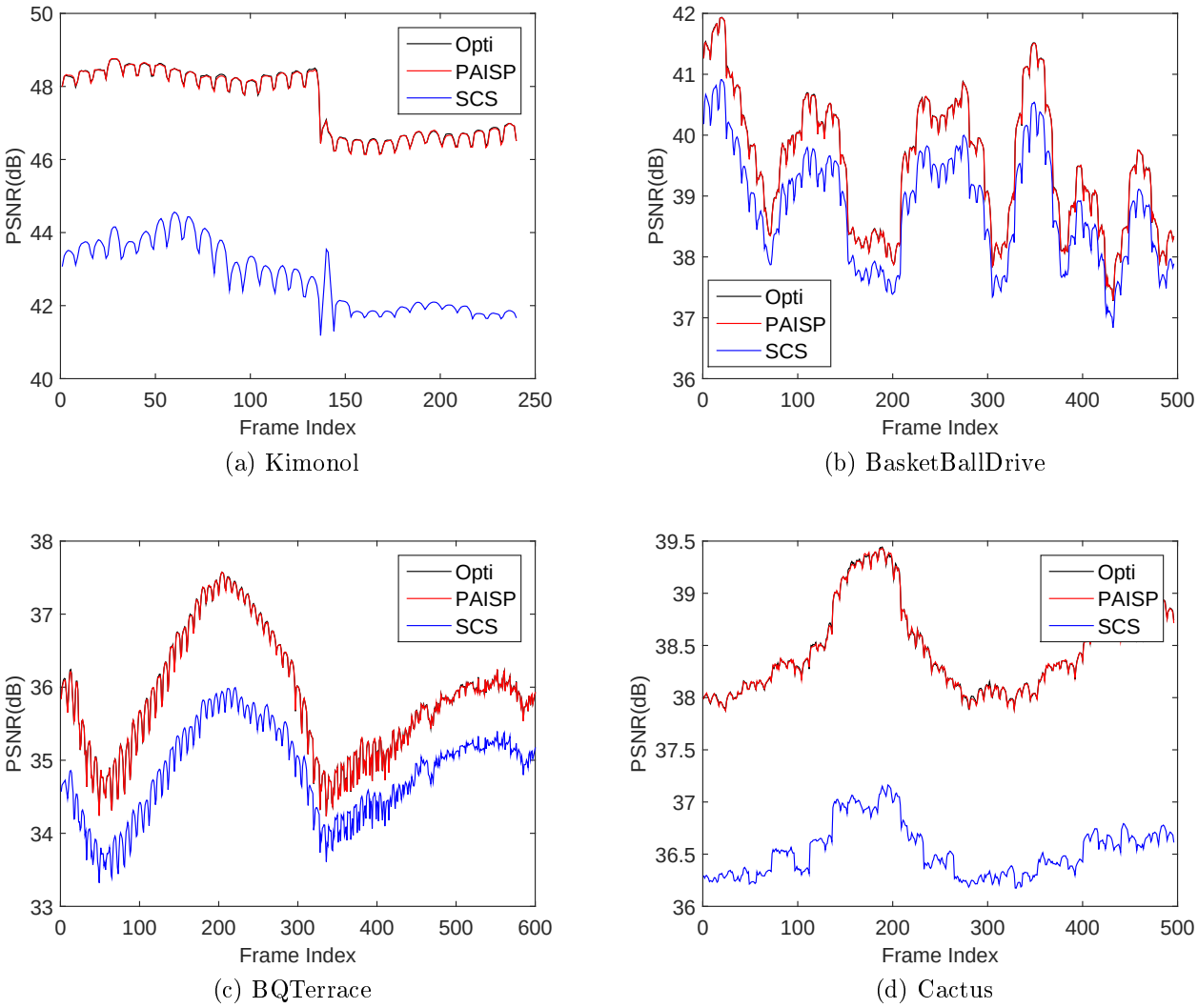
For sequences of class E (video conference content), where many chunks have very small variance (because of low spatial or temporal activity), the performance gain of the proposed approaches is really significant compared to SCS. The many small variance chunks prevent SCS to achieve good allocation performance and justify the large gains in these cases. High gains

Class	Name	PSNR (dB)			
		SCS	Opt.Alloc. or PAISP (Dichotomy)	PAISP or PALPA	Gain
B	Kimonol	42.79	47.57	47.56	4.78
	BasketballDrive	38.83	39.54	39.53	0.71
	BQ Terrace	34.83	35.86	35.85	1.03
	Cactus	36.53	38.47	38.46/38.44	1.94
	ParkScene	41.83	44.06	44.06/44.03	2.23
Av. PSNR Class B		38.96	41.10	41.09/41.08	2.13
C	PartyScene	40.89	42.94	42.94	2.05
	BQMall	41.24	44.91	44.90/44.91	3.67
	BasketballDrill	44.96	47.32	47.31	2.36
	RaceHorses	42.81	46.21	46.21	3.4
Av. PSNR Class C		42.48	45.35	45.34	2.87
D	BQSquare	39.38	44.55	44.55	5.17
	RaceHorses	43.89	49.03	49.03	5.14
	BlowingBubbles	42.26	47.90	47.90	5.64
	BasketballPass	45.03	49.55	49.55	4.52
Av. PSNR Class D		42.64	47.76	47.76	5.12
E	FourPeople	40.74	47.13	47.11/47.13	6.39
	Jonny	40.56	48.43	48.40/48.43	7.87
	KristenAndSara	39.77	46.95	46.94/46.95	7.18
Av. PSNR Class E		40.36	47.50	47.48/47.50	7.14
F	SlideShow	35.28	46.83	46.82/46.80	11.55

Table 4.1: Simulation results with per-subchannel power constraints

are also observed for Kimono1, due to the flatness of the variance vector. For video sequences in Class D, since the size of the video is relative small, the subchannels to be used have higher SNR constraints and are relatively flat compared to the vector of variances. In this case optimal precoding or PAISPs/PALPA can allocate power in a efficient way. For SlideShow in Class F, the high gains come from two aspects. First, many chunks have small variance, in this case the optimal and suboptimals designs can allocate power more efficiently. Second, there are chunks in some GoPs with zero variance, hence in such case, some subchannels do not need to be used to transmit these chunks. With SCS these available subchannels cannot be used, contrary to the optimal precoding matrix design method, PAISPs, or PALPA, which increases significantly the PSNR.

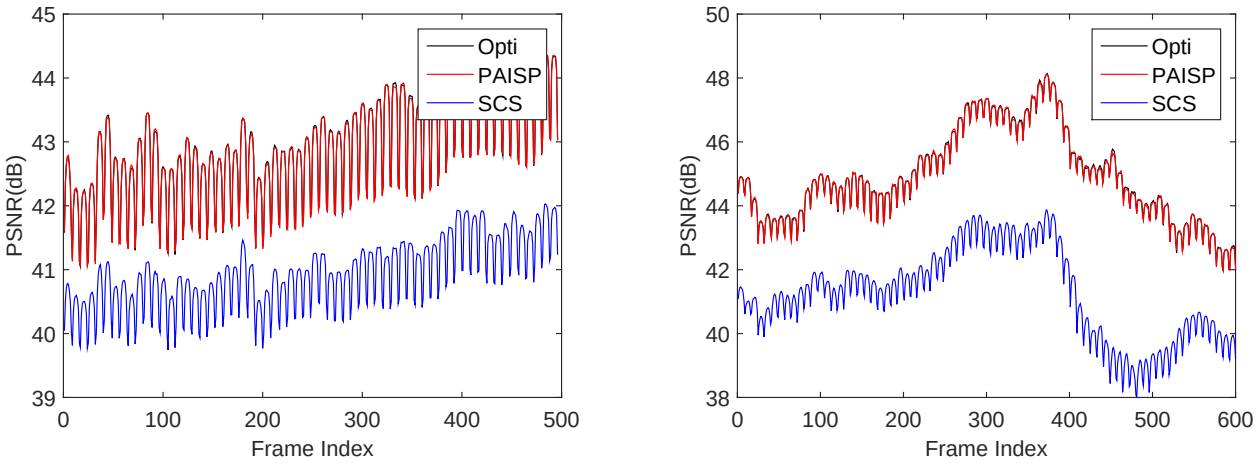
Finally, Figure 4.11 shows reconstructed first frames of Kimonol with PAISP and SCS respectively. Full sequences are available at *https://drive.google.com/drive/folders/1DIkGdk1IIZVvz3AiM68RUO6Yn2-DGR5S?usp=sharing.*



(a) Kimonol (b) BasketBallDrive

(c) BQTerrace (d) Cactus

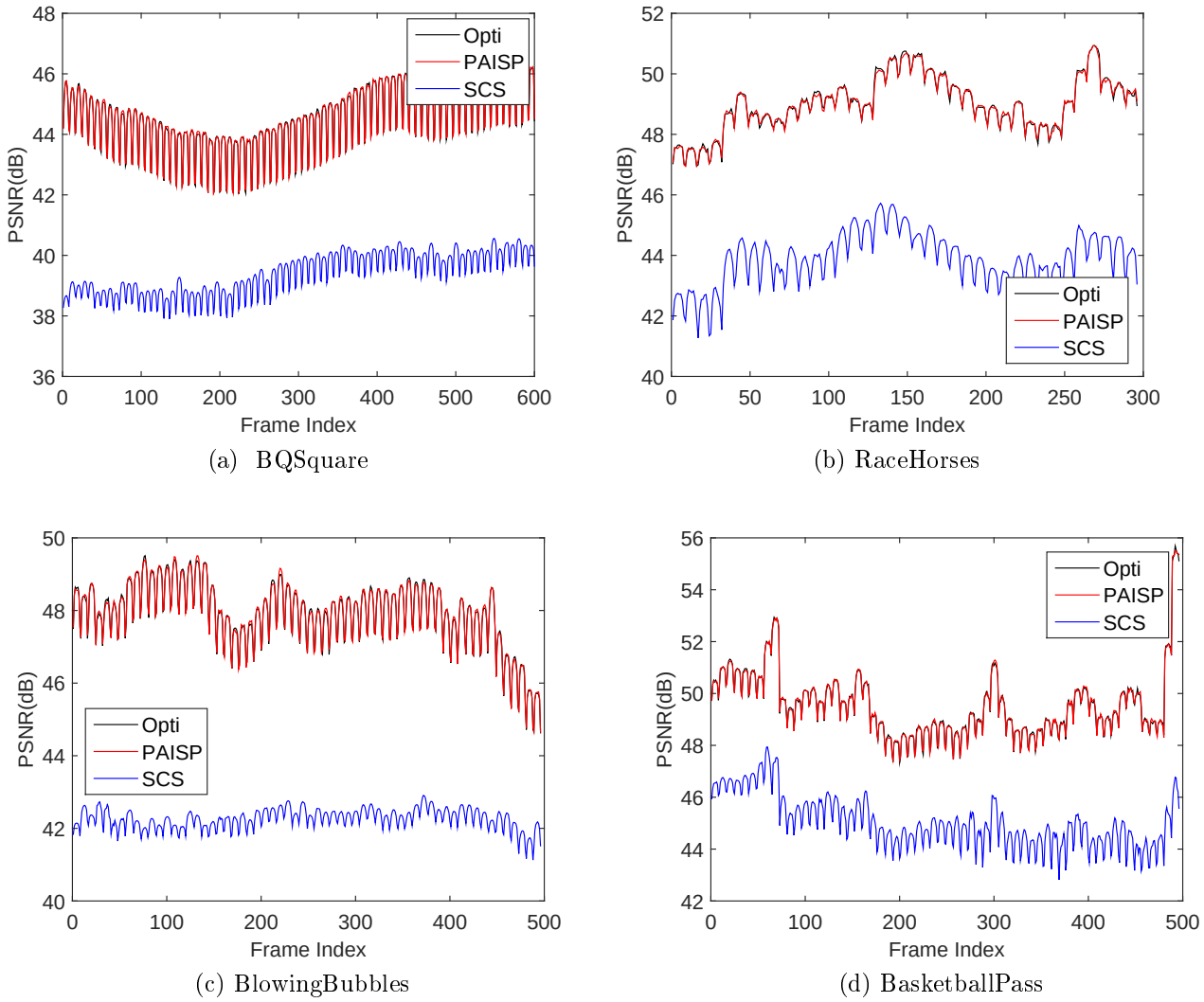
Figure 4.7: PSNR of Class B



(a) PartyScene (b) BQMall

(c) BastballDrill (d) RaceHorses

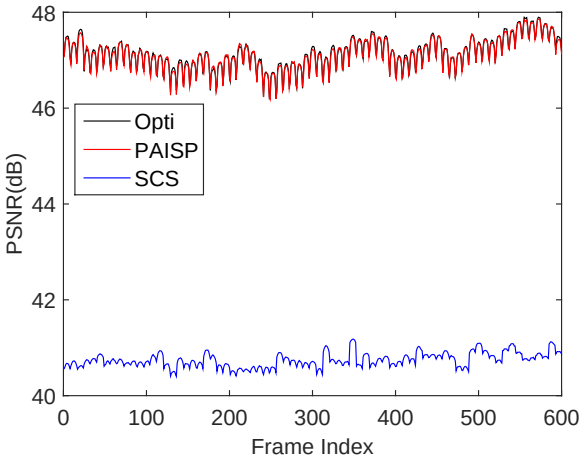
Figure 4.8: PSNR of Class C



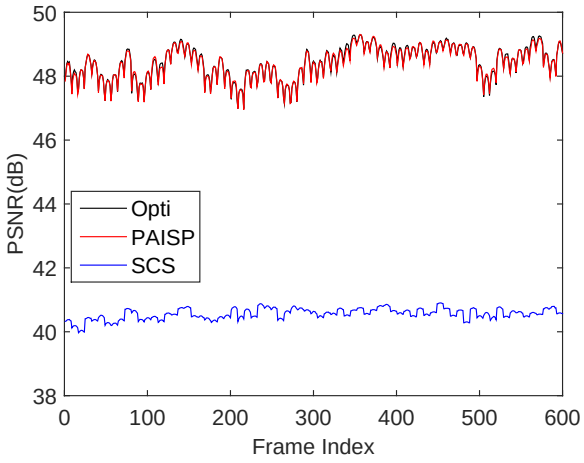
(a) BQSquare (b) RaceHorses

(c) BlowingBubbles (d) BasketballPass

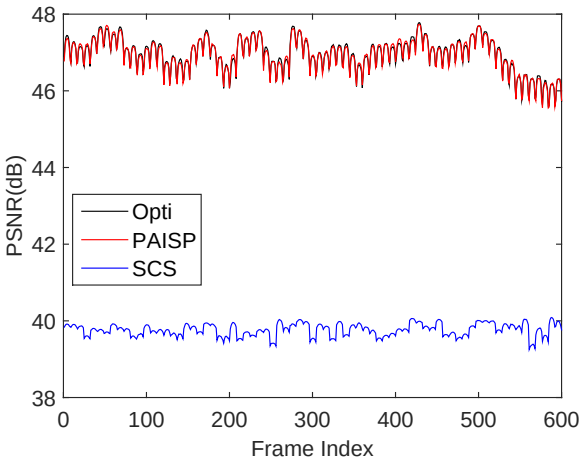
Figure 4.9: PSNR of Class D



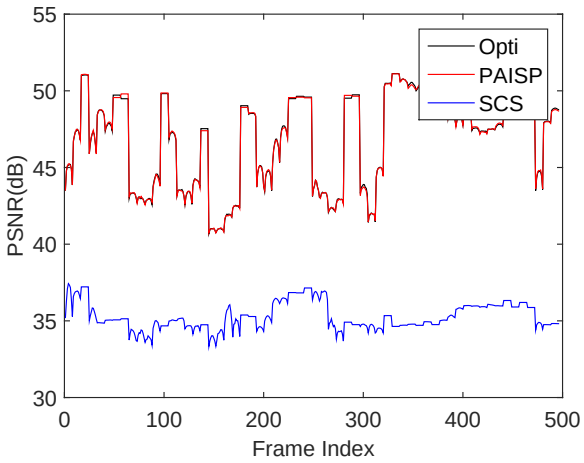
(a) FourPeople



(b) Johnny



(c) KristenaAndSara



(d) SlideShow

Figure 4.10: PSNR of Class E and Class F



(a) PAISP, PSNR = 48.03dB



(b) SCS, PSNR = 43.08dB

Figure 4.11: Reconstructed first frames of Kimonol with PAISP and SCS. (a) PAISP, (b) SCS

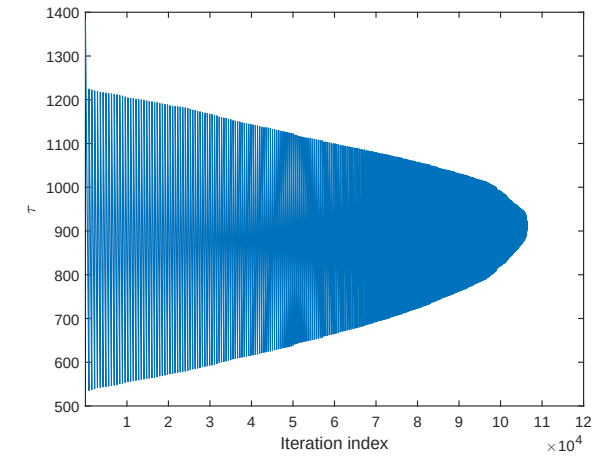
4.6.2 Complexity comparison

The computation cost of the optimal power allocation, of PAISP, of PAISP with Dichotomy and of PALPA are compared on simulations performed using MatlabR2014b on an Intel(R)Xeon(R)CPU E5-1603 v3 @ 2.8GHz. Table 4.2 provides the speed-up factor (ratio of precoding matrix computation times) of PAISPs and PALPA compared to the optimal precoding matrix design method. The parameters of PAISP have been taken as $\alpha = 0.75$ and $\beta = 0.5$. For RaceHorses of class C, SlideShow of Class F, the speed-up of the suboptimal algorithms is close to one, since in most of the GoPs it is not necessary to perform vector splitting. For videos of Class D, there is no split within all GoPs, therefore the four methods again perform similarly. But for the video sequences of class B, class E, and the video BQMall of class C, the speed up is significant, especially for PALPA and PAISP. The reasons is that the suboptimal algorithms can quickly find the split positions. On the other hand, the complexity to obtain an orthogonal matrix (Lemma 1) is $O(n^2)$ [ZZ95] where n is the length of subvector. Since the suboptimal algorithms may lead to more split positions than the optimal algorithm, the size of the subblocks is decreased and the computation costs related to the solution of the SHIE problem are also decreased. We also evaluated the speed-up factor of the heuristic approach in [LP76][2]. We can see the proposed suboptimal methods are also faster than [LP76].

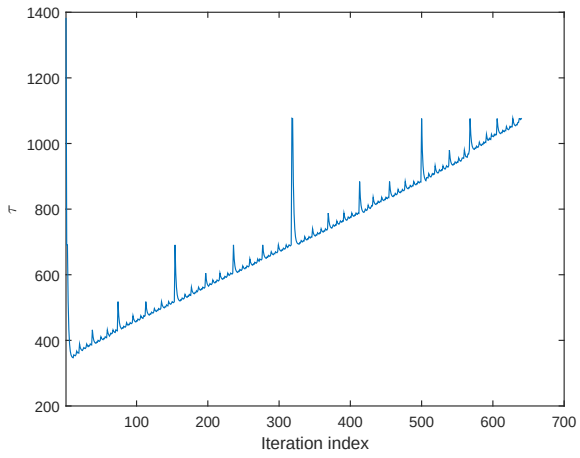
To further illustrate these results, consider the first GoP of Kimonol sequence. The sorted variances of the chunk vector components and the per-subchannel SNR constraints are shown in Figure 4.2. Figure 4.12 represents the largest index τ at which condition (3.41) is violated at each iteration of the optimal allocation algorithm, of PAISP, and of PALPA. We also plot the test positions of PAISP Dichotomy (variable c in Algorithm 4.2) at each iteration. The optimal allocation algorithm requires 106438 iterations, whereas PAISP requires 640 iterations, PAISP Dichotomy 2171 iterations and PALPA 848 iterations. This explains the efficiency of the proposed suboptimal methods for the Kimono1 sequence.

