



HAL
open science

Zero-error network information theory: graphs, coding for computing and source-channel duality

Nicolas Charpenay

► **To cite this version:**

Nicolas Charpenay. Zero-error network information theory: graphs, coding for computing and source-channel duality. Information Theory [math.IT]. Université de Rennes, 2023. English. NNT: . tel-04384481

HAL Id: tel-04384481

<https://theses.hal.science/tel-04384481>

Submitted on 10 Jan 2024

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.

Public Domain

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE N° 601

*Mathématiques, Télécommunications, Informatique, Signal, Systèmes,
Électronique*

Spécialité : *Signal, Image, Vision*

Par

Nicolas CHARPENAY

**Théorie de l'information zéro-erreur dans les réseaux :
graphes, codage pour le calcul de fonctions et dualité source-
canal**

Thèse présentée et soutenue à Rennes, le mardi 28 novembre 2023

Unité de recherche : IRISA UMR 6074 et Centre Inria de l'Université de Rennes

Rapporteurs avant soutenance :

Mikael SKOGLUND Professeur à KTH
Michèle WIGGER Professeure à Telecom Paris

Composition du Jury :

Président :	Jean-Marie GORCE	Professeur des Universités à INSA Lyon (<i>à préciser après la soutenance</i>)
Rapporteurs :	Mikael SKOGLUND	Professeur à KTH
	Michèle WIGGER	Professeure à Telecom Paris
Examinatrice :	Derya MALAK	Maître de Conférences à EURECOM
Dir. de thèse :	Aline ROUMY	Directrice de Recherche Inria
Co-dir. de thèse :	Maël LE TREUST	Chargé de Recherche CNRS
Invité :	Tristan TOMALA	Professeur à HEC Paris

TABLE OF CONTENTS

1	Résumé	7
1.1	Problèmes zéro-erreur basés sur la correction	8
1.2	Problèmes zéro-erreur basés sur les graphes	10
1.3	Problèmes zéro-erreur pour le calcul de fonction	11
1.4	Organisation du manuscrit	11
	Glossary	13
2	Introduction	15
2.1	Correction-based zero-error problems	16
2.2	Graph-based zero-error problems	17
2.3	Coding for computing zero-error problems	18
2.4	Organization of the manuscript	19
3	State of the art	21
3.1	Source coding	21
3.2	Channel coding	23
3.2.1	Channel coding in the vanishing error regime	25
3.2.2	Zero-error channel coding	25
3.2.3	Related works	28
3.3	Slepian-Wolf problem	28
3.3.1	Zero-error Slepian-Wolf problem	30
4	Contributions for correction-based zero-error problems	33
4.1	Correction-based zero-error problems	33
4.1.1	Zero-error source coding when side information may be present at the decoder	33
4.1.2	Gray-Wyner problem	36
4.1.3	Zero-error source coding for a Gray-Wyner network when side in- formation may be present at the decoder	38

TABLE OF CONTENTS

4.2	Main results for the zero-error source coding problem when side information may be present at the decoder	41
4.2.1	Definitions for the achievability proof	42
4.2.2	Coding scheme	43
4.2.3	Zero-error property and rate analysis	45
4.3	Main results for the zero-error source coding problem for a Gray-Wyner network when side information may be present at the decoder	45
4.3.1	Comparison with previous results	47
5	Contributions for graph-based zero-error problems	49
5.1	Graph-based zero-error problems	50
5.1.1	Zero-error channel coding with typical input constraint	50
5.1.2	Linearization in the zero-error Slepian-Wolf problem	53
5.1.3	Linearization in zero-error channel coding	56
5.2	Main results on the linearization of \bar{H} and $C(\cdot, P_V)$	60
5.3	Main results on capacity-achieving distributions	62
5.4	Linking the linearizations of C_0 and $C(\cdot, P_V)$	64
5.5	Main examples and counterexamples of linearization	65
5.5.1	Perfect graphs	65
5.5.2	The Schläfli graph	68
6	Contributions for coding for computing zero-error problems	71
6.1	Coding for computing problems	71
6.1.1	Coding for computing in the vanishing error regime	71
6.1.2	Zero-error coding for computing with side information at the encoder	73
6.2	General case	75
6.3	Pairwise shared side information	76
6.4	Example	78
	Conclusion	81
A	Proofs for correction-based zero-error results	83
A.1	Proof of Lemma 4.2.5	83
A.1.1	Preliminaries	83
A.1.2	Probability of decoding ambiguity	83
A.1.3	Rate on the common channel	85