Figure 4.13 represents the same information as Figure 4.12 for a chunk vector in the 7th GoP of BasketballDrill. The optimal allocation algorithm requires only 360 iterations, whereas PAISP requires 142 iterations, PAISP Dichotomy 61 iterations and PALPA 61 iterations. In this case the suboptimal algorithms do not reduce significantly the time to find the precoding

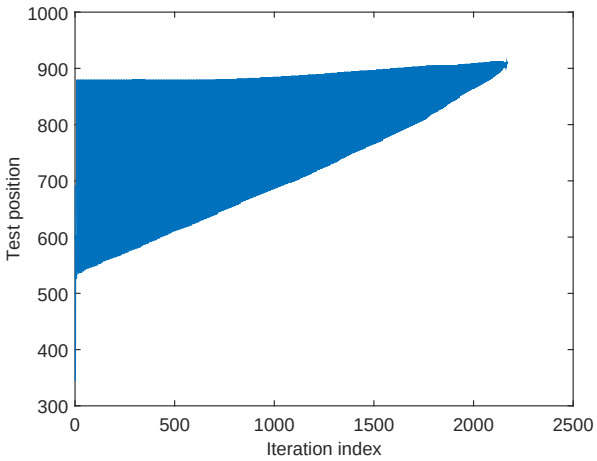
[2]In [LP76], there is an another algorithm to find the orthogonal matrix, but the complexity is higher than [ZZ95]. In order to have a fair comparison, we adopte the method of [ZZ95] in the implementation[LP76].



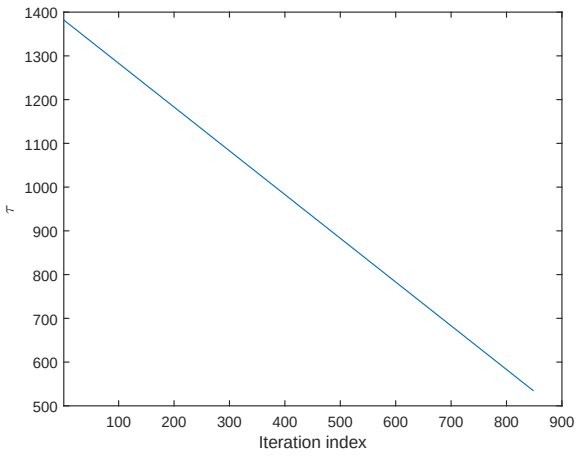
(a) Optimal allocation algorithm



(b) PAISP



(c) PAISP Dichotomy



(d) PALPA

Figure 4.12: Evolution of the value of the largest index τ at which Condition (3.41) is violated and the test positions for Dichotomy as a function of the iteration index for a chunk vector of the first GoP of Kimonol (Class B): (a) for the optimal allocation algorithm, (b) for PAISP, (c) for PAISP Dichotomy, and (d) for PALPA.

Class	Name	Speed-Up			
		PAISP	PALPA	PAISP (Dichotomy)	Approach in [LP76]
B	Kimonol	39	33	6	4
	BasketballDrive	20	122	3	2
	BQ Terrace	12	168	3	2
	Cactus	19	117	2	1.4
	ParkScene	10	44	4	3
C	PartyScene	5	3	2	1.0
	BasketballDrill	1.3	1.1	1.0	1.0
	BQMall	9	5	3	4
	RaceHorses	1.1	1.1	1.0	1.0
E	FourPeople	6	13	3	2
	Jonny	12	21	4	3
	KristenAndSara	12	11	3	2
F	SlideShow	1.0	1.0	1.0	1.0

Table 4.2: Speed-up factor provided by PAISPs and PALPA and approach in [LP76] compared to the optimal power allocation method of Section 3.4 for precoding matrix design.

Class	Name	PSNR (dB) and Speed-Up		
		$\alpha = 0.75$ $\beta = 0.5$	$\alpha = 0.5$ $\beta = 0.3$	$\alpha = 0.25$ $\beta = 0.1$
B	Kimonol	47.56dB	47.55dB	47.37dB
		39x	31x	29x
E	FourPeople	47.11dB	47.12dB	46.71dB
		6x	6x	6x
	Jonny	48.40dB	48.34dB	48.03dB
		12x	13x	10x
	KristenAndSara	46.94dB	46.95dB	46.53dB
		12x	11x	10x

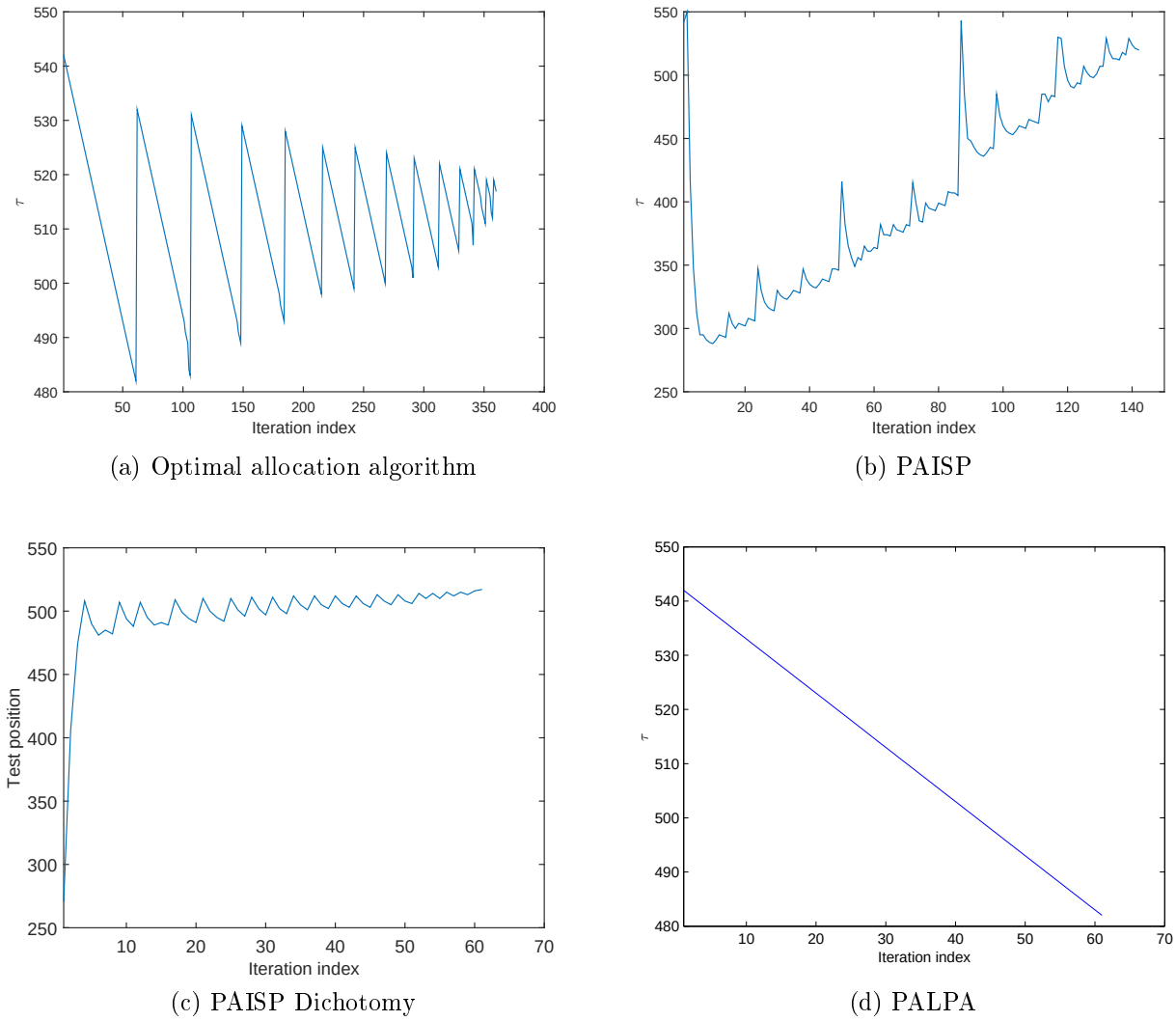
Table 4.3: PAISP: Influence of α and β

matrix.

In conclusion, PAISP is faster than PAISP with Dichotomy, but it should be tuned using appropriate values of α and β. Table 4.3 illustrates the influence of the values of α and β on the PSNR and on the speed-up factor for a subset of the considered video sequences. The values $\alpha = 0.75$ and $\beta = 0.5$ provide a good compromise between PSNR degradation and speed-up.

4.6.3 Mismatch

In this section, we test the robustness of the proposed suboptimal methods in a multicast scenario. The simulation condition is the same as in Section 3.6.5.



(a) Optimal allocation algorithm

(b) PAISP

(c) PAISP Dichotomy

(d) PALPA

Figure 4.13: Evolution of the value of the largest index τ at which Condition (3.41) is violated and the test positions for Dichotomy as a function of the iteration index for a chunk vector in the 7th GoP of BasketballDrill (class C): (a) for the optimal allocation algorithm, (b) for PAISP, (c) for PAISP dichotomy, and (d) for PALPA.

Video	receiver 1 PSNR(dB)		receiver 2 PSNR(dB)	
	Point-to-Point	MultiCast	Point-to-Point	MultiCast
Kimonol	47.56	47.40	42.14	42.14
BQMall	44.90	44.84	38.29	38.29

Table 4.4: linearly degraded multicast channel per-subchannel power constraint

receiver 1	receiver 2	Precoder	receiver1 PSNR(dB)	receiver2 PSNR(dB)
Chan.1	5% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.63	44.63
		Ave. SNR	41.78	41.75
		Chan.1	44.90	38.08
—	10% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.37	44.37
		Ave.SNR	39.74	39.79
		Chan.1	44.90	35.27
—	15% flip from Chan.1	$p_i/\sigma^2_{o,i}$	44.07	44.07
		Ave.SNR	38.25	38.26
		Chan.1	44.90	33.48
—	20% flip from Chan.1	$p_i/\sigma^2_{o,i}$	43.75	43.75
		Ave.SNR	37.20	37.19
		Chan.1	44.90	32.26

Table 4.5: Mismatch For video sequence BQMall in general multicast channel

First, one considers a channel model where the covariance of channel noise of Receiver 2 is 5 times larger than that of Receiver 1. The procedure to compute the precoding matrix is described in Section 3.5.1.2, except that the precoding matrix computation considers the equivalent channel and is done using a suboptimal method. Since the suboptimal methods have similar performance, we only do the simulation with PAISP. The results are shown in Table 4.4. We can see that the PSNR loss in this case of linear degraded channel is negligible.

Then results for a channel model for Receiver 2 obtained by flipping a fraction of subchannels of the channel of Receiver 1 are shown in Table **??**. Likewise the experiments reported in Section 3.6.5.2, the design method proposed in Section 3.5.2.1 provides the best results, and the PSNR only decreases by 0.3dB each time the probability of flipping increases by 5%.

4.7 Conclusions

This chapter has presented four suboptimal precoding matrix design, when the video has to be transmitted over parallel additive white Gaussian noise (AWGN) channel, with different characteristics. Among these techniques, PAISP, PAISP with Dichotomy and PALPA may sig-

nificantly reduce the matrix design complexity comparing to optimal approach, with a marginal degradation in terms of video PSNR. Comparing to SCS which can be considered as the most straightforward and natural extension of Parcast [LHL+14b], the optimal method, the both versions of PAISP and PALPA have significant average PSNR gains, ranging from 2.13 dB for class B videos to 11.55 dB for class F videos. From Table 4.1 and results shown in Section 3.6.4, we also can see that PAISPs and PALPA achieve better performance than [LP76]. Moreover, PAISPs and PALPA are faster than the heuristic method in [LP76] and than the optimal method. Moreover, the robustness of PAISP is also tested in the case of mismatch, , it is shown that the PSNR loss is rather small.

Chapter 5

Impulse error mitigation for LVC schemes

5.1 Introduction and main contributions

The characteristics of the transmission channel have been better taken into account into SoftCast-based video transmission. The first papers considered wideband additive white Gaussian noise (AWGN) channels [JSKG11]. Fading channels and MIMO channels [HLL+17, ZWW+15, ZLMW17, LHL+14a, LHL+14b] have then been considered. Optimal precoding schemes for per-subchannel power constrained channels have been designed in [ZAC+16]. Nevertheless, all the above-mentioned papers consider mainly channels affected by Gaussian noise only. Several types of communication channels may be also prone to impulse noise, such as the Digital Subscriber Line (DSL) [Ned03] and the Power Line Telecommunications (PLT) channels [ZD02]. Impulse noise has a high magnitude (its power may be 50dB above that of the background noise), and when it is bursty, may corrupt the channel for more than 1 ms [ZD02]. If impulses are not corrected, the communication performance may be significantly degraded [ANQC14, LNE13], even if LVC schemes are more robust than classical video coding scheme to noise and channel mismatch [JK10b].

In this chapter, we address the problem of impulse noise mitigation when the LVC-encoded video is transmitted using an Orthogonal Frequency-Division Multiplexing (OFDM) scheme for multi-carrier modulation over a wideband channel prone to impulse noise. As in [ANQC14], the impulse noise is modeled in the time domain by independent and identically distributed (iid) Bernoulli-Gaussian variables. A Fast Bayesian Matching Pursuit (FBMP) [SPZ08] algorithm, adapted to OFDM systems by [ANQC14], is employed for impulse noise mitigation. This

approach requires the provisioning of some OFDM subchannels to estimate the impulse noise locations and amplitudes. Since nothing can be transmitted on provisioned subchannels, this leads to a decrease of the number of chunks which may be transmitted and to a decrease of the video quality at receivers in absence of impulse noise.

To adresse this problem, we propose a phenomenological model (PM) structure to describe the residual noise in the OFDM subchannels after impulse noise estimation and removal. It amounts to a parametric model takes as input the channel and noise characteristics as well as the number of provisioned subchannels. We have combined the PM with a model of the evolution of the PSNR at the receiver in absence of impulse noise as a function of the channel and video characteristics and of the proportion of transmitted chunks. It is then possible to optimize the proportion of subchannels to provision. The parameters of the PM have been adjusted for different channel characteristics and one has observed that it leads to accurate estimates of the optimal proportion of subchannel to provision once the characteristics of the video are known. Simulation results show that impulse noise may then be efficiently mitigated with a limited impact on the PSNR that may be obtained in absence of impulse noise.

The rest of the chapter is organized as follows. Related results are described in Section 5.2. The SoftCast-based coding and transmission system is introduced in Section 5.3. The application of FBMP for impulse noise mitigation is described in Section 5.4. Section 5.5 presents the way the optimal number of subchannels to provision can be determined for impulse noise correction. Simulation results are described in Section 5.6 before drawing some conclusions in Section 5.7.

Table 5.1 summarizes the main notations used in this chapter.

5.2 Related work

One approach to mitigate the impulse noise is to use a clipping/blanking nonlinearity [Zhi08]. A threshold is chosen to decide whether there is a impulse noise in the received symbol. If the magnitude of the received symbol is above that threshold, this symbol is clipped or blanked. The improvement is little when the signal to impulse noise ratio (SINR) is large [Zhi06]. Another approach is to use sparse vector recovery algorithms to estimate the impulse noise characteristics (position and amplitude) [CR08, Lam11, LNE13, ANQC14]. In [CR08] the estimation of

Symbol	Set	Represents
n_{Ck}	$\mathbb{N}$	Dimension of source vector
n_{SC}	$\mathbb{N}$	Nb of parallel subchannels
t	$\mathbb{R}^{n_{\mathrm{Ck}}}$	a chunk vector
λ_i	$\mathbb{R}_+$	variance of a chunk
Λ	$\mathbb{R}_+^{n_{\mathrm{Ck}} \times n_{\mathrm{Ck}}}$	Chunk vector covariance matrix
u	$\mathbb{R}^{n_{\mathrm{Ck}}}$	scaled chunk vector
$\widetilde{u}$	$\mathbb{C}^{n_{\mathrm{Ck}}}$	a complex scaled chunk vector
F	$\mathbb{C}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	DFT matrix
v_{g}	$\mathbb{C}^{n_{\mathrm{SC}}}$	Gaussian noise vector
N_{g}	$\mathbb{R}_{++}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	Covariance of Fv_{g}
$2\sigma_i^2$	$\mathbb{R}_{++}$	variance circular complex Gaussian noise of i-th subchannel
N	$\mathbb{R}_{++}^{n_{\mathrm{SC}} \times n_{\mathrm{SC}}}$	$N_{\mathrm{g}} = 2N = 2\mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_{n_{\mathrm{SC}}}^2\right)$
v_{I}	$\mathbb{C}^m$	Impulse noise vector
δ	$\mathbb{R}$	Bernoulli variable
w	$\mathbb{C}$	Impulse noise variable
p_{I}	$\mathbb{R}$	Probability of impulse noise
$2\sigma_{\mathrm{I}}^2$	$\mathbb{R}$	Variance of impulse noise
p_{T}	$\mathbb{R}_+$	Total power constraint
G	$\mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}}$	Channel precoding matrix
H	$\mathbb{R}^{n_{\mathrm{Ck}} \times n_{\mathrm{SC}}}$	Decoding matrix
ℓ	$\mathbb{N}$	Nb of transmitted components
q	$\mathbb{N}$	Nb of zero components in Gt
r_{d}	$\mathbb{R}_+$	Ratio of subchannels provisioning q/n_{SC}
Ψ	$\mathbb{C}^{q \times n_{\mathrm{SC}}}$	Parity-check matrix
s	$\mathbb{C}^q$	Syndrome vector
N_s	$\mathbb{R}^{q \times q}$	Last q rows of N_{g}
v_{r}	$\mathbb{C}^{n_{\mathrm{SC}}}$	Vector of impulse noise correction residual
σ_{r}^2	$\mathbb{R}$	Variance of the impulse noise correction residual

Table 5.1: Main notations

impulse noise is performed borrowing tools from compressive sensing (see [CRA+16] and the references there in). The quality of the recovery depends on the restricted isometry property of the conjugate transpose of precoding matrix [CR08]. In other papers, an *a priori* statistical information is exploited to estimate the sparse impulse noise samples, *e.g.*, the sparse Bayesian learning approach [WR04, Tip01] is applied in [Lam11, LNE13, ANQC14]. A adaptation of the FBMP algorithm [SPZ08] for OFDM systems has been proposed in [ANQC14]. This leads to a reduced run time by one order of magnitude with a performance similar to that of FBMP.

Compared to the state-of-the-art, our contributions are (*i*) to adapt the FBMP-based impulse mitigation technique to LVC schemes in the context of OFDM for wideband channels prone to impulse noise; (*ii*) to propose a PM of the residual noise after impulse noise mitigation; (*iii*) to provide an algorithm for the selection of the number of subchannels to provision which minimizes the receiver Mean-Square Error (MSE). Simulation results illustrate the performance improvements provided by the impulse noise mitigation scheme once the optimal number of provisioned subchannels has been chosen.

5.3 Linear Video Coding and OFDM Transmission Scheme

The SoftCast-based LVC architecture [JK10b] is first briefly recalled before focusing on the OFDM-based transmission scheme used to convey the LVC-processed video frames. The considered coding and transmission scheme is represented in Figure 5.1.

5.3.1 Joint source-channel coding

We focus on the luminance part of the video. The chrominance components undergo a similar processing. For the power allocation between luminance and chrominance components, the weighting approach adopted in MPEG may be employed [OSS+12]. The input digital video signal is organized in Group of Pictures (GoP) of n_{F} frames with $n_{\mathrm{R}} \times n_{\mathrm{C}}$ pixels each. Each GoP is processed independently. First, a full-frame 2D-DCT is applied on each frame of the GoP, then a temporal 1D-DCT is applied on the transformed frames to perform temporal decorrelation. The transformed GoP coefficients are grouped into *chunks* of $n_{\mathrm{r}} \times n_{\mathrm{c}}$ coefficients from nearby spatio-temporal subbands. Each GoP contains thus $n_{\mathrm{Ck}} = n_{\mathrm{F}} \frac{n_{\mathrm{R}} n_{\mathrm{C}}}{n_{\mathrm{r}} n_{\mathrm{c}}}$ chunks.

The elements of a given chunk are assumed iid and to follow the same zero-mean Gaussian

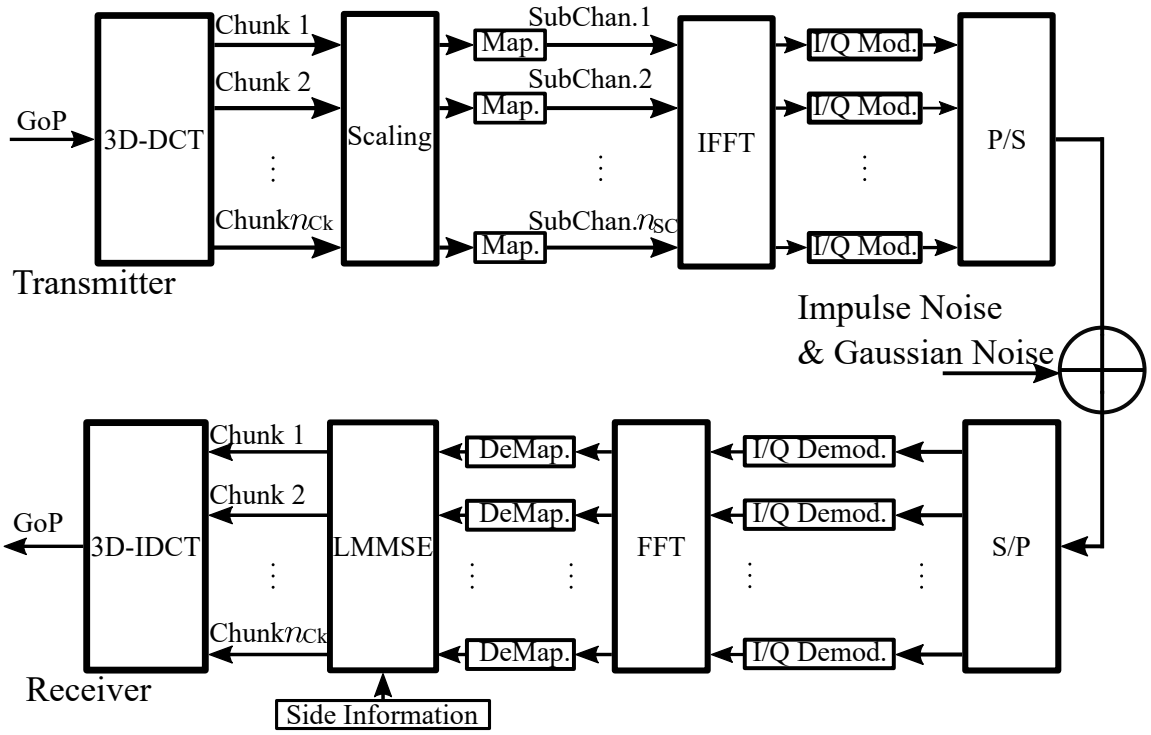


Figure 5.1: SoftCast-based linear video coding, transmission, and decoding architecture

distribution. Without loss of generality, the chunks, are assumed to be sorted according to decreasing variance λ_i, $i = 1, \ldots, n_{\mathrm{Ck}}$. Usually, due to channel transmission power constraints and bandwidth limitations, only ℓ chunks among the n_{Ck} are sent. A bitmap indicating the transmitted chunks, as well as their mean values, and their variances are robustly transmitted (using, *e.g.*, a strong FEC) as metadata on the channel. The resulting rate overhead may be typically neglected, provided that the number of chunks remains small compared to the number of pixels of each frame [JK10b].

A scaling of the selected chunks is then performed to help minimizing the reconstruction MSE at receiver, assuming that each scaled chunk has been transmitted over a different subchannel. This optimization may be performed assuming a total power constraint as in [JK10b] or considering a per-subchannel power constraint as in [ZAC+16]. In the first case, the resilience to packet losses may be improved by a Hadamard matrix to transform the chunks into equal-energy *slices*.

5.3.2 Transmission

One considers an OFDM-based transmission scheme with n_{SC} subchannels on which quasi-analog signaling (64k-QAM as in [JK10b]) or analog QAM is used. A total power p_{T} is available

for the transmission of each OFDM symbol.

To perform scaling and transmission, $n_{\mathrm{r}} \times n_{\mathrm{c}}$ *chunk vectors* t_i, $i = 1, \ldots, n_{\mathrm{r}} \times n_{\mathrm{c}}$, each of dimension n_{Ck}, are formed by selecting for each vector one coefficient per chunk. The t_is can be seen as realizations of $n_{\mathrm{r}} \times n_{\mathrm{c}}$ iid zero-mean Gaussian vectors with covariance matrix $\Lambda = \mathrm{diag}(\lambda_1 \ldots \lambda_{n_{\mathrm{Ck}}})$. The chunk vectors are multiplied by a precoding matrix $G \in \mathbb{R}^{n_{\mathrm{SC}} \times n_{\mathrm{Ck}}}$ designed in such a way that $u_i = Gt_i$ satisfies a power constraint $p_{\mathrm{T}}/2$. Then $n_{\mathrm{r}} \times n_{\mathrm{c}}/2$ vectors of complex symbols are formed by combining pairs of consecutive scaled chunk vectors

$$\begin{aligned} \widetilde{u}_i &= u_{2i-1} + j u_{2i} \\ &= G\left(t_{2i-1} + j t_{2i}\right) \end{aligned} \tag{5.1}$$

with $\widetilde{u}_i \sim C\mathcal{N}\left(0, 2G\Lambda G^{\mathrm{T}}\right)$, $i = 1, \ldots, n_{\mathrm{r}} \times n_{\mathrm{c}}/2$. In what follows, the index i is omitted, since all vectors $\widetilde{u}_i$ have similar distribution and undergo the same processing. The $\widetilde{u}_i$s are transformed into $n_{\mathrm{r}} \times n_{\mathrm{c}}/2$ OFDM symbols, each satisfying the power constraint p_{T}, using an inverse Discrete Fourier Transform (IDFT). A cyclic prefix may be inserted and the symbols are then quadrature-mixed to passband.

One focuses on a transmission power and bandwidth constrained scenario where $n_{\mathrm{SC}} \leqslant n_{\mathrm{Ck}}$. The number of transmitted chunks ℓ is thus such that $\ell \leqslant n_{\mathrm{SC}}$.

5.3.3 Channel model

The transmitted signal is assumed to be corrupted by Gaussian noise and impulsive noise. Using the model introduced in [ANQC14], the input vector $y \in \mathbb{C}^{n_{\mathrm{SC}}}$ of the FFT at receiver may be represented as

$$y \;=\; F^H \widetilde{u} + v_{\mathrm{I}} + v_{\mathrm{g}} \tag{5.2}$$

where F^H is IDFT matrix, v_{g} is a Gaussian noise vector and v_{I} is an impulse noise vector. After the DFT, $Fv_{\mathrm{g}} \sim \mathcal{CN}\left(0, N_{\mathrm{g}}\right)$ can be modeled as a zero-mean complex circular Gaussian noise vector with $N_{\mathrm{g}} = 2N$ and $N = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_{n_{\mathrm{SC}}}^2\right)$, and v_{I} is an impulse noise vector. The components of v_{I} are iid and such that $v_{\mathrm{I},k} = \delta_k w_k$, where δ_k is the realization of a Bernoulli variable with parameter $p_{\mathrm{I}} = \Pr\{\delta_k = 1\}$ and $w_k \sim \mathcal{CN}\left(0, 2\sigma_{\mathrm{I}}^2\right)$ with $\sigma_{\mathrm{I}}^2 > \sigma_i^2$, $i = 1, \ldots, n_{\mathrm{SC}}$.

5.3.4 Baseline receiver

At first baseline receiver is introduced, without considering impulse noise. The components of the vector y go through an I/Q demodulator, a DFT to get

$$Fy = \widetilde{u} + Fv_{\mathrm{g}},$$

which is fed to a demapper. An LMMSE estimate $\hat{t}_i$ of the chunk vector t_i is then evaluated using a decoding matrix $H \in \mathbb{R}^{n_{\mathrm{Ck}} \times n_{\mathrm{SC}}}$.

Finally, using the side information (*i.e.,* the map of selected chunks, the chunk mean values, and their variances), the GoP is reconstructed by applying 3D-Inverse DCT (3D-IDCT) on the estimated chunk vectors.

5.3.5 Power allocation and chunk selection

Here, the precoding and decoding matrix design is briefly recalled without accounting for the presence of the impulse noise. Without loss of generality, one assumes that the subchannel indexing is such that $\sigma_1^2 \leqslant \cdots \leqslant \sigma_{n_{\mathrm{SC}}}^2$.

Under total power constraint $p_{\mathrm{T}}/2$, and in absence of impulsive noise, the precoding matrix G and the decoding matrix H minimizing

$$\varepsilon_{\mathrm{NI}} = E\left[\left\|t - \hat{t}\right\|_2^2\right] \tag{5.3}$$

have been shown in [LP76, ZAC+16] to be

$$G = \begin{bmatrix} \mathrm{diag}\left(g_1 \dots g_\ell\right) & \mathbf{0}_{\ell \times (n_{\mathrm{Ck}} - \ell)} \\ \mathbf{0}_{(n_{\mathrm{SC}} - \ell) \times \ell} & \mathbf{0}_{(n_{\mathrm{SC}} - \ell) \times (n_{\mathrm{Ck}} - \ell)} \end{bmatrix} \tag{5.4}$$

and

$$H = \Lambda G^T \left(G \Lambda G^T + N\right)^{-1}, \tag{5.5}$$

where $\ell \leqslant n_{\mathrm{SC}}$ is the largest integer such that

$$\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 > 0, \; i = 1, \dots, \ell \tag{5.6}$$

with

$$\sqrt{\gamma} = \frac{\sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2}}{\frac{p_T}{2} + \sum_{i=1}^{\ell} \sigma_i^2} \tag{5.7}$$

and

$$g_i = \left(\sqrt{\frac{\lambda_i \sigma_i^2}{\gamma}} - \sigma_i^2 \right)^{1/2} / \sqrt{\lambda_i}, \ i = 1, \ldots, \ell. \tag{5.8}$$

In absence of impulse error, the resulting MSE on the chunk coefficients is then computed as

$$\varepsilon_{\mathrm{NI}} = \sum_{i=\ell+1}^{n_{\mathrm{Ck}}} \lambda_i + \sqrt{\gamma} \sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_i^2}. \tag{5.9}$$

If $\ell < n_{\mathrm{SC}}$, $n_{\mathrm{SC}} - \ell$ subchannels are not used for the transmission of scaled chunk coefficients. They will be helpful to mitigate the effect of the impulse noise, as shown Section 5.4.

In the plain Softcast, a Hadamard transform is performed after chunk scaling. This transform may be replaced by an orthogonal random matrix [LHL+14b], which will not modify the MSE [LP76, ZAC+16], but may improve the robustness to losses of sub-channels. Here, this additional transform is not considered.

5.4 Impulse Noise Correction

This section introduces the considered impulse noise correction algorithm adapted to the LVC and transmission scheme presented in Section 5.3.

When, $\ell < n_{\mathrm{SC}}$, one observes from (5.4) that the last $q = n_{\mathrm{SC}} - \ell$ rows of G are null, corresponds to sub-channels that are not used to transmit chunk coefficients. Here, in the proposed impulse noise mitigation scheme, it may be necessary to discard chunks even when there is enough available transmission power, since this operation improves the robustness to impulse noise, as shown later on. Sub-channels not used to transmit chunk coefficients (called provisioned subchannels in what follows), are used to estimate the characteristics of the impulse noise. Consider the matrix $\Psi \in \mathbb{C}^{q \times n_{\mathrm{SC}}}$ formed by the q last rows of F. From (5.4), one has

$$\Psi F^H G = 0. \tag{5.10}$$

Combining (5.2), (5.10), and (5.1), one may evaluate the *syndrome vector*

$$\begin{aligned} s &= \Psi y \\ &= \Psi F^H \widetilde{u} + \Psi \left(v_{\mathrm{I}} + v_{\mathrm{g}} \right) \\ &= \Psi v_{\mathrm{I}} + \Psi v_{\mathrm{g}}. \end{aligned} \tag{5.11}$$

Since the rows of Ψ are the q last rows of F, $\Psi v_{\mathrm{g}} \sim \mathcal{CN}\left(0, N_s\right)$, with $N_s = 2\mathrm{diag}\left(\sigma^2_{n_{\mathrm{SC}}-q+1}, \ldots, \sigma^2_{n_{\mathrm{SC}}}\right)$. Therefore to mitigate the effect of the impulse noise, one has to estimate the sparse vector v_{I} from noisy measurements of Ψv_{I}. This is a typical compressive sensing estimation problem for which many solutions have been proposed. Here, one considers the Fast Bayesian Matching Pursuit algorithm (FBMP) to get an estimate

$$\widehat{v_{\mathrm{I}}} = \mathrm{E}\left(v_{\mathrm{I}} | s \right) \tag{5.12}$$

of v_{I}, see [SPZ08] for more details. Finally, the vector $\widehat{y}$' after impulse noise mitigation is

$$\begin{aligned} \widehat{y} &= Fy - F\widehat{v_{\mathrm{I}}} \\ &= \widetilde{u} + F\left(v_{\mathrm{I}} - \widehat{v_{\mathrm{I}}} \right) + F v_{\mathrm{g}}. \end{aligned} \tag{5.13}$$

The components of $\widehat{y}$ are then used to get LMMSE estimates of t_{2i-1} and t_{2i}, $i = 1, \ldots, n_{\mathrm{r}} \times n_{\mathrm{c}}/2$ using (5.5), where N is replaced by the covariance matrix of $F\left(v_{\mathrm{I}} - \widehat{v_{\mathrm{I}}} \right) + F v_{\mathrm{g}}$ divided by 2 to account for (5.1).

Figure (5.2) represents the modified SoftCast-based LVC, transmission, and decoding architecture, with impulse noise estimation and mitigation at receiver. In the impulse noise estimation (INE) block, $\widehat{v_{\mathrm{I}}}$ is first estimated from the elements of the last q subchannels, see (5.11). Then, in the impulse noise mitigation (INM) block, $F\widehat{v_{\mathrm{I}}}$ is removed from Fy according to (5.13). In what follows, this scheme is called LVC With Subchannel Provisioning and Impulse Correction (LVC-WSP-IC).

The main difficulty lies in the optimization of the number q of sub-channels provisioned for impulse noise mitigation. The number q is not only due to the power constraint, the q lowest

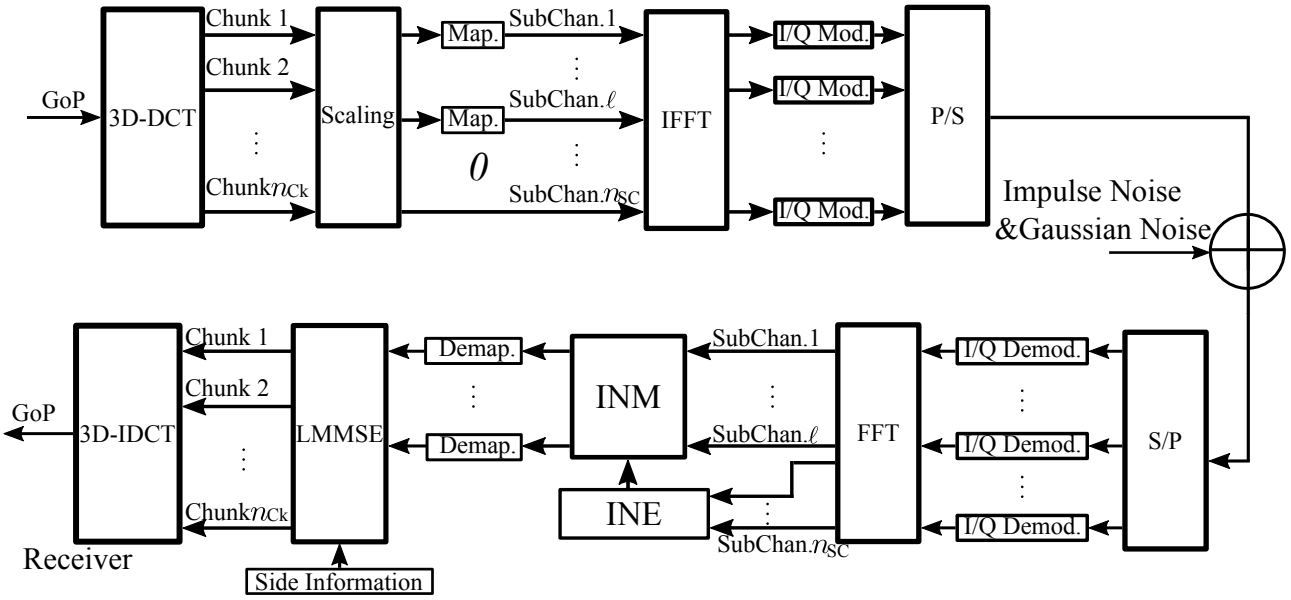


Figure 5.2: Modified SoftCast-based LVC (with discarded chunks), transmission, and decoding architecture (with impulse noise correction)

variance chunks may be also discarded in our proposed scheme even there is enough available power. A solution to this problem is detailed in Section 5.5.

5.5 Sub-channel provisioning for impulse noise mitigation

The efficiency of the FBMP algorithm increases with the number q of observations of linear combinations of the impulse errors (5.11). Nevertheless, increasing q reduces the number of subchannels on which chunk coefficients can be transmitted. A trade-off has thus to be found between efficiency of impulse noise correction and transmission performance. This requires a model of the residual noise after impulse noise mitigation and an evaluation of the impact of subchannel provisioning on the performance of the SoftCast-based LVC and transmission scheme.

5.5.1 Residual noise after impulse noise mitigation

One may rewrite (5.13) as

$$\widehat{y} = \widetilde{u} + Fv_{\mathrm{r}} + Fv_{\mathrm{g}}, \tag{5.14}$$

where $v_{\mathrm{r}} = v_{\mathrm{I}} - \widehat{v}_{\mathrm{I}}$ represents the impulse noise residual vector after mitigation. This residual can be seen as an additional noise component to the background Gaussian noise affecting the

sub-channels. This additional noise component has to be taken into account in the design of the precoding and decoding matrices described in Section 5.3.5.