A.1.4	Rate on the secondary channel	86
A.2	Proof of Theorem 4.3.1	88
A.2.1	Outline	88
A.2.2	Encoding algorithm	89
A.2.3	Decoding algorithm and zero-error property	90
A.2.4	Probability of a decoding ambiguity	91
A.2.5	Rate analysis	91
A.3	Proof of Theorem 4.3.2	92
A.4	Proof of Lemma 4.3.3	93
B	Proofs for graph-based zero-error results	95
B.1	Proof dependencies	95
B.2	Main proofs for the linearization of \overline{H} and $C(\cdot, P_V)$	95
B.2.1	Proof of Lemma 5.2.1	95
B.2.2	Proof of Theorem 5.2.4	96
B.2.3	Proof of Lemma B.2.1	97
B.2.4	Proof of Proposition 5.2.5	100
B.3	Main proofs for the capacity-achieving distributions	101
B.3.1	Proof of Lemma 5.1.4	101
B.3.2	Proof of Theorem 5.1.5	102
B.3.3	Proof of Proposition 5.3.2	105
B.3.4	Proof of Theorem 5.3.3	106
B.3.5	Proof of Lemma 5.3.7	110
B.4	Main proofs for the link between linearizations of $C(\cdot, P_V)$ and C_0	111
B.4.1	Proof of Theorem 5.4.1	111
B.4.2	Proof of Lemma B.4.1	112
B.4.3	Proof of Lemma B.4.2	113
B.4.4	Proof of Theorem 5.4.2	114
B.4.5	Proof of Lemma B.4.3	116
B.4.6	Proof of Lemma B.4.4	117
B.5	Main proofs for the examples and counterexamples	118
B.5.1	Proof of Theorem 5.5.8	118
B.6	Secondary proofs	119
B.6.1	Proof of Lemma B.2.2	119

TABLE OF CONTENTS

B.6.2	Proof of Lemma B.2.3	121
B.6.3	Proof of Lemma B.2.4	122
B.6.4	Proof of Lemma B.2.5	122
B.6.5	Proof of Lemma B.2.7	123
B.6.6	Proof of Lemma B.2.8	124
B.6.7	Proof of Lemma B.3.1	125
B.6.8	Proof of Lemma B.4.5	126
B.6.9	Proof of Lemma B.5.2	127
B.6.10	Proof of Lemma B.6.1	127
B.6.11	Proof of Lemma B.6.2	128
B.6.12	Proof of Lemma B.6.4	129
B.6.13	Proof of Lemma B.6.5	129
C	Proofs for coding for computing zero-error results	131
C.1	Proof of Theorem 6.3.2	131
C.2	Proof of Theorem 6.3.6	133
C.3	Proof of Lemma C.2.2	134
C.4	Proof of Lemma C.2.3	135
	Bibliography	137

RÉSUMÉ

Le codage de source zéro-erreur regroupe une grande variété de problèmes en théorie de l'information, où une source doit être compressée par un encodeur avec un débit à minimiser, puis doit être retrouvée par le décodeur avec probabilité d'erreur zéro. Dans tous les problèmes étudiés ici, la distribution de source est connue et le décodeur a accès à une information adjacente, i.e. une variable aléatoire corrélée à la source qui aide au décodage et à minimiser le débit. L'étude de tels problèmes trouve son utilité dans des applications demandant une garantie forte sur la qualité de reconstruction de la source, en particulier la télédétection et le calcul distribué. Dans un problème de codage de source zéro-erreur, trouver une formule pour le débit optimal permet d'établir une borne théorique sur les performances de tous les schémas de codage possibles. De plus, si cette formule est à une seule lettre (i.e. ne fait pas apparaître de limite), elle peut être facilement calculée. On a aussi l'existence d'un schéma de codage qui permet d'atteindre asymptotiquement ce débit optimal. Pour ces raisons, notre premier objectif dans ce manuscrit sera de trouver des expressions à une seule lettre pour les débits optimaux des problèmes de codage de source zéro-erreur.

La contrainte zéro-erreur diffère de la contrainte d'erreur asymptotiquement nulle, cette dernière ne demandant qu'une limite nulle pour la probabilité d'erreur au décodage quand la longueur du bloc codant tend vers l'infini. Dans les cadres où le décodeur n'a pas d'information adjacente, on peut facilement montrer que les débits optimaux sous les contraintes zéro-erreur et erreur asymptotiquement nulle coïncident. Toutefois, lorsqu'on impose la contrainte zéro-erreur quand le décodeur a une information adjacente – ce sera le cas dans les problèmes étudiés ici, la nature de certains problèmes passe de statistique à combinatoire (e.g. le problème de Slepian-Wolf zéro-erreur). Il s'ensuit que les outils adéquats pour de tels problèmes sont à la fois tirés de la théorie de l'information et de la combinatoire ; et peuvent aller d'un dispositif de correction qui donne la propriété zéro-erreur à des codes existants, à des codes zéro-erreur entièrement bâtis avec des contraintes tirées de la théorie des graphes.