One assumes that v_{r} and v_g are uncorrelated. As for the covariance of v_{r}, it can be observed that

$$\begin{aligned} \mathrm{Cov}\left(v_{\mathrm{r}}|s\right) &= \mathrm{Cov}\left(\left(v_{\mathrm{I}}-\widehat{v}_{\mathrm{I}}\right)|s\right) \\ &= \mathrm{Cov}\left(\left(v_{\mathrm{I}}-\mathrm{E}\left(v_{\mathrm{I}}|s\right)\right)|s\right) \\ &= \mathrm{Cov}\left(v_{\mathrm{I}}|s\right). \end{aligned} \tag{5.15}$$

The covariance of the estimation error (5.15) can be closely approximated [SPZ08] as

$$\mathrm{Cov}\left(v_{\mathrm{I}}|s\right) \approx \sum_{\widetilde{\delta}\in\Delta^*} p\left(\widetilde{\delta}|s\right) \mathrm{Cov}\left(v_{\mathrm{I}}|s,\widetilde{\delta}\right) \tag{5.16}$$

where $\widetilde{\delta}$ is a binary vector which non-zero entries indicate the estimated locations of impulses and Δ^* is the set of the D vectors $\widetilde{\delta}$ that achieve the largest values of $p\left(\widetilde{\delta}|s\right)$. A large value of D provides a better estimate of v_{I} and of $\mathrm{Cov}\left(v_{\mathrm{I}}|s\right)$ but at the price of a higher complexity. Moreover, from [SPZ08], one has

$$\mathrm{Cov}\left(v_{\mathrm{I}}|s,\widetilde{\delta}\right) = R\left(\widetilde{\delta}\right) - R\left(\widetilde{\delta}\right)\Psi^H\left(\Psi R\left(\widetilde{\delta}\right)\Psi^H + N_{\mathrm{s}}\right)^{-1}\Psi R\left(\widetilde{\delta}\right), \tag{5.17}$$

where $R\left(\widetilde{\delta}\right) = \mathrm{Cov}\left(v_{\mathrm{I}}|\widetilde{\delta}\right)$ is a sparse diagonal matrix, since the components of the impulse noise are iid and p_{I} is small in general. Using the matrix inversion lemma, one gets

$$\mathrm{Cov}\left(v_{\mathrm{I}}|s,\widetilde{\delta}\right) = R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\left(I - R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\Psi^H\left(\Psi R\left(\widetilde{\delta}\right)\Psi^H + N_{\mathrm{s}}\right)^{-1}\Psi R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\right)R\left(\widetilde{\delta}\right)^{\frac{1}{2}} \tag{5.18}$$

$$= R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\left(I + R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\Psi^H N_{\mathrm{s}}^{-1}\Psi R\left(\widetilde{\delta}\right)^{\frac{1}{2}}\right)^{-1} R\left(\widetilde{\delta}\right)^{\frac{1}{2}}, \tag{5.19}$$

Consider a column Ψ_j, $j = 1,\ldots,n_{\mathrm{SC}}$ of Ψ. As shown in [ANQC14] one has

$$\Psi_j^H\Psi_{j'} = \begin{cases} \frac{q}{n_{\mathrm{SC}}} & j = j', \\ \frac{q}{n_{\mathrm{SC}}}\left|\frac{\sin\left(\pi(j-j')\frac{q}{n_{\mathrm{SC}}}\right)}{q\sin\left(\pi(j-j')\frac{1}{n_{\mathrm{SC}}}\right)}\right| & j \neq j'. \end{cases} \tag{5.20}$$

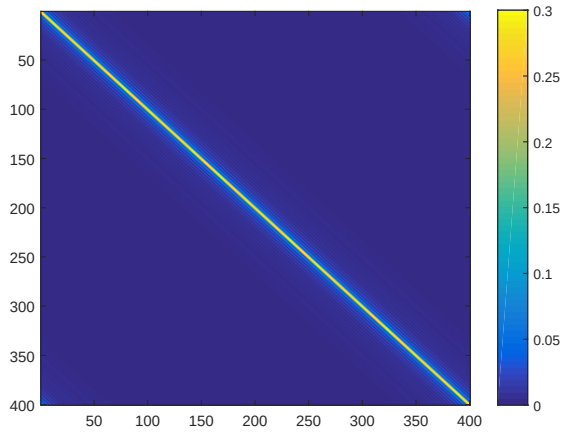


Figure 5.3: $\Psi^H\Psi$

Hence, any pair of columns Ψ_j and $\Psi_{j'}$ of Ψ, such that $j-j'$ is a multiple of $\frac{n_{\mathrm{SC}}}{q}$, are orthogonal. For other columns Ψ_j and $\Psi_{j'}$ with $j \neq j'$, $\Psi_j^H\Psi_{j'}$ is small compared to q/n_{SC}. Hence when q and n_{SC} are sufficiently large, one may approximate $\Psi^H\Psi$ as diagonal

$$\Psi^H\Psi \approx \frac{q}{n_{\mathrm{SC}}} I. \tag{5.21}$$

Figure 5.3 illustrates the norm of each entry of $\Psi^H\Psi$ when $n_{\mathrm{SC}} = 400$ and $q = 120$. It shows that the elements of $\Psi^H\Psi$ which around the diagonal are large and that the other terms are much smaller.

Now, one first assumes that $\sigma^2_{n_{\mathrm{SC}}-q+1} = \cdots = \sigma^2_{n_{\mathrm{SC}}} = \sigma_g^2$, leading to $N_s = 2\sigma_g^2 I$ of size $q \times q$. Using (5.19) and (5.21), one deduces that $\mathrm{Cov}\left(v_{\mathrm{I}}|s,\delta\right)$ is diagonal. In the general case, when $\sigma^2_{n_{\mathrm{SC}}-q+1} \leqslant \cdots \leqslant \sigma^2_{n_{\mathrm{SC}}}$, $\Psi^H N_{\mathrm{s}}^{-1}\Psi$ can again be approximated by a diagonal matrix provided that n_{SC} and q are large enough, see for example Figure 5.4. Consequently, $\mathrm{Cov}\left(v_{\mathrm{I}}|s,\delta\right)$ is close to diagonal and will be considered as diagonal in what follows.

Since $R\left(\delta\right)$ is sparse, $\mathrm{Cov}\left(v_{\mathrm{I}}|s,\delta\right)$ is also a sparse diagonal matrix, as well as, $\mathrm{Cov}\left(v_{\mathrm{I}}|s\right)$ provided that D remains small. The neglected contributions to $\mathrm{Cov}\left(v_{\mathrm{I}}|s\right)$ in (5.16) are terms with a very low associated *a posteriori* probability $p\left(\widetilde{\delta}|s\right)$. Consequently, $v_{\mathrm{r}}|s$ can be modeled as zero-mean Gaussian with a sparse diagonal covariance matrix.

Furthermore, the covariance matrix of $Fv_{\mathrm{r}}|s$ is

$$\mathrm{Cov}\left(Fv_{\mathrm{r}}|s\right) \quad = \quad F\mathrm{Cov}\left(v_{\mathrm{r}}|s\right)F^H \tag{5.22}$$

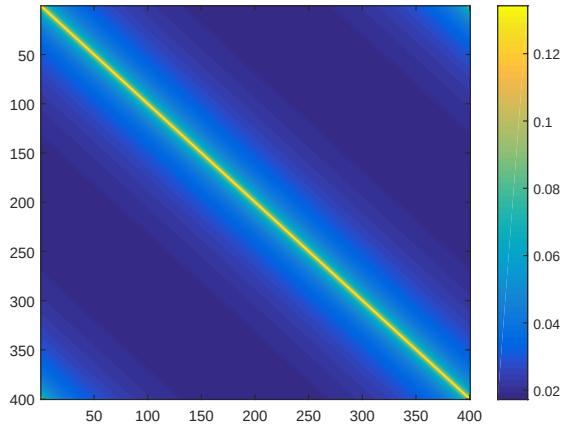


Figure 5.4: $\Psi^H N_{\mathrm{s}}^{-1}\Psi$. $n_{\mathrm{SC}} = 400$, $q = 120$, and $N_{\mathrm{s}} = \mathrm{diag}\,(0.1, 0.2, \ldots, 11.9, 12)$.

and has its diagonal elements equal to $\sigma_{\mathrm{r}}^2 = \mathrm{Tr}\left(\mathrm{Cov}\left(v_{\mathrm{r}}|s\right)\right)/n_{\mathrm{SC}}$. Clearly the off-diagonal entries in $\mathrm{Cov}\left(Fv_{\mathrm{r}}|s\right)$ are not zero, but they are neglected in what follows to get

$$\mathrm{Cov}\left(Fv_{\mathrm{r}}|s\right) \approx \sigma_{\mathrm{r}}^2 I. \tag{5.23}$$

Considering (5.1), (5.14), and (5.23), Each vector Gt_{2i} and Gt_{2i+1} is corrupted respectively by the real and imaginary parts of Fv_{r} and Fv_{g}, with $Fv_{\mathrm{r}} \sim \mathcal{CN}\left(0, \sigma_{\mathrm{r}}^2 I\right)$ and $Fv_{\mathrm{g}} \sim \mathcal{CN}\left(0, 2\mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_{n_{\mathrm{SC}}}^2\right)\right)$. Assuming that Fv_{r} and Fv_{g} are uncorrelated, each component of Gt_{2i} and Gt_{2i+1} will be corrupted by a zero-mean Gaussian noise with variance $\sigma_i^2 + \sigma_{\mathrm{r}}^2/2$. Using this in the design of the optimal precoding matrix (5.4) and decoding matrix (5.5), the updated MSE (see (5.3)) of the received chunk vector $E\left[\left\|\left(t - \widehat{t}\right)\right\|_2^2\right]$ is

$$\varepsilon = \sum_{i=\ell+1}^{n_{\mathrm{Ck}}} \lambda_i + \sqrt{\gamma} \sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_{\mathrm{c},i}^2}, \tag{5.24}$$

where $\sigma_{\mathrm{c},i}^2 = \left(\sigma_i^2 + \frac{\sigma_{\mathrm{r}}^2}{2}\right)$ and $\ell \leqslant n_{\mathrm{SC}}$ is the largest integer such that

$$\sqrt{\frac{\lambda_i \sigma_{\mathrm{c},i}^2}{\gamma}} - \sigma_{\mathrm{c},i}^2 > 0, \;\; i = 1, \ldots, \ell \tag{5.25}$$

with

$$\sqrt{\gamma} = \frac{\sum_{i=1}^{\ell} \sqrt{\lambda_i \sigma_{\mathrm{c},i}^2}}{\frac{p_{\mathrm{T}}}{2} + \sum_{i=1}^{\ell} \sigma_{\mathrm{c},i}^2}. \tag{5.26}$$

The number of chunks ℓ to be transmitted which minimizes ε has to be determined account-

ing for the fact that σ_{r}^2 depends on $q = n_{\mathrm{SC}} - \ell$, N_{g}, σ_{I}^2, and p_{I} [SPZ08].

5.5.2 Estimation of σ_{r}^2

An explicit expression of the evolution of σ_{r}^2 as a function of n_{SC}, q, N_{g}, σ_{I}^2, and p_{I} is very difficult to obtain. Thus, in this section, we will resort to a phenomenological model (PM) of σ_{r}^2 as a function of these parameters. First experiments have been conducted to characterize the structure of the model. Then the value of the model parameters are estimated via least-square estimation.

Two main channels have been considered, the first with $n_{\mathrm{SC}} = 256$ subchannels and the second of $n_{\mathrm{SC}} = 416$ subchannels, respectively. For both channels, Gaussian background noise with $N_{\mathrm{g}} = 2\sigma_{\mathrm{g}}^2 I$ and impulsive noise with $\sigma_{\mathrm{I}}^2 = 100$ are introduced. The variance of the background noise is adjusted in such a way that the *impulsive to background noise ratio* (INR) in dB, *i.e.*, $10\log_{10}\left(\sigma_{\mathrm{I}}^2/\sigma_{\mathrm{g}}^2\right)$ ranges from 10 dB to 30 dB with a step of 2 dB. The impulse probability p_{I} ranges from 0.5% to 3% with a step of 0.5%. Under these channel conditions, the variance of the residual noise σ_{r}^2 is evaluated once the FBMP algorithm has been employed for impulse estimation. It is obtained as the average of $\|v_{\mathrm{I}} - \widehat{v}_{\mathrm{I}}\|_2^2$, where $\widehat{v}_{\mathrm{I}}$ is obtained from the FBMP algorithm, see (5.12). One evaluates σ_{r}^2 considering different proportions of unused subchannels $r_{\mathrm{d}} = \frac{q}{n_{\mathrm{SC}}}$ ranging from 0.15 to 0.75 with a step of 0.05. Figures 5.5, 5.6 and 5.7 gather the evaluation of $\log_{10}\left(\sigma_{\mathrm{r}}^2\right)$ as a function of the different system parameters. Since the FBMP only uses the syndrome (5.11), which does not depend on the transmitted chunks, all evaluations are performed assuming that all-zero chunks are transmitted.

One observes in Figure 5.5 that $\log_{10}\left(\sigma_{\mathrm{r}}^2\right)$ evolves almost linearly with $\mathrm{INR}_{\mathrm{dB}} = 10\log_{10}\left(\mathrm{INR}\right)$ for the different values of r_{d} and p_{I} considered. Hence its evolution may be approximated as

$$\log_{10}\left(\sigma_{\mathrm{r}}^2\right) = \alpha_0\left(p_{\mathrm{I}}, r_{\mathrm{d}}\right) + \alpha_1\left(p_{\mathrm{I}}, r_{\mathrm{d}}\right)\mathrm{INR}_{\mathrm{dB}}, \tag{5.27}$$

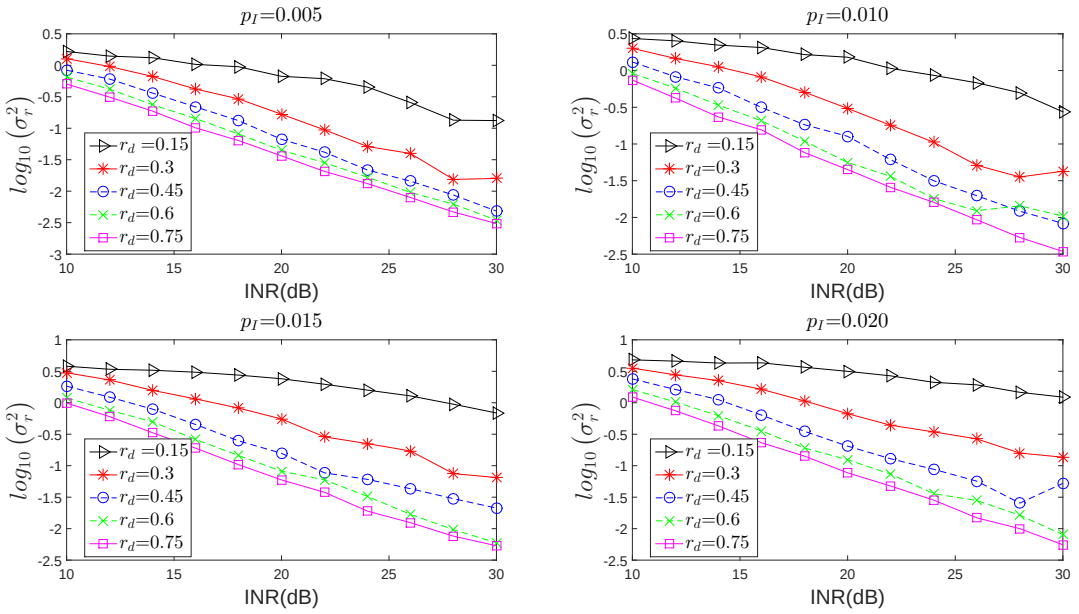


Figure 5.5: $\log_{10}(\sigma_{\mathrm{r}}^2)$ as a function of the INR in dB for different values of p_{I} and r_{d}

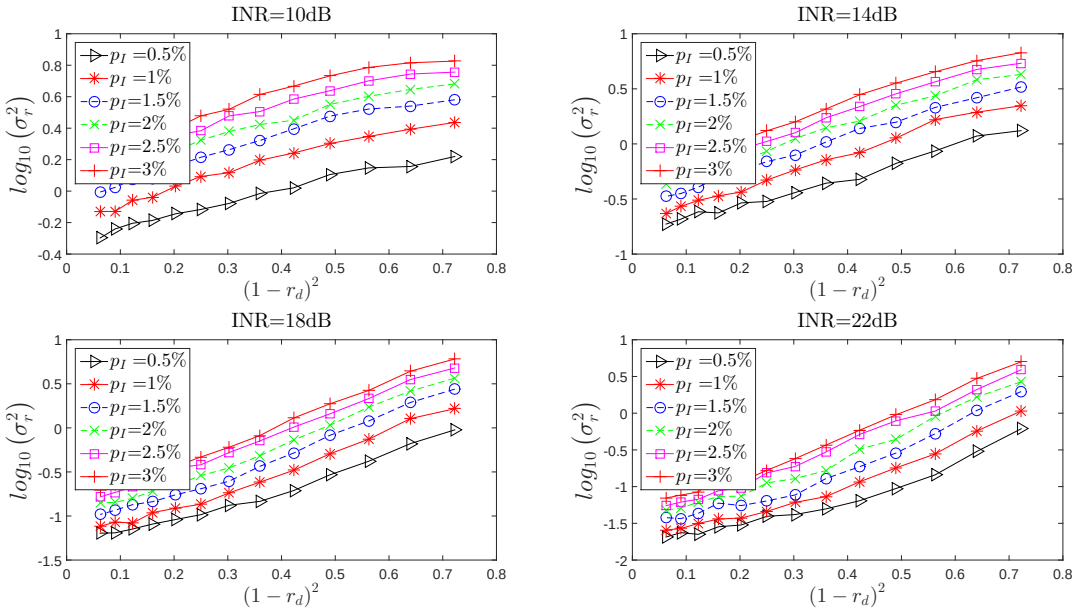


Figure 5.6: $\log_{10}(\sigma_{\mathrm{r}}^2)$ as a function of $(1-r_{\mathrm{d}})^2$ for different values of the INR in dB and p_{I}

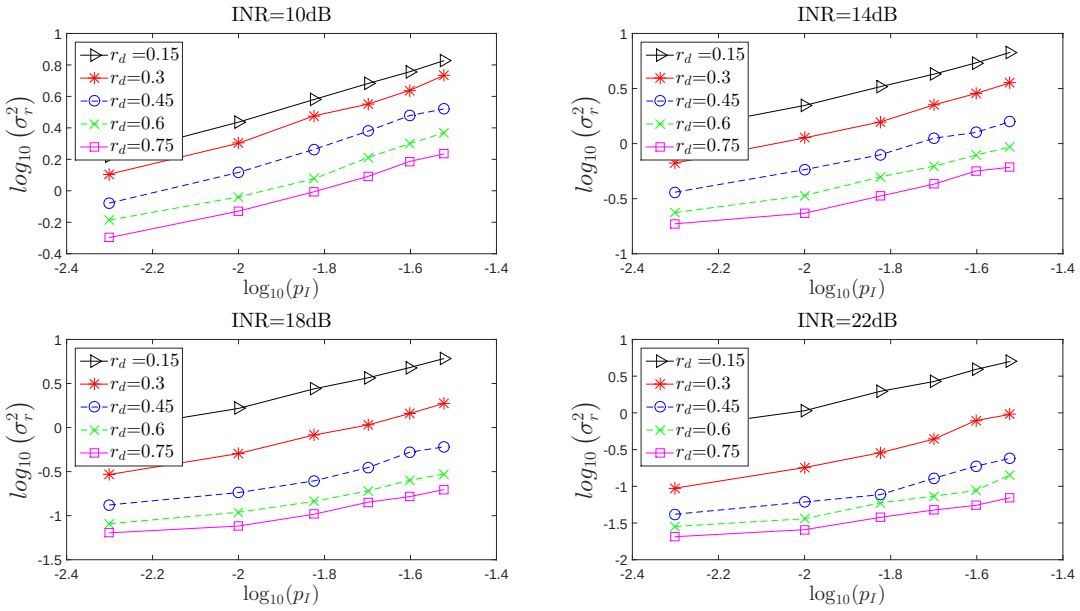


Figure 5.7: $\log_{10}(\sigma_{\mathrm{r}}^2)$ as a function of $\log_{10}(p_{\mathrm{I}})$ for different values of the INR in dB and of r_{d}

where $\alpha_0(p_{\mathrm{I}}, r_{\mathrm{d}})$ and $\alpha_1(p_{\mathrm{I}}, r_{\mathrm{d}})$ are functions of p_{I} and r_{d} to be determined. Figure 5.6 shows that $\log_{10}(\sigma_{\mathrm{r}}^2)$ also evolves linearly with $(1-r_{\mathrm{d}})^2$ for different values of $\mathrm{INR}_{\mathrm{dB}}$ and p_{I}. This evolution may thus be modeled as

$$\log_{10}\left(\sigma_{\mathrm{r}}^2\right) = \beta_0\left(p_{\mathrm{I}}, \mathrm{INR}_{\mathrm{dB}}\right) + \beta_1\left(p_{\mathrm{I}}, \mathrm{INR}_{\mathrm{dB}}\right)(1-r_{\mathrm{d}})^2, \tag{5.28}$$

where $\beta_0(p_{\mathrm{I}}, \mathrm{INR}_{\mathrm{dB}})$ and $\beta_1(p_{\mathrm{I}}, \mathrm{INR}_{\mathrm{dB}})$ are functions of p_{I} and $\mathrm{INR}_{\mathrm{dB}}$ to be determined. Finally, in Figure 5.7, $\log_{10}(\sigma_{\mathrm{r}}^2)$ has been represented as a function of $\log_{10}(p_{\mathrm{I}})$ and shows again an almost linear behavior. Therefore one may approximate $\log_{10}(\sigma_{\mathrm{r}}^2)$ as

$$\log_{10}\left(\sigma_{\mathrm{r}}^2\right) = \mu_0\left(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}}\right) + \mu_1\left(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}}\right)\log_{10}\left(p_{\mathrm{I}}\right), \tag{5.29}$$

where $\mu_0(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}})$ and $\mu_1(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}})$ are functions of r_{d} and $\mathrm{INR}_{\mathrm{dB}}$ to be determined.

The proposed PM of $\log_{10}(\sigma_{\mathrm{r}}^2)$ has to be simultaneously consistent with (5.27), (5.28), and (5.29). Starting from (5.29), one has considered the same structure for each $\mu_i(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}})$, $i = 0, 1$, namely

$$\begin{aligned}\mu_i\left(r_{\mathrm{d}}, \mathrm{INR}_{\mathrm{dB}}\right) &= \mu_{i,0} + \mu_{i,1}\mathrm{INR}_{\mathrm{dB}} + \mu_{i,2}\left(1-r_{\mathrm{d}}\right)^2 \\ &\quad + \mu_{i,3}\left(1-r_{\mathrm{d}}\right)^2 \mathrm{INR}_{\mathrm{dB}}\end{aligned} \tag{5.30}$$

Considering this structure, using (5.27), one deduces

$$\begin{aligned}\alpha_0(p_{\rm I}, r_{\rm d}) &= \mu_{0,0} + \mu_{1,0}\log_{10}(p_{\rm I}) + \mu_{0,2}(1-r_{\rm d})^2 \\ &+ \mu_{1,2}\log_{10}(p_{\rm I})(1-r_{\rm d})^2, \\ \alpha_1(p_{\rm I}, r_{\rm d}) &= \mu_{0,1} + \mu_{1,1}\log_{10}(p_{\rm I}) + \mu_{0,3}(1-r_{\rm d})^2 \\ &+ \mu_{1,3}\log_{10}(p_{\rm I})(1-r_{\rm d})^2,\end{aligned}$$

and using (5.28), one gets

$$\begin{aligned}\beta_0(p_{\rm I}, {\rm INR}_{\rm dB}) &= \mu_{0,0} + \mu_{1,0}\log_{10}(p_{\rm I}) + \mu_{0,1}{\rm INR}_{\rm dB} \\ &+ \mu_{1,1}\log_{10}(p_{\rm I}){\rm INR}_{\rm dB}, \\ \beta_1(p_{\rm I}, {\rm INR}_{\rm dB}) &= \mu_{0,2} + \mu_{1,2}\log_{10}(p_{\rm I}) + \mu_{0,3}{\rm INR}_{\rm dB} \\ &+ \mu_{1,3}\log_{10}(p_{\rm I}){\rm INR}_{\rm dB}.\end{aligned}$$

Considering all collected data, and using the PM (5.29), one may easily get a least-square estimate of the value of the parameter vectors $\boldsymbol{\mu}_i = (\mu_{i,0}, \ldots, \mu_{i,3})$, $i = 0, 1$. A different set of parameter vectors is associated to each channel. One gets

$$\begin{aligned}\boldsymbol{\mu}_0^{256} &= (2.6,\ -0.14,\ -1.71,\ 0.29) \\ \boldsymbol{\mu}_1^{256} &= (0.71,\ -0.003,\ -0.92,\ 0.1)\end{aligned} \tag{5.31}$$

for the channel with 256 subchannels, and

$$\begin{aligned}\boldsymbol{\mu}_0^{416} &= (2.6,\ -0.12,\ -1.79,\ 0.27) \\ \boldsymbol{\mu}_1^{416} &= (0.72,\ 0.007,\ -0.93,\ 0.09)\end{aligned} \tag{5.32}$$

for the channel with 416 subchannels. One observes that both sets of parameters have very close values.

The PM output for the values of the parameter vectors considered in (5.31) and (5.32) are compared with the experimental values, which are shown in Figures 5.8 and 5.9. In most of the cases, model output is very close to the values of $\log_{10}(\sigma_{\rm r}^2)$ obtained experimentally. The

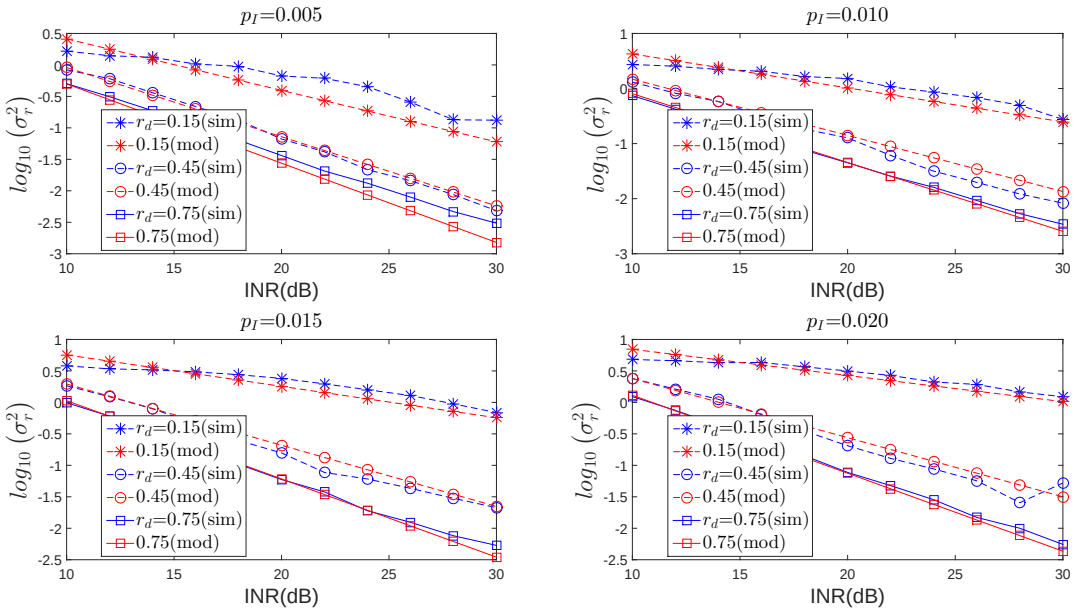


Figure 5.8: $\log_{10}(\sigma_{\mathrm{r}}^2)$ obtained from the PM (5.29) (mod) and from the simulations (sim) for the channel with 256 subchannels

gap is less than 0.5 in all cases. Consequently, the PM (5.29) provides a good estimate of σ_{r}^2 and can be used in (5.24) to evaluate the total distortion.

5.5.3 Optimization of sub-channel provisioning

This section describes the way the optimal proportion r_{d} of provisioned subchannels is evaluated, as a function of the system parameters, namely the background and impulsive noise characteristics, and the characteristics of the chunks. Here, one assumes a point-to-point communication.

For that purpose, one assumes that the number of subchannels n_{SC}, the noise variances affecting each subchannel $\left(2\sigma_1^2, \ldots, 2\sigma_{n_{\mathrm{SC}}}^2\right)$, the probability p_{I}, and the variance $2\sigma_{\mathrm{I}}^2$ of the impulse noise are all known. These parameters may be estimated by the receiver and fed back to the transmitter. Moreover, the transmitter already knows the vector of chunk variances $\lambda = (\lambda_1, \ldots, \lambda_{n_{\mathrm{CK}}})$.

Contrary to Section 5.5.2, where all σ_i^2, $i = 1, \ldots, n_{\mathrm{SC}}$ are assumed equal to σ_{g}^2 and the INR equal to $\sigma_{\mathrm{I}}^2/\sigma_g^2$, here, one considers that the INR depends on the subchannel index. As seen in

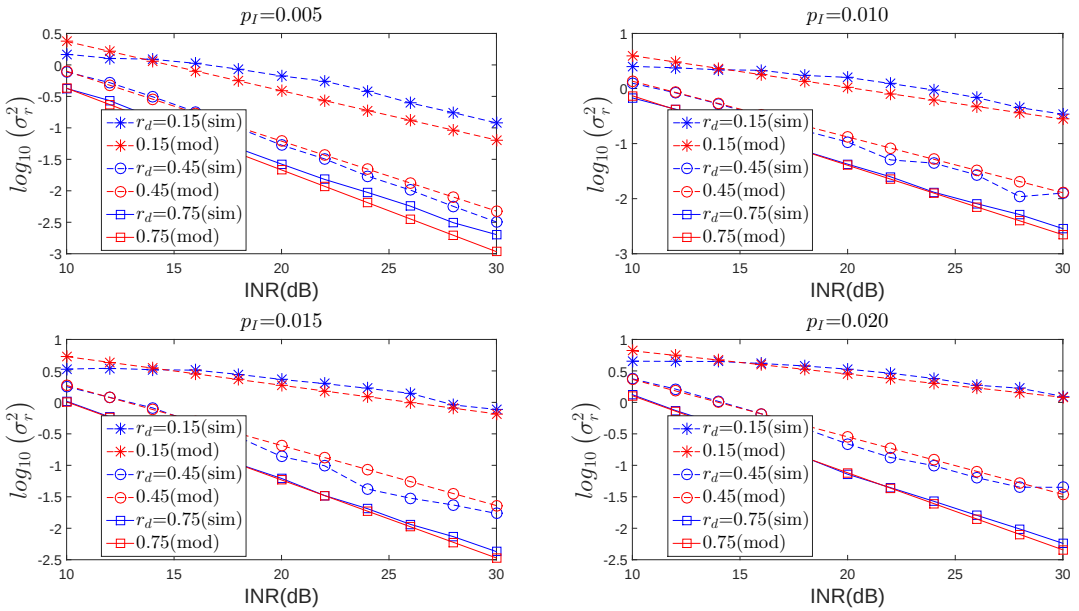


Figure 5.9: $\log_{10}(\sigma_{\mathrm{r}}^2)$ obtained from the PM (5.29) (mod) and from the simulations (sim) for the channel with 416 subchannels

(5.19), the covariance of impulse noise residual depends on the *average* INR, where the average is evaluated considering the noise variance of the last $n_{\mathrm{SC}} - \ell$ subchannels which are used for the impulse estimation. One thus introduces

$$\overline{\mathrm{INR}} = \frac{(n_{\mathrm{SC}} - \ell)\,\sigma_{\mathrm{I}}^2}{\sum_{i=\ell+1}^{n_{\mathrm{SC}}} \sigma_i^2}, \tag{5.33}$$

which will be used in the PM (5.29) in place of INR in what follows. A consequence of this substitution is that $\overline{\mathrm{INR}}$ depends on ℓ, the number of subchannels allocated for the chunk transmission. One also introduces the average SNR defined as

$$\overline{\mathrm{SNR}} = \frac{p_{\mathrm{T}}}{\sum_{i=1}^{n_{\mathrm{SC}}} \sigma_i^2}, \tag{5.34}$$

to evaluate the channel quality.

Then for a given value of r_{d},

1. one evaluates the target number ℓ_{t} of subchannels available for chunk transmission and $\overline{\mathrm{INR}}$ using (5.33),
2. σ_{r}^2 is then deduced from the PM (5.29),
3. the chunk reconstruction MSE $\varepsilon(r_{\mathrm{d}})$ is finally obtained from (5.24).

At Step 3, the actual number ℓ of subchannels used for chunk transmission may be less than the target number ℓ_{t}.

The minimization of $\varepsilon(r_{\mathrm{d}})$ may then be performed, *e.g.*, by exhaustive search, or by gradient descent to find

$$\widehat{r}_{\mathrm{d}} = \arg\min_{r_{\mathrm{d}}} \varepsilon(r_{\mathrm{d}}). \tag{5.35}$$

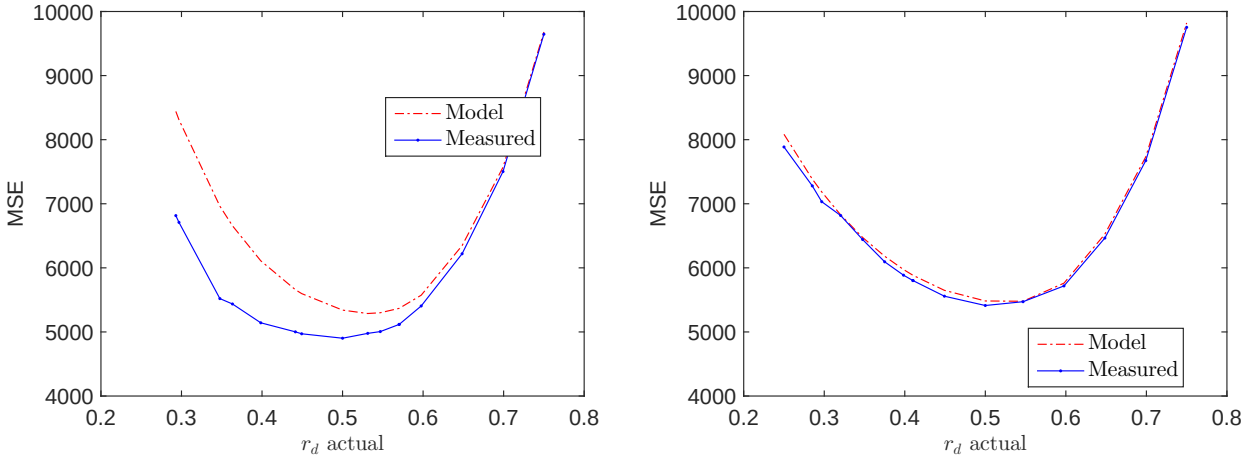
The version of the LVC scheme implementing the Optimal Subchannel Provisioning (OSP) with the Impulse noise Correction (IC) is denoted LVC-OSP-IC in what follows.

To illustrate the accuracy of the PM in the subchannel provisioning approach, one considers first a simple example where a group of chunks of the first GoP of the BQSquare and RaceHorses videos (their characteristics are detailed in Section 5.6.2) is transmitted considering different $\overline{\mathrm{SNR}}$ and p_{I} with $\sigma_{\mathrm{I}}^2 = 100$. The background channel noise is assumed zero-mean complex circular Gaussian with covariance matrix $\tau\mathrm{diag}(0.08,\ 0.084, 0.088, \ldots, 0.08 + (n_{\mathrm{SC}} - 1) \times 0.004)$, where τ is a scaling coefficient. At transmitter, various target values of r_{d} are chosen. The precoding and decoding matrices are updated accordingly to account for σ_{r}^2, as evaluated by the PM, see Section 5.5.1. At receiver side, the MSE (5.3) as estimated by (5.24) and the measured MSE for different values of r_{d} and different channel conditions are compared. Note that the actual value of r_{d} may be different from the target value, due to the total power constraint. This is especially true at low SNR.

Figure 5.10 illustrates the evolution of the predicted and measured MSE as a function of actual values of r_{d}. Two values of τ are considered, which leads to two different values of the total power constraint p_{T} and of the INR. The match is good, especially when the background noise is large (corresponding to large values of τ). One sees that using the PM, one is able to get a very good estimate of $\widehat{r}_{\mathrm{d}}$, without the need for a time-consuming exhaustive search for the value of r_{d} that minimizes the MSE.

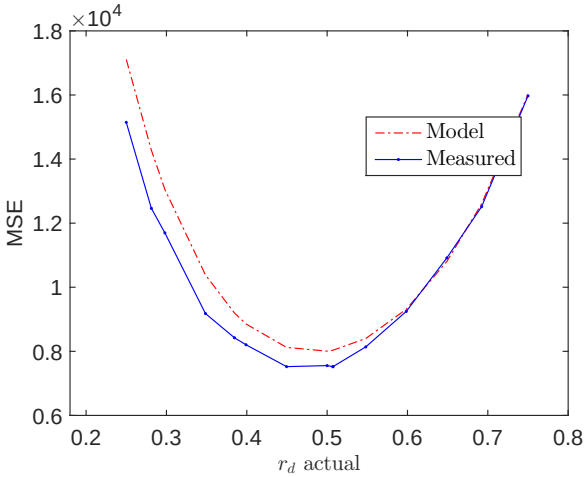
5.6 Simulation

Several variants of LVC schemes described in Section 5.6.1 are considered in what follows. The first are the baseline LVC with No Impulse noise Correction (LVC-NIC) as well as an LVC With Subchannel Provisioning but No Impulse noise Correction (LVC-WSP-NIC). The aim is

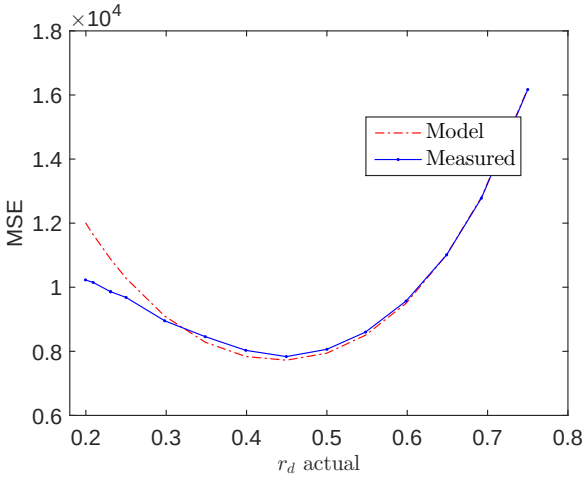


(a) BQSquare when $\tau = 1$

(b) BQSquare when $\tau = 7$



(c) RaceHorses when $\tau = 1$



(d) RaceHorses when $\tau = 7$

Figure 5.10: MSE as predicted by the model (5.24) and measured on simulations for one chunk vector of BQSquare and RaceHorses. $\overline{\text{SNR}}$= 9 dB, $p_{\text{I}} = 0.01$, and $\sigma_{\text{I}}^2 = 100$. (a) BQSquare when $\tau = 1$; (b) BQSquare when $\tau = 7$; (c) RaceHorses when $\tau = 1$; (d) RaceHorses when $\tau = 7$.

to analyze the impact of the impulse noise and of the subchannel provisioning on the LVC coding and transmission scheme. Considering an LVC scheme with Subchannel Provisioning and Impulse Correction (LVC-WSP-IC) (Section 5.4) and a scheme with Optimal Subchannel provisioning (LVC-OSP-IC) (Section 5.5.3) allows one to study the impact of impulse correction. The simulation parameters are detailed in Section 5.6.2. Simulation results are described in Section 5.6.3.

5.6.1 Compared LVC schemes

Four LVC schemes are compared in what follows.