La nature combinatoire des problèmes zéro-erreur est liée à l'incertitude de l'encodeur sur l'information adjacente du décodeur. Dans les problèmes à erreur asymptotiquement nulle, on peut supposer que la suite des réalisations de l'information adjacente du décodeur sera typique par rapport à la distribution de source, avec grande probabilité ; même si l'encodeur n'a pas accès à ces réalisations. En revanche, la contrainte zéro-erreur rend ce manque d'information plus difficile à gérer ; en particulier, il faut que les codes utilisés permettent une reconstruction parfaite de la source pour *n'importe quelle* réalisation de l'information adjacente du décodeur.

Par conséquent, nous classifions les problèmes zéro-erreur étudiés ici en trois catégories ; selon l'information que possède l'encodeur sur l'information adjacente du décodeur, et selon ce que le décodeur lui-même doit retrouver :

- Les problèmes zéro-erreur *basés sur la correction*, dans lesquels l'encodeur a accès à toutes les informations adjacentes du problème, et peut simuler chaque décodage ;
- Les problèmes zéro-erreur *basés sur les graphes*, dans lesquels l'encodeur est “moins informé” sur l'information adjacente du décodeur ;
- Les problèmes zéro-erreur *pour le calcul de fonction*, dans lesquels le décodeur doit calculer une fonction de la source et de son information adjacente ; au lieu de simplement reconstruire la source.

1.1 Problèmes zéro-erreur basés sur la correction

Le problèmes zéro-erreur *basés sur la correction* peuvent être abordés avec l'aide d'un dispositif de correction qui garantit la propriété zéro-erreur. Il consiste en l'utilisation d'un livre-code adapté à la contrainte d'erreur asymptotiquement nulle, la simulation de chaque décodage, et l'envoi de toute les suite de source s'il y a erreur dans l'un des décodages. Ce dispositif de correction a un impact négligeable sur le débit en raison de la faible probabilité d'erreur, et permet la conversion des codes à erreur asymptotiquement nulle en codes zéro-erreur. Pour cette raison, les outils que nous utilisons avec le dispositif de correction sont la typicalité, et les codes aléatoires et de parité.

Dans le premier problème de cette classe, présenté en Figure 4.1, l'information adjacente peut être présente au décodeur. Dans ce cadre, l'encodeur dispose d'une information adjacente, mais ignore si le décodeur y a accès aussi ; l'encodage est divisé en deux étapes. À la première étape, l'encodeur envoie de l'information sans savoir si l'information adjacente est présente au décodeur. Le message envoyé à la première étape doit être exploitable

indépendamment de l'éventuelle présence de l'information adjacente. De plus, un décodage zéro-erreur doit être réalisé si le décodeur a accès à l'information adjacente. À la deuxième étape, un supplément d'information est envoyé par l'encodeur dans le cas où l'information adjacente n'est pas présente au décodeur, et ce dernier doit ensuite retrouver la source avec zéro-erreur. Ces deux étapes sont modélisées par deux décodeurs différents et deux canaux parfaits. Nous voulons répondre à la question suivante : quelles sont les paires de débits réalisables pour les deux étapes d'encodage, qui permettent un décodage zéro-erreur ? Notre première contribution consiste à déterminer la région réalisable, à l'aide d'une partition en cosets des suites de source, couplée à un dispositif de correction ; prouvant de ce fait que cette région réalisable en zéro-erreur coïncide avec la région réalisable en erreur asymptotiquement nulle.

Le deuxième problème basé sur la correction est illustré en Figure 4.3 et est une généralisation du premier. Dans ce problème où l'encodeur connaît deux sources différentes et leurs informations adjacentes respectives, deux incertitudes se présentent : la première est que l'encodeur ignore quelle source sera demandée par le décodeur, et la deuxième est que l'information adjacente de la source demandée peut être présente au décodeur. L'encodage est alors divisé en trois étapes. À la première étape, l'encodeur envoie de l'information sans savoir quelle source sera demandée ; nous cherchons à envoyer l'information commune aux deux différentes sources, d'où l'appellation "réseau de Gray-Wyner" pour ce problème. À la deuxième étape, l'encodeur est informé de la source demandée par le décodeur, mais ignore toujours si ce dernier dispose d'une information adjacente ; si tel est le cas, alors un décodage zéro-erreur doit être réalisé. La troisième étape est un supplément d'information envoyé par l'encodeur dans le cas où le décodeur ne dispose pas de l'information adjacente ; le décodeur doit ensuite retrouver la source avec zéro-erreur. Pour ce problème, notre deuxième contribution consiste en une borne interne et une borne externe sur la région réalisable. Pour la borne interne, nous utilisons le schéma de codage suivant : des suites auxiliaires sont utilisées pour capturer l'information commune entre les suites de réalisations des deux sources, ensuite nous effectuons un étiquetage aléatoire de ces suites, et les ensembles de suites de source sont partitionnées en cosets. Les algorithmes de décodages sont basés sur les V -enveloppes des suites d'information adjacente, et la propriété zéro-erreur est garantie par un dispositif de correction.