In the LVC-NIC scheme, the number of transmitted chunks is only constrained by the bandwidth and total power constraints. Nevertheless, the effect of the impulse noise is taken into account by an increase of the variance of the background noise from σ_i^2 to $p_{\mathrm{I}}\sigma_{\mathrm{I}}^2 + \sigma_i^2$. The precoding and decoding matrices are adapted accordingly.

In the LVC-WSP-NIC scheme, a proportion r_{d} of subchannels is not used for chunk transmission. Consequently, the remaining chunks benefit from more transmission power. At receiver, only the subchannels on which chunks have been transmitted are considered. The other subchannels are not considered for impulse noise correction. The effect of the impulse noise is again captured by an increase of the background noise from σ_i^2 to $p_{\mathrm{I}}\sigma_{\mathrm{I}}^2 + \sigma_i^2$. The precoding and decoding matrices are adapted accordingly.

In the LVC-WSP-IC scheme, a proportion r_{d} of subchannels is used for impulse noise correction. The value of r_{d} is not optimized. This scheme is used to analyze the impact of the choice of r_{d} on the reconstruction MSE.

Finally, in the LVC-OSP-IC scheme, an optimal proportion r_{d} of subchannels is used for impulse noise correction. The optimization is performed using the PM model described in Section 5.5.3.

In all cases, metadata have to be transmitted to indicate the indexes and variances of the chunks, the subchannel noise variances of the reference channel, as well as the variance and probability of the impulse noise. With this information, each receiver (in case of multicast or broadcast) may rebuilt the precoding matrix and evaluate the decoding matrix optimized for the observed channel conditions. The amount of side information is of the same order of

magnitude as that of plain SoftCast [JK10a] and is neglected in what follows.

5.6.2 Simulation parameters

The considered channels consist of $n_{\mathrm{SC}} = 256$ or $n_{\mathrm{SC}} = 416$ subchannels. Each subchannel has a bandwidth $f_{\mathrm{SC}} = 24.414$ kHz. Using analog QAM and root-raised-cosine Nyquist filters with $\beta_{\mathrm{r}} = 30\,\%$ roll-off, one obtains a per-subchannel transmission rate

$$r_{\mathrm{SC}} = \frac{2 f_{\mathrm{SC}}}{1 + \beta_{\mathrm{r}}}, \tag{5.36}$$

here equal to $r_{\mathrm{SC}} = 37.56 \times 10^3$ real-valued symbols per second.

The two video sequences are taken from the MPEG test set used for the standardization of HEVC [OSS+12], namely BQSquare (Class D) and RaceHorse (Class C). One considers only the luminance component of each video. The frame rate is r_{F}. Considering GoPs of constant size n_{F}, and chunks of $n_{\mathrm{c}} \times n_{\mathrm{r}}$ pixels, each GoP contains

$$n_{\mathrm{Ck}} = \frac{n_{\mathrm{C}} n_{\mathrm{R}}}{n_{\mathrm{c}} n_{\mathrm{r}}} n_{\mathrm{F}}$$

chunks. The number of chunks a subchannel can transmit for the duration of a GoP is

$$\nu_{\mathrm{Ck}} = \frac{n_{\mathrm{F}}}{r_{\mathrm{F}}} \frac{r_{\mathrm{SC}}}{n_{\mathrm{r}} n_{\mathrm{c}}}. \tag{5.37}$$

For the typical values of the parameters considered in the simulations, $\nu_{\mathrm{Ck}} > 1$, *i.e.*, several chunks have to be transmitted on the same subchannel for the duration of a GoP. Therefore, given the number of subchannels n_{SC}, at most

$$n_{\mathrm{TrCk}} = \nu_{\mathrm{Ck}} n_{\mathrm{SC}}$$

chunks can be transmitted. For that purpose, the n_{TCk} chunks are ordered by decreasing variance and are partitioned into

$$n_{\mathrm{gCk}} = \frac{n_{\mathrm{Ck}}}{\nu_{\mathrm{Ck}}}$$

groups of ν_{Ck} chunks of similar variance. Consequently, ν_{Ck} precoding (and decoding) matrices

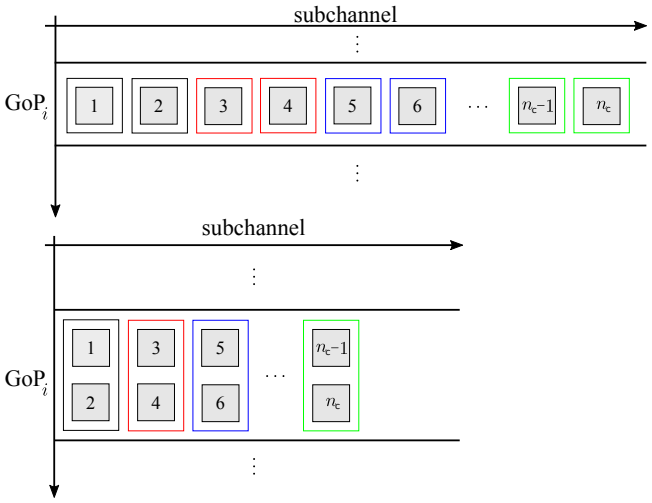


Figure 5.11: Organizations of the transmission of chunks of the i-th GoP: (top) A vector of chunks; (bottom) chunks of similar variance are gathered into groups of $\nu_{\mathrm{Ck}} = 2$ chunks, each group of chunk being transmitted over a dedicated subchannel.

have to be designed considering the n_{gCk} chunks of same index in each groups of chunks. Figure 5.11 illustrates the way chunks are transmitted over the available subchannels.

For impulse noise correction, the parameter D used in the FBMP (Section 5.4) is chosen equal to 5, which represents a compromise between complexity and performance as shown in [SPZ08].

The simulation parameters are shown in Table 5.2.

5.6.3 Simulation results

5.6.3.1 Impact of r_{d} on the efficiency of impulse noise correction

The average PSNR of the first 5 GoPs of BQSquarre and RaceHorses is evaluated for SNRs ranging from 0 dB to 20 dB (the SNR is evaluated without taking into account the impulse noise). The power p_{T} for one OFDM symbol is set equal to 2560. The variance and the probability of impulse noise are $\sigma_{\mathrm{I}}^2 = 100$ and $p_{\mathrm{I}} = 0.01$ or $p_{\mathrm{I}} = 0.02$. Then INR is computed from SNR, p_{T} and n_{SC}.

The schemes considered first are LVC-NIC, LVC-WSP-NIC, and LVC-WSP-IC. Additionally, one considers reference situations in absence of impulse noise. The corresponding schemes are denoted as LVC Without Impulse noise (LVC-WoI) and LVC With Subchannel Provisoning and Without Impluse noise (LVC-WSP-WoI), respectively.

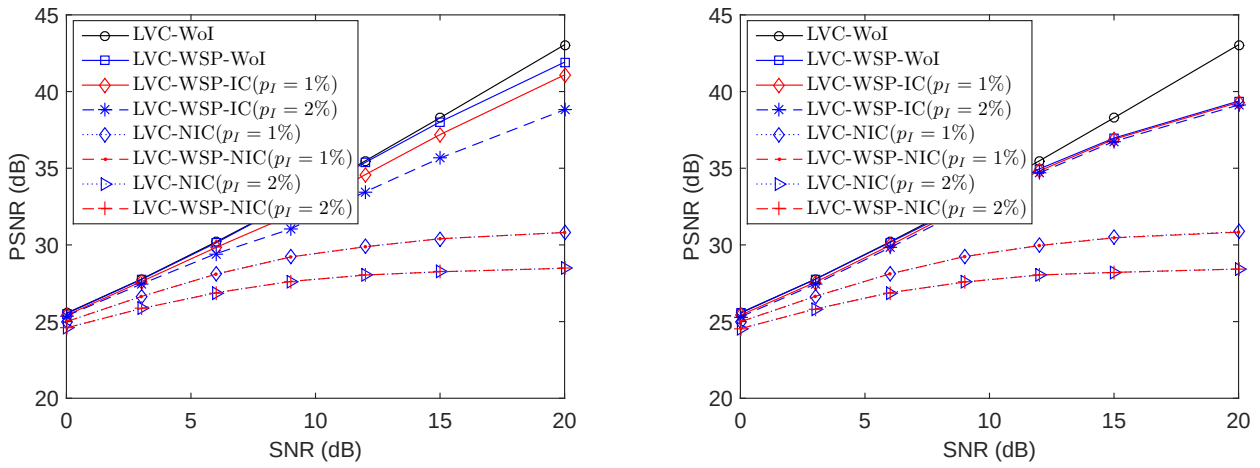
Simulation results are shown in Figures 5.12 and 5.13. Compared to the reference situation

Variable	Value		Signification
	BQSquare	RaceHorses	
$n_{\mathrm{R}} \times n_{\mathrm{C}}$	240×416	480×832	Frame size
r_{F}	60	30	Frame rate [Hz]
n_{F}	8	8	Nb of frames per GoP
$n_{\mathrm{r}} \times n_{\mathrm{c}}$	30×32	30×32	Size of a chunk
n_{Ck}	832	3328	Nb of chunks per GoP
n_{SC}	256	416	Nb of subchannels
ν_{Ck}	3	8	Nb of chunks a subchannel transmits per GoP
n_{gCk}	256	416	Nb of group of chunks
n_{TrCk}	768	3328	Nb of chunks transmitted per GoP

Table 5.2: Simulation parameters

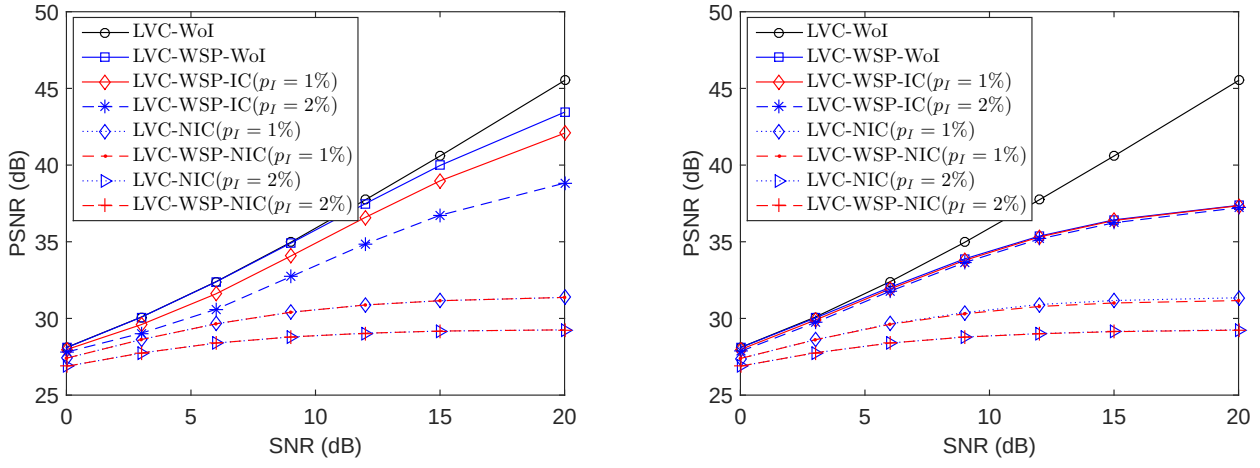
LVC-WoI, when $p_{\mathrm{I}} = 0.01$, without correction, the impulse noise leads to a PSNR decrease ranging from 0.6 dB at low channel SNR to 13 dB at high SNR. When $p_{\mathrm{I}} = 0.02$, the PSNR decrease is between 1 dB at low channel SNR and 15 dB at high SNR. Depending on the number of provisioned subchannels, the impulse noise correction may only be partial (when p_{I} is large and r_{d} is small) or complete (when r_{d} is large enough). In the second case, LVC-WSP-IC and LVC-WSP-WoI perform similarly. Nevertheless, r_{d} should not be chosen too large, since the PSNR in absence of impulse noise at high SNR decreases when r_{d} increases, due to the additionally dropped chunks.

Figures 5.14 and 5.15 represent the gains obtained by LVC-WSP-IC compared to LVC-NIC at different SNRs and for different target values of r_{d} taken in $\mathcal{R} = \{0.25, 0.33, 0.41, 0.5, 0.66, 0.75\}$. One observes that the optimal value of r_{d} depends on the value of the channel SNR. At low SNRs, r_{d} should be large, whereas at large SNRs, r_{d} may be reduced. This is mainly due to the fact that at low SNR, the INR is low and impulse noise identification is difficult with few syndrome samples. At high SNR, the INR increases, and it becomes easier to perform impulse noise identification .



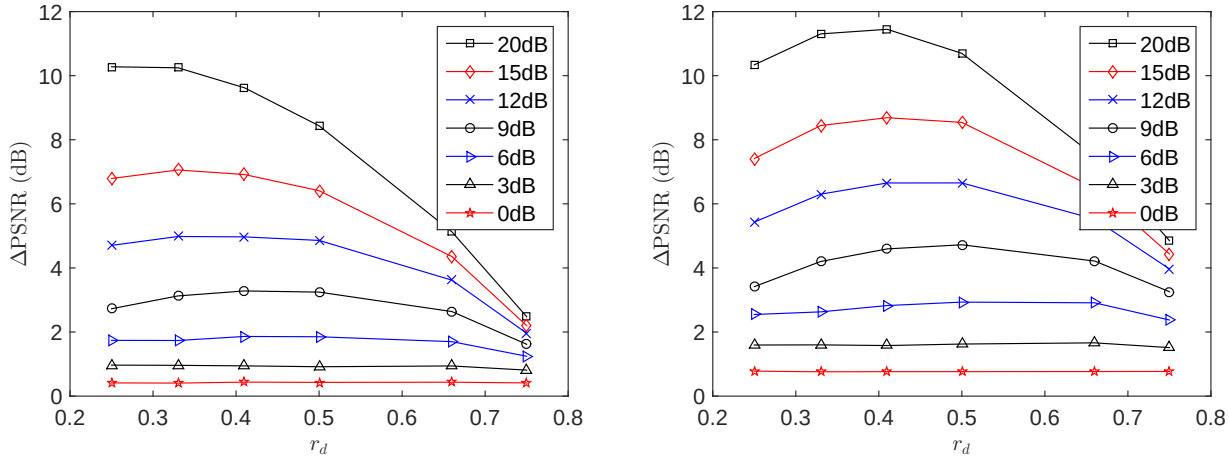
(a) $r_{\mathrm{d}} = 0.25$ (b) $r_{\mathrm{d}} = 0.5$

Figure 5.12: PSNR evolution as a function of SNR (dB) and r_{d} for BQSquare when $\sigma_{\mathrm{I}}^2 = 100$



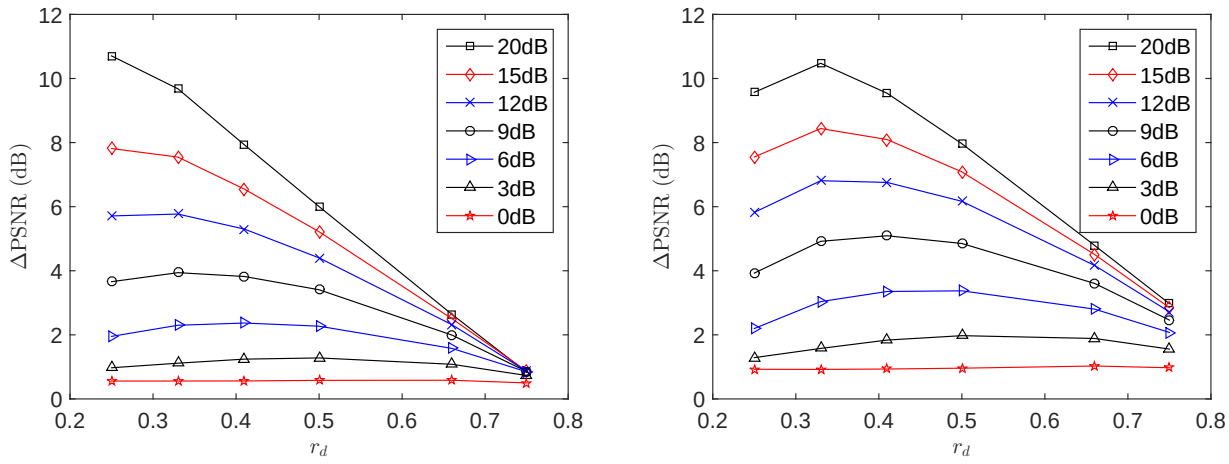
(a) $r_{\mathrm{d}} = 0.25$ (b) $r_{\mathrm{d}} = 0.5$

Figure 5.13: PSNR evolution as a function of SNR (dB) and r_{d} for RaceHorse when $\sigma_{\mathrm{I}}^2 = 100$



(a) $p_{\mathrm{I}} = 0.01$ (b) $p_{\mathrm{I}} = 0.02$

Figure 5.14: PSNR gain of LVC-WSP-IC compared to LVC-NIC for different r_{d} for BQSquare when $\sigma_{\mathrm{I}}^2 = 100$,



(a) $p_{\mathrm{I}} = 0.01$ (b) $p_{\mathrm{I}} = 0.02$

Figure 5.15: PSNR gain of LVC-WSP-IC compared to LVC-NIC for different r_{d} for RaceHorses when $\sigma_{\mathrm{I}}^2 = 100$

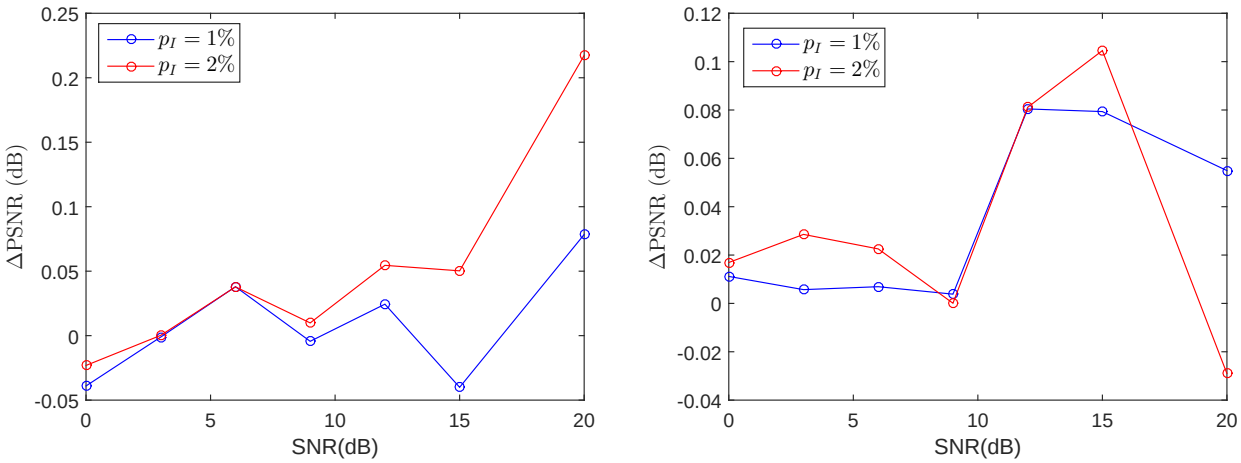
5.6.3.2 Optimal subchannel provisioning

Considering similar channel conditions as in the previous section, Figure 5.16 represents the PSNR differences between LVC-OSP-IC and LVC-WSP-IC. The choice of r_{d} in the case of LVC-OSP-IC is taken as the value minimizing the PM. In the case of LVC-WSP-IC, it corresponds to the value in $\mathcal{R}$ maximizing the PSNR for each SNR value. In most of the cases, LVC-OSP-IC provides better results (positive PSNR difference), since the search for the optimal r_{d} is in a larger set which is [0.15 0.75]. In some cases, LVC-OSP-IC may not perform as well as LVC-WSP-IC due to a mismatch of the PM. Nevertheless, the PNSR loss remains less than 0.05 dB.

Finally, Figure 5.17 shows reconstructed frames with LVC-NIC and LVC-OSP-IC when $\sigma_{\mathrm{I}}^2 = 100$, $p_{\mathrm{I}} = 0.01$, and SNR = 15 dB. A gain of 7.5 dB is observed for both video sequences when the impulse noise correction is performed. Full sequences are available at *https://drive.google.com/drive/folders/13LB5nR3nY79bF3CEMUl41HY4Bc_ekhBF*.

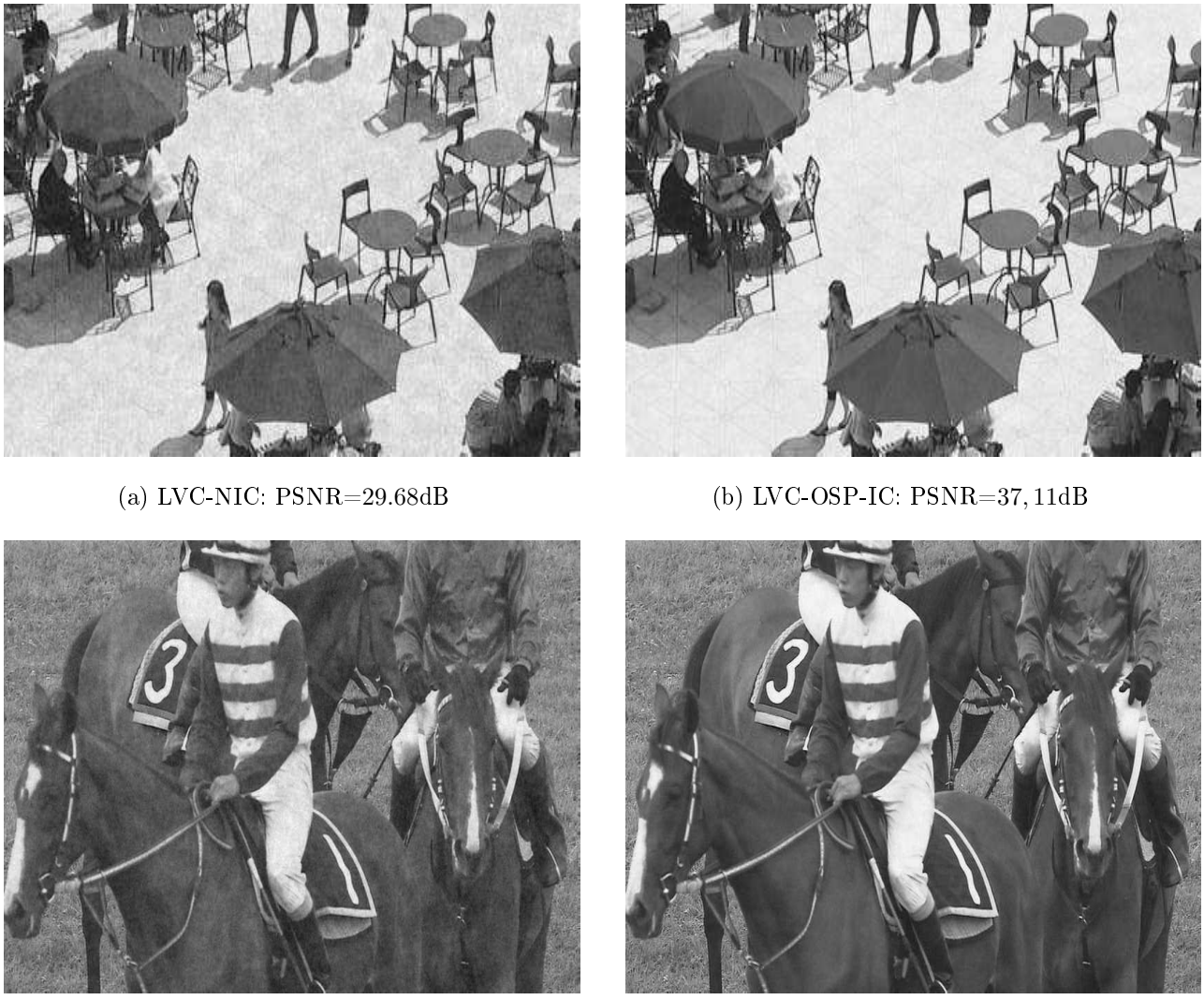
5.6.3.3 Analysis of the effect of mismatched channel conditions

The aim of this section is to analyze the effect of a mismatch between the characteristics of the channel impulse noise (used for the design of the precoding matrix and in the subchannel provisioning mechanisms) and the characteristics of the impulse noise perceived by a receiver. This aspect is important when LVC streams are broadcast to several users experiencing different



(a) BQSquare (b) RaceHorse

Figure 5.16: PSNR differences between LVC-OSP-IC and the best PSNR achieved by LVC-WSP-IC when r_{d} is searched in $\mathcal{R} = \{0.25, 0.33, 0.41, 0.5, 0.66, 0.75\}$



(a) LVC-NIC: PSNR=29.68dB (b) LVC-OSP-IC: PSNR=37, 11dB

(c) LVC-NIC: PSNR=30.83dB (d) LVC-OSP-IC: PSNR=38.64dB

Figure 5.17: First frame of BQSquare and RaceHorses. $\sigma_{\mathrm{I}}^2 = 100$, $p_{\mathrm{I}} = 0.01$ and SNR=15dB. (a) and (c) by using LVC-NIC; (b) and (d) by using LVC-OSP-IC.

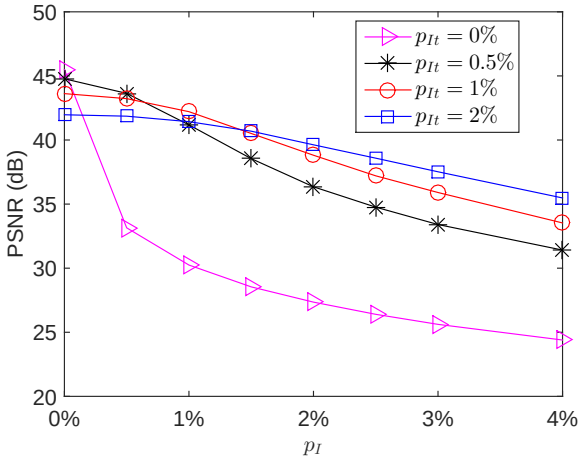


Figure 5.18: Effect of the mismatch between the target impulse noise probability p_{It} and the actual impulse noise probability p_{I} when $\sigma_{\mathrm{I}}^2 = 100$ and SNR $= 20\,\mathrm{dB}$ (for the RaceHorses video sequence)

channel conditions.

In the following experiments, the variance of impulse noise is $\sigma_{\mathrm{I}}^2 = 100$ and the channel SNR is set equal to 20 dB. One considers several target impulse noise probabilities p_{It} chosen equal to 0%, 0.5%, 1%, or 2% for the LVC-OSP-IC scheme. At receiver side, the parameters of impulse noise correction (FBMP algorithm) involve the actual channel impulse noise probability. Moreover the variance of the impulse noise residual variance which is estimated by the PM (5.29) is used to update the decoding matrix. The parameters of the RaceHorses video are in Table 5.2.

PSNR results for actual impulse noise probabilities ranging from 0% to 4% are shown in Figure 5.18. As expected, the PSNR decreases as p_{I} increases. The performance is best when p_{I} matches p_{It}. Choosing a large p_{It} improves the robustness to a larger p_{I}, but the price to be paid is a lower PSNR when p_{I} is smaller than p_{It}. Figure 5.18 shows that even if a small $p_{\mathrm{It}} = 0.5\%$ is chosen, in case of mismatch, the PSNR decrease is much smoother than in absence of subchannel provisioning for impulse noise mitigation.

5.7 Conclusion

This chapter considers the transmission of SoftCast-based encoded videos over channels affected by impulse noise. Albeit LVC schemes are relatively robust to impulse noise, at high channel SNR, a significant performance loss in terms of PSNR may be observed.

This chapter uses a Fast Bayesian Matching Poursuit algorithm for impulse noise identification and removal. This requires the provisioning of some subchannels on which no information is transmitted. At receiver side, the samples received on these subchannels are used to estimate the realizations of the impulse noise. The price to pay for the subchannel provisioning is a decrease of the nominal PSNR that may be obtained in absence of impulse noise. A trade-off has thus to be found between impulse noise correction efficiency and nominal PSNR reduction.

A phenomenological model has been proposed to describe the variance of the noise residual after impulse noise estimation and removal. This model takes as input the channel and noise characteristics as well as the number of provisioned subchannels. Combined with a model of the PSNR evolution as a function of the channel and video characteristics and of the number of provisioned subchannels, it is possible to optimize the proportion of subchannel to provision. The parameters of the PM have been adjusted for different channel characteristics and one has observed that it leads to accurate estimates of the optimal proportion of subchannel to provision once the characteristics of the video are known.

The performance of the optimal subchannel provisioning algorithm combined with the impulse noise mitigation algorithm has been evaluated on two reference video sequences. Provided that the amount of provisioned subchannels is not too large, the nominal PSNR decrease remains moderate, while the impulse noise can be efficiently mitigated.

Future work will be dedicated to the evaluation of the optimal amount of subchannels to provision in case of LVC for video multicast over channels prone to impulse noise.

Chapter 6

Conclusions and Perspectives

6.1 Conclusions

SoftCast-based LVC schemes have attracted a lot of attention in the last decade. Such schemes provide decoded video quality at different receivers to be commensurate with their experienced channel quality.

Nevertheless, SoftCast-based video transmission under per-subchannel power constraints and transmission over channels prone to impulse noise have not been considered. Per-subchannel power constraints can be found in PLT channels or when considering multi-antenna transmission over different antennas. In the second problem, if impulse noise has high amplitude or is bursty, may lead to significantly degraded received video.

We have addressed these two issues in this thesis. At first, by using multi-level water filling and a solution to an inverse eigenvalue problem, we have proposed an optimal power allocation algorithm for SoftCast-based video transmission under per-subchannel power constraint. Furthermore, inspired by multi-water filling, we also have proposed three lower complexity sub-optimal power allocation algorithms. They can reduce significantly the execution time and have negligible performance loss compared to optimal allocation algorithms. All of them have significant performance gain over a straightforward extension method of SoftCast in point-to-point communication and in multicast situation.

For the second issue, the mitigation of impulse noise, it is necessary to perform subchannels provisioning and over which there are no data transmitted. Then we use a Fast Bayesian Matching Poursuit algorithm to estimate the impulse noise samples and correct them. The

problem here is that a trade-off has to be found between the subchannels provisioning for impulse noise correction and subchannels for data transmission. To address this problem, we proposed a phenomenological model of the impulse noise correction residual error. This model allows one to estimate the optimal number of subchannels to provision as a function of the chunk vector variances and channel conditions, *e.g.,* SNR, variance and probability of impulse noise. Simulation results have shown the accuracy of this method to find an optimal trade-off. Moreover, the video performance with impulse noise mitigation is significantly better than when there is no impulse noise correction. In the multicast situation, the results show that a small number of subchannels provisioned is helpful to improve the robustness of video transmission under various channel conditions.

6.2 Perspectives

Three short and medium term research directions are detailed in what follows.

6.2.1 Precoding matrix design for multicast

In Section 3.5, the problem of precoding matrix design in a multicast scenario has been presented. We have not provided an optimal solution to this problem. The approach we considered was to compute the precoding matrix for multicast in an analytic way (using KKT conditions to find a closed-form expression) and several approximations, for example the SNR of subchannels is assumed to be large. Nevertheless, on some toy examples, we have seen that the solutions obtained via numerical optimization provide a better performance than those obtained using our proposed approximated solution. Numerical methods have thus to be considered to compute precoding matrix for SoftCast-based video transmission in multicast or broadcast situation.

On the other hand, the objective function to be minimized was the average of MSE of the receivers. An alternative criterion may be to minimize the worst MSE among receivers (min-max approach). In this case, under total power constraint, from (3.8), the precoding and decoding design problem can be reformulated as

$$\begin{array}{ll} \min\limits_{G,H_1,\ldots,H_k} \max\limits_{i=1,\ldots,k} & \varepsilon_i = \mathrm{tr}\left(H_i G\Lambda G^T H_i^T - H_i G\Lambda - \Lambda^T G^T H_i^T\right) + \mathrm{tr}\left(H_i N_i H_i^T + \Lambda\right) \\ \text{s.t} & \mathrm{tr}\left(G\Lambda G^T\right) \leqslant p_{\mathrm{T}}, \end{array} \tag{6.1}$$

where k is the number of receivers.

A problem similar to (6.1) has been considered in [KR13] in the context of relay-assisted multicast. An optimal iterative method has been proposed. First, some G is chosen that satisfies the total power constraint. Then, optimal H_is are computed from (3.10). Next, given the H_is, (6.1) is converted to a convex Semi-Definite Programming (SDP) problem. Let us introduce the constraints

$$G\Lambda G^T \preccurlyeq \Phi$$

and

$$H_i G\Lambda G^T H_i^T - H_i G\Lambda - \Lambda^T G^T H_i^T \preccurlyeq \Psi_i$$

and a real valued slack variable t. Here, $A \preccurlyeq B$ indicates that $B - A$ is positive semi-definite. Then using Shur's complement, (6.1) can rewritten as

$$\begin{array}{ll} \min\limits_{G,\Psi_i,\Phi} & t \\ \text{s.t.} & \mathrm{tr}\left(\Psi_i\right) + \mathrm{tr}\left(H_i N_i H_i^T + \Lambda\right) \leqslant t \\ & \begin{bmatrix} \Psi_i + H_i G\Lambda + \Lambda^T G^T H_i^T & H_i G\Lambda^{\frac{1}{2}} \\ \left(H_i G\Lambda^{\frac{1}{2}}\right)^T & I \end{bmatrix} \succcurlyeq 0 \\ & i = 1, \ldots, k \\ & \mathrm{tr}\left(\Phi\right) \leqslant p_{\mathrm{T}} \\ & \begin{bmatrix} \Phi & G \\ G^T & \Lambda^{-1} \end{bmatrix} \succcurlyeq 0. \end{array} \tag{6.2}$$

The SDP problem (6.2) may then be solved using classical tools such as the CVX toolbox. An alternative low-complexity solution has been proposed in [KR13]. One has to determine whether this solution may also be applied in the context of SoftCast.

6.2.2 Optimization of the amount of Metadata

Another issue is that we do not know yet the impact of accuracy of metadata on the whole performance. It is necessary to optimize the allocation between the metadata and chunk coefficients. For example if they are transmitted together under bandwidth limitation, we may increase the chunk size to reduce the metadata, or in another way to keep the chunk size and decrease the bandwidth for metadata.

6.2.3 Application of Deep Learning to SoftCast schemes

On the whole, joint source-channel coding scheme is a prominent coding technology. The key point is to exploit the channel conditions in source compression and use source information in channel error protection. This research domain has been considered since a while and has a lot of successful results.

Recently, deep learning and convolutional neural network (CNN) have been applied to joint source-channel coding problems. For example to text transmission [FRG18] and to image transmission [BKG18], where the latter has used the linearity property of SoftCast. In [YFS18], CNN are used at decoder side of SoftCast-based video transmission schemes. At encoder side, since SoftCast does not use the digital scheme for compression, there is still redundant information and consequently the number of symbols need to be transmitted is large. To address this problem, the Shannon-Kotel'nikov Mapping is applied as shown in Section 2.2.6. We also can try to use Linear factor model [GBCB16, Chapter 13] to find a sparse vector that can reconstruct the original vector in a linear way. In this case, the number of symbols need to be transmitted can be reduced.

Bibliography

[ACAB07] Thomas André, Marco Cagnazzo, Marc Antonini, and Michel Barlaud, *Jpeg2000-compatible scalable scheme for wavelet-based video coding*, Journal on Image and Video Processing **2007** (2007), no. 1, 9–9.

[ANQC14] Tareq Y Al-Naffouri, Ahmed A Quadeer, and Giuseppe Caire, *Impulse noise estimation and removal for ofdm systems*, IEEE Transactions on Communications **62** (2014), no. 3, 976–989.

[BKG18] Eirina Bourtsoulatze, David Burth Kurka, and Deniz Gunduz, *Deep joint source-channel coding for wireless image transmission*, arXiv preprint arXiv:1809.01733 (2018).

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.

[CCA+07] Marco Cagnazzo, Filippo Castaldo, Thomas André, Marc Antonini, and Michel Barlaud, *Optimal motion estimation for wavelet motion compensated video coding*, IEEE Transactions on Circuits and Systems for Video Technology **17** (2007), no. 7, 907–911.

[Chu98] M. T. Chu, *Inverse eigenvalue problems*, SIAM review **40** (1998), no. 1, 1–39.

[CK15] M. Cagnazzo and M. Kieffer, *Shannon-Kotelnikov mappings for SoftCast-based joint source-channel video coding*, Proc. IEEE ICIP, IEEE, 2015, pp. 1085–1089.

[CLCW14a] Hao Cui, Chong Luo, Chang Wen Chen, and Feng Wu, *Robust linear video transmission over rayleigh fading channel*, IEEE Transactions on Communications **62** (2014), no. 8, 2790–2801.

[CLCW14b] ———, *Robust uncoded video transmission over wireless fast fading channel*, INFOCOM, 2014 Proceedings IEEE, IEEE, 2014, pp. 73–81.

[Cov72] Thomas Cover, *Broadcast channels*, IEEE Transactions on Information Theory **18** (1972), no. 1, 2–14.

[CR08] Emmanuel J Candes and Paige A Randall, *Highly robust error correction byconvex programming*, IEEE Transactions on Information Theory **54** (2008), no. 7, 2829–2840.

[CRA+16] Rafael E Carrillo, Ana B Ramirez, Gonzalo R Arce, Kenneth E Barner, and Brian M Sadler, *Robust compressive sensing of sparse signals: a review*, EURASIP Journal on Advances in Signal Processing **2016** (2016), no. 1, 108.

[CSY+13] H. Cui, Z. Song, Z. Yang, C. Luo, R. Xiong, and F. Wu, *Cactus: A hybrid digital-analog wireless video communication system*, Proc. ACM Int. Conf. on Modeling, Analysis & Simulation of Wireless and Mobile Systems, 2013, pp. 273–278.

[CT06] Thomas M Cover and Joy A Thomas, *Elements of information theory*, John Wiley & Sons, 2006.

[CXL+15] Hao Cui, Ruiqin Xiong, Chong Luo, Zhihai Song, and Feng Wu, *Denoising and resource allocation in uncoded video transmission*, IEEE Journal of Selected Topics in Signal Processing **9** (2015), no. 1, 102–112.

[CZX+17] Zhe Chen, Xu Zhang, Yuedong Xu, Jie Xiong, Yu Zhu, and Xin Wang, *Muvi: Multiview video aware transmission over mimo wireless systems*, IEEE Transactions on Multimedia **19** (2017), no. 12, 2788–2803.

[ETS15] ETSI TR 103 343 V1.1.1, *Powerline telecommunications (PLT); powerline HDMI® analysis for very short range link HD and UHD applications*, Tech. report, ETSI, Dec. 2015.

[FKAWO18a] Takuya Fujihashi, Toshiaki Koike-Akino, Takashi Watanabe, and Philip V Orlik, *Freecast: Graceful free-viewpoint video delivery*, IEEE Transactions on Multimedia (2018).

[FKAWO18b] ———, *High-quality soft video delivery with gmrf-based overhead reduction*, IEEE Transactions on Multimedia **20** (2018), no. 2, 473–483.

[FLWZ14] Nianfei Fan, Yu Liu, Quan Wang, and Lin Zhang, *Hybrid digital-analog video multicast scheme based on h. 264/avc and softcast*, 2014 International Symposium on Wireless Personal Multimedia Communications (WPMC), IEEE, 2014, pp. 277–282.