1.2 Problèmes zéro-erreur basés sur les graphes

Dans les problèmes zéro-erreur *basés sur les graphes*, les contraintes zéro-erreur sont représentées par un graphe caractéristique, et déterminer le débit optimal dans de tels problèmes revient à résoudre des questions difficiles de théorie des graphes. Le problème de Slepian-Wolf zéro-erreur illustré en Figure 3.4 est l’un d’eux, son débit optimal est donné par l’entropie complémentaire \bar{H} de son graphe caractéristique, et trouver une expression à une seule lettre pour \bar{H} est une question ouverte. Toutefois, afin de mieux comprendre la nature du problème Slepian-Wolf zéro-erreur, nous donnons de nouveaux résultats structurels qui lient ce problème avec d’autres, en particulier la capacité zéro-erreur d’un canal C_0 (voir la Figure 3.2).

Notre étude principale concerne le problème de “linéarisation”. Dans le cadre du problème Slepian-Wolf zéro-erreur, si l’encodeur est en train de réaliser plusieurs tâches de compression indépendantes avec leurs informations adjacentes respectives, alors le débit optimal est donné par $\bar{H}(\wedge \cdot)$, i.e. \bar{H} d’un graphe avec une structure de produit ET. Comme illustré par Tuncel et al. dans [71], “séparer” les tâches indépendantes donne un schéma d’atteignabilité, d’où $\bar{H}(\wedge \cdot) \leq \sum \bar{H}(\cdot)$. Un autre cas d’intérêt est celui où l’encodeur dispose d’une information partielle $g(Y)$ sur l’information adjacente du décodeur. Le débit optimal est donné par $\bar{H}(\sqcup \cdot)$, i.e. \bar{H} d’un graphe avec une structure d’union disjointe. “Séparer” les réalisations de la source selon les valeurs de l’information adjacente de l’encodeur donne aussi un schéma d’atteignabilité, d’où $\bar{H}(\sqcup^{P_{g(Y)}} \cdot) \leq \sum_z P_{g(Y)}(z) \bar{H}(\cdot)$. Lorsqu’on a égalité dans l’un ou l’autre de ces cas, on dit qu’on a “linéarisation” de \bar{H} .

Notre contribution consiste à prouver les équivalences des linéarisations de $C_0(\wedge \cdot)$, $C_0(\sqcup \cdot)$, $C(\wedge \cdot, P_V)$, $C(\sqcup \cdot, P_V)$, $\bar{H}(\wedge \cdot)$, et $\bar{H}(\sqcup \cdot)$; où $C(\cdot, P_V)$ est le débit optimal du problème de codage canal zéro-erreur où l’encodeur doit utiliser des mots de codes typiques par rapport à la distribution P_V . Par conséquent, le schéma de codage “séparé” est optimal dans le problème de codage canal zéro-erreur, si et seulement si il est optimal dans le problème de Slepian-Wolf zéro-erreur avec le même graphe caractéristique. De plus, dans chacun de ces problèmes, les optimalités des schémas de codage “séparés” respectifs pour le produit \wedge et pour l’union disjointe \sqcup sont équivalentes. Pour prouver cela, nous définissons les distributions atteignant la capacité d’un canal en régime zéro-erreur, et nous déterminons plusieurs résultats sur celles-ci ; cela nous permet de lier C_0 avec $C(\cdot, P_V)$ et \bar{H} . De plus, nous donnons des exemples et contre-exemples de linéarisation pour toutes ces équivalences. Tout d’abord, cela nous donne une formule à une seule lettre pour \bar{H}

dans des cas où elle manquait ; en particulier pour les produits de graphes parfaits, qui ne sont pas nécessairement parfaits. Enfin, les contre-exemples développés illustrent que les schémas de codage “séparés” ne sont pas toujours optimaux: dans le problème de Slepian-Wolf zéro-erreur, le débit peut strictement décroître quand on compresse ensemble des sources indépendantes.

1.3 Problèmes zéro-erreur pour le calcul de fonction

Les problèmes zéro-erreur *pour le calcul de fonction* sont une généralisation des problèmes basés sur les graphes, et font aussi usage des graphes caractéristiques. Toutefois, la fonction à retrouver impacte aussi les outils adéquats de la théorie des graphes à utiliser. Dans le problème illustré en Figure 6.2, le décodeur doit retrouver une fonction f de la source et de son information adjacente. L’encodeur ne connaît pas les réalisations de cette dernière, mais en observe une version déterministiquement dégradée, représentée par la fonction