[FRG18] Nariman Farsad, Milind Rao, and Andrea Goldsmith, *Deep learning for joint source-channel coding of text*, arXiv preprint arXiv:1802.06832 (2018).

[FWZ+12] Xiaopeng Fan, Feng Wu, Debin Zhao, Oscar C Au, and Wen Gao, *Distributed soft video broadcast (dcast) with explicit motion*, Data Compression Conference (DCC), 2012, IEEE, 2012, pp. 199–208.

[FWZA13] X. Fan, F. Wu, D. Zhao, and O. C. Au, *Distributed wireless visual communication with power distortion optimization*, IEEE Trans. Circuits and Systems for Video Technology **23** (2013), no. 6, 1040–1053.

[FXWZ12] X. Fan, R. Xiong, F. Wu, and D. Zhao, *Wavecast: Wavelet based wireless video broadcast using lossy transmission*, Proc. IEEE VCIP (San Diego, CA), 2012, pp. 1–6.

[FXZW15] Xiaopeng Fan, Ruiqin Xiong, Debin Zhao, and Feng Wu, *Layered soft video broadcast for heterogeneous receivers.*, IEEE Trans. Circuits Syst. Video Techn. **25** (2015), no. 11, 1801–1814.

[GARRM05] Bernd Girod, Anne Margot Aaron, Shantanu Rane, and David Rebollo-Monedero, *Distributed video coding*, Proceedings of the IEEE **93** (2005), no. 1, 71–83.

[GBCB16] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio, *Deep learning*, vol. 1, MIT press Cambridge, 2016.

[Gob65] T Goblick, *Theoretical limitations on the transmission of data from analog sources*, IEEE Trans. Information Theory **11** (1965), no. 4, 558–567.

[GRV03] Michael Gastpar, Bixio Rimoldi, and Martin Vetterli, *To code, or not to code: Lossy source-channel communication revisited*, IEEE Transactions on Information Theory **49** (2003), no. 5, 1147–1158.

[HFR09] Fredrik Hekland, Pal Anders Floor, and Tor A Ramstad, *Shannon-Kotelnikov mappings in joint source-channel coding*, IEEE Transactions on Communications **57** (2009), no. 1, 94–105.

[HLL+15] Dongliang He, Chong Luo, Cuiling Lan, Feng Wu, and Wenjun Zeng, *Structure-preserving hybrid digital-analog video delivery in wireless networks*, IEEE Transactions on Multimedia **17** (2015), no. 9, 1658–1670.

[HLL+17] D. He, C. Lan, C. Luo, E. Chen, F. Wu, and W. Zeng, *Progressive pseudo-analog transmission for mobile video streaming*, IEEE Trans. Multimedia **19** (2017), no. 8, 1894–1907.

[JK10a] S. Jakubczak and D. Katabi, *Softcast: Clean-slate scalable wireless video*, Proc. 48th Annual Allerton Conference on Communication, Control, and Computing, 2010, pp. 530–533.

[JK10b] ———, *Softcast: one-size-fits-all wireless video*, Proc. ACM SIGCOMM (New-York, NY), 2010, pp. 449–450.

[JSKG10a] S. Jakubczak, J. Sun, D. Katabi, and V. Goyal, *Analog transmission of degradable content over wireless channels*, Tech. report, Technical report, MIT, 2010. http://people. csail. mit. edu/szym/softcast/tr. pdf, 2010, It compare the performance of analog and digital scheme in unicast and broadcast, and all is in the Total power constraint.

[JSKG10b] Szymon Jakubczak, J Sun, Dina Katabi, and V Goyal, *Analog transmission of degradable content over wireless channels*, Technical report (2010).

[JSKG11] S. Jakubczak, J. Z. Sun, D. Katabi, and V. K. Goyal, *Performance regimes of uncoded linear communications over AWGN channels*, Proc. 45th Annual Conference on Information Sciences and Systems (CISS), 2011.

[KR13] Muhammad RA Khandaker and Yue Rong, *Precoding design for mimo relay multicasting*, IEEE Transactions on Wireless Communications **12** (2013), no. 7, 3544–3555.

[Lam11] Lutz Lampe, *Bursty impulse noise detection by compressed sensing*, Power Line Communications and Its Applications (ISPLC), 2011 IEEE International Symposium on, IEEE, 2011, pp. 29–34.

[LHL+14a] X. L. Liu, W. Hu, C. Luo, Q. Pu, and F. Wu, *Compressive image broadcasting in MIMO systems with receiver antenna heterogeneity*, Signal Processing: Image Communication **29** (2014), no. 3, 361–374.

[LHL+14b] X. L. Liu, W. Hu, C. Luo, Q. Pu, F. Wu, and Y. Zhang, *Parcast+: Parallel video unicast in MIMO-OFDM WLANs*, IEEE Trans. Multimedia **16** (2014), no. 7, 2038–2051.

[LHP+12] Xiao Lin Liu, Wenjun Hu, Qifan Pu, Feng Wu, and Yongguang Zhang, *Parcast: Soft video delivery in mimo-ofdm wlans*, Proceedings of the 18th annual international conference on Mobile computing and networking, ACM, 2012, pp. 233–244.

[LLX+17] Fei Liang, Chong Luo, Ruiqin Xiong, Wenjun Zeng, and Feng Wu, *Hybrid digital-analog video delivery with shannon-Kotelnikov mappings*, IEEE Transactions on Multimedia (2017), To appear.

[LLZW18] Cuiling Lan, Chong Luo, Wenjun Zeng, and Feng Wu, *A practical hybrid digital-analog scheme for wireless video transmission*, IEEE Transactions on Circuits and Systems for Video Technology **28** (2018), no. 7, 1634–1647.

[LNE13] Jing Lin, Marcel Nassar, and Brian L Evans, *Impulsive noise mitigation in powerline communications using sparse bayesian learning*, IEEE Journal on Selected Areas in Communications **31** (2013), no. 7, 1172–1183.

[LP76] K. H. Lee and D. P. Petersen, *Optimal linear coding for vector channels*, IEEE Trans. On Communications **24** (1976), no. 12, 1283–1290.

[LXF+18] Hangfan Liu, Ruiqin Xiong, Xiaopeng Fan, Debin Zhao, Yongbing Zhang, and Wen Gao, *Cg-cast: Scalable wireless image softcast using compressive gradient*, IEEE Transactions on Circuits and Systems for Video Technology (2018).

[MOA11] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of majorization and its applications*, Springer, New-York, NY, 2011.

[MP02] Udar Mittal and Nam Phamdo, *Hybrid digital-analog (hda) joint source-channel codes for broadcasting and robust communications*, IEEE Transactions on Information Theory **48** (2002), no. 5, 1082–1102.

[Ned03] N. H. Nedev, *Analysis of the impact of impulse noise in digital subscriber line systems*, Ph.D. thesis, The University of Edinburgh, 2003.

[OSS+12] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, *Comparison of the coding efficiency of video coding standards–including high efficiency video coding (HEVC)*, IEEE Trans. Circuits and Systems for Video Technology **22** (2012), no. 12, 1669–1684.

[PCL03] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, *Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization*, IEEE trans. Signal Processing **51** (2003), no. 9, 2381–2401.

[PJ07] D. P. Palomar and Y. Jiang, *MIMO transceiver design via majorization theory*, vol. 3, Foundations and trends in communications and information theory, no. 4, Jun 2007.

[PLC04] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, *Optimum linear joint transmit-receive processing for MIMO channels with qos constraints*, IEEE trans. Signal Processing **52** (2004), no. 5, 1179–1197.

[PWS94] M. O. Polley, S. J. Wee, and W. F. Schreiber, *Hybrid channel coding for multiresolution HDTV terrestrial broadcasting*, Proc. IEEE ICIP, vol. 1, 1994, pp. 243–247.

[Sch95] W. Schreiber, *Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions*, Proc. of the IEEE **83** (1995), no. 6, 958–981.

[SMW07] H. Schwarz, D. Marpe, and T. Wiegand, *Overview of the scalable video coding extension of the H.264/AVC standard*, IEEE Trans. on Circuits Syst. Video Technol. **17** (2007), no. 9, 1103–1120.

[SOHW12] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, *Overview of the high efficiency video coding (HEVC) standard*, IEEE Trans. Circuits Syst. Video Technol. **22** (2012), no. 12, 1649–1668.

[SPA02] Mikael Skoglund, Nam Phamdo, and Fady Alajaji, *Design and performance of vq-based hybrid digital-analog joint source-channel codes*, IEEE Transactions on Information Theory **48** (2002), no. 3, 708–720.

[SPA06] ———, *Hybrid digital–analog source–channel coding for bandwidth compression/expansion*, IEEE Transactions on Information Theory **52** (2006), no. 8, 3757–3763.

[SPX+17] Xiaodan Song, Xiulian Peng, Jizheng Xu, Guangming Shi, and Feng Wu, *Distributed compressive sensing for cloud-based wireless image transmission*, IEEE Transactions on Multimedia **19** (2017), no. 6, 1351–1364.

[SPZ08] Philip Schniter, Lee C Potter, and Justin Ziniel, *Fast bayesian matching pursuit*, Information Theory and Applications Workshop, 2008, IEEE, 2008, pp. 326–333.

[ST03] Andrew Secker and David Taubman, *Lifting-based invertible motion adaptive transform (limat) framework for highly scalable video compression*, IEEE transactions on image processing **12** (2003), no. 12, 1530–1542.

[SXM+14] Z. Song, R. Xiong, S. Ma, X. Fan, and W. Gao, *Layered image/video SoftCast with hybrid digital-analog transmission for robust wireless visual communication*, Proc. IEEE ICME, 2014, pp. 1–6.

[SYLL18] Jian Shen, Lei Yu, Li Li, and Houqiang Li, *Foveation based wireless soft image delivery*, IEEE Transactions on Multimedia **20** (2018), no. 10.

[Tip01] Michael E Tipping, *Sparse bayesian learning and the relevance vector machine*, Journal of machine learning research **1** (2001), no. Jun, 211–244.

[VA99] P. Viswanath and V. Anantharam, *Optimal sequences and sum capacity of synchronous CDMA systems*, IEEE trans. Information Theory **45** (1999), no. 6, 1984–1991.

[WR04] David P Wipf and Bhaskar D Rao, *Sparse bayesian learning for basis selection*, IEEE Transactions on Signal processing **52** (2004), no. 8, 2153–2164.

[WSBL03] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, *Overview of the H.264/AVC video coding standard*, IEEE Trans. on Circuits and Systems for Video Technology **13** (2003), no. 7, 560–576.

[WSO07] M. Wien, H. Schwarz, and T. Oelbaum, *Performance analysis of SVC*, IEEE Trans. Circuits and Systems for Video Technology **17** (2007), no. 9, 1194–1203.

[XLM+14] Ruiqin Xiong, Hangfan Liu, Siwei Ma, Xiaopeng Fan, Feng Wu, and Wen Gao, *G-cast: Gradient based image softcast for perception-friendly wireless visual communication*, Data Compression Conference (DCC), 2014, IEEE, 2014, pp. 133–142.

[XWF+13] R. Xiong, F. Wu, X. Fan, C. Luo, S. Ma, and W. Gao, *Power-distortion optimization for wireless image/video SoftCast by transform coefficients energy modeling with adaptive chunk division*, Proc. IEEE VCIP, 2013, pp. 1–6.

[XWX+16] R. Xiong, F. Wu, J. Xu, X. Fan, C. Luo, and W. Gao, *Analysis of decorrelation transform gain for uncoded wireless image and video communication*, IEEE Trans. Image Processing **25** (2016), no. 4, 1820–1833.

[XZW+17a] R. Xiong, J. Zhang, F. Wu, J. Xu, and W. Gao, *Power distortion optimization for uncoded linear transformed transmission of images and videos*, IEEE Trans. Image Processing **26** (2017), no. 1, 222–236.

[XZW+17b] ———, *Power distortion optimization for uncoded linear transformed transmission of images and videos*, IEEE Trans. Image Processing **26** (2017), no. 1, 222–236.

[YAA+13] L. Yonge, J. Abad, K. Afkhamie, L. Guerrieri, S. Katar, H. Lioe, P. Pagani, R. Riva, D. M. Schneider, and A. Schwager, *An overview of the HomePlug AV2 technology*, Journal of Electrical and Computer Engineering **2013** (2013), 1–20.

[YFS18] Wenbin Yin, Xiaopeng Fan, and Yunhui Shi, *Convolutional neural networks based soft video broadcast*, Pacific Rim Conference on Multimedia, Springer, 2018, pp. 641–650.

[YL07] W. Yu and T. Lan, *Transmitter optimization for the multi-antenna downlink with per-antenna power constraints*, IEEE Trans. Signal Processing **55** (2007), no. 6, 2646–2660.

[YLL14] L. Yu, H. Li, and W. Li, *Wireless scalable video coding using a hybrid digital-analog scheme*, IEEE Trans Circuits and Systems for Video Technology **24** (2014), no. 2, 331–345.

[YLL15] Lei Yu, Houqiang Li, and Weiping Li, *Wireless cooperative video coding using a hybrid digital–analog scheme*, IEEE Transactions on Circuits and Systems for Video Technology **25** (2015), no. 3, 436–450.

[ZAC+16] S. Zheng, M. Antonini, M. Cagnazzo, L. Guerrieri, M. Kieffer, I. Nemoianu, R. Samy, and B. Zhang, *SoftCast with per-carrier power-constrained channels*, Proc. IEEE ICIP, 2016, pp. 2122–2126.

[ZD02] Manfred Zimmermann and Klaus Dostert, *Analysis and modeling of impulsive noise in broad-band powerline communications*, IEEE transactions on Electromagnetic compatibility **44** (2002), no. 1, 249–258.

[ZFXZ13] Ailing Zhang, Xiaopeng Fan, Ruiqin Xiong, and Debin Zhao, *Distributed soft video broadcast with variable block size motion estimation*, Visual Communications and Image Processing (VCIP), 2013, IEEE, 2013, pp. 1–5.

[Zhi06] Sergey V Zhidkov, *Performance analysis and optimization of OFDM receiver with blanking nonlinearity in impulsive noise environment*, IEEE Transactions on Vehicular Technology **55** (2006), no. 1, 234–242.

[Zhi08] ———, *Analysis and comparison of several simple impulsive noise mitigation schemes for ofdm receivers*, IEEE Transactions on Communications **56** (2008), no. 1.

[ZLCW16] Xiao Zhao, Hancheng Lu, Chang Wen Chen, and Jun Wu, *Adaptive hybrid digital–analog video transmission in wireless fading channel*, IEEE Transactions on Circuits and Systems for Video Technology **26** (2016), no. 6, 1117–1130.

[ZLMW17] Z. Zhang, D. Liu, X. Ma, and X. Wang, *Ecast: An enhanced video transmission design for wireless multicast systems over fading channels*, IEEE Systems Journal **11** (2017), no. 4, 2566–2577.

[ZWL+18] Jing Zhang, Anhong Wang, Jie Liang, Haidong Wanga, Suyue Li, and Xiong Zhangb, *Distortion estimation-based adaptive power allocation for hybrid digital-analog video transmission*, IEEE Transactions on Circuits and Systems for Video Technology (to appear) (2018).

[ZWW+15] F. Zhang, A. Wang, H. Wang, S. Li, and X. Ma, *Channel-aware video SoftCast scheme*, Proc. IEEE ChinaSIP (Chengdu, China), 2015, pp. 578–581.

[ZZ95] H. Zha and Z. Zhang, *A note on constructing a symmetric matrix with specified diagonal entries and eigenvalues*, BIT Numerical Mathematics **35** (1995), no. 3, 448–452.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Prise en compte des contraintes de canal dans les schémas de codage vidéo conjoint du source-canal

Mots clés : codage conjoint du source-canal, transmission vidéo, codage vidéo linéaire, allocation de la puissance, réduction de l'impact du bruit impulsif

Résumé : Les schémas de Codage Vidéo Linéaire (CVL) inspirés de SoftCast ont émergé dans la dernière décennie comme une alternative aux schémas de codage vidéo classiques. Ces schémas de codage source-canal conjoint exploitent des résultats théoriques montrant qu'une transmission (quasi-)analogique est plus performante dans des situations de multicast que des schémas numériques lorsque les rapports signal-à-bruit des canaux (C-SNR) diffèrent d'un récepteur à l'autre. Dans ce contexte, les schémas de CVL permettent d'obtenir une qualité de vidéo décodée proportionnelle au C-SNR du récepteur.
Une première contribution de cette thèse concerne l'optimisation de la matrice de précodage de canal pour une transmission de type OFDM de flux générés par un CVL lorsque les contraintes de puissance diffèrent d'un sous-canal à l'autre. Ce type de contrainte apparait en sur des canaux DSL, ou dans des dispositifs de transmission sur courant porteur en ligne (CPL). Cette thèse propose une solution optimale à ce problème de type multi-level water filling et nécessitant la solution d'un problème de type Structured Hermitian Inverse Eigenvalue. Trois algorithmes sous-optimaux de complexité réduite sont également proposés. Des nombreux résultats de simulation montrent que les algorithmes sous-optimaux ont des performances très proches de l'optimum et réduisent significativement le temps de codage. Le calcul de la matrice de précodage dans une situation de multicast est également abordé.
Une seconde contribution principale consiste en la réduction de l'impact du bruit impulsif dans les CVL. Le problème de correction du bruit impulsif est formulé comme un problème d'estimation d'un vecteur creux. Un algorithme de type Fast Bayesian Matching Pursuit (FBMP) est adapté au contexte CVL. Cette approche nécessite de réserver des sous-canaux pour la correction du bruit impulsif, entrainant une diminution de la qualité vidéo en l'absence de bruit impulsif. Un modèle phénoménologique (MP) est proposé pour décrire l'erreur résiduelle après correction du bruit impulsif. Ce modèle permet de d'optimiser le nombre de sous-canaux à réserver en fonction des caractéristiques du bruit impulsif. Les résultats de simulation montrent que le schéma proposé améliore considérablement les performances lorsque le flux CVL est transmis sur un canal sujet à du bruit impulsif.

Title : Accounting for Channel Constraints in Joint Source-Channel Video Coding Schemes

Keywords : joint source-channel coding, video transmission, linear video coding, power allocation, impulse noise mitigation

Abstract : SoftCast based Linear Video Coding (LVC) schemes have been emerged in the last decade as a quasi analog joint-source-channel alternative to classical video coding schemes. Theoretical analyses have shown that analog coding is better than digital coding in a multicast scenario when the channel signal-to-noise ratios (C-SNR) differ among receivers. LVC schemes provide in such context a decoded video quality at different receivers proportional to their C-SNR.
This thesis considers first the channel precoding and decoding matrix design problem for LVC schemes under a per-subchannel power constraint. Such constraint is found, e.g., on Power Line Telecommunication (PLT) channels and is similar to per-antenna power constraints in multi-antenna transmission system. An optimal design approach is proposed, involving a multi-level water filling algorithm and the solution of a structured Hermitian Inverse Eigenvalue problem. Three lower-complexity alternative suboptimal algorithms are also proposed. Extensive experiments show that the suboptimal algorithms perform closely to the optimal one and can reduce significantly the complexity. The precoding matrix design in multicast situations also has been considered.
A second main contribution consists in an impulse noise mitigation approach for LVC schemes. Impulse noise identification and correction can be formulated as a sparse vector recovery problem. A Fast Bayesian Matching Pursuit (FBMP) algorithm is adapted to LVC schemes. Subchannels provisioning for impulse noise mitigation is necessary, leading to a nominal video quality decrease in absence of impulse noise. A phenomenological model (PM) is proposed to describe the impulse noise correction residual. Using the PM model, an algorithm to evaluate the optimal number of subchannels to provision is proposed. Simulation results show that the proposed algorithms significantly improve the video quality when transmitted over channels prone to impulse noise.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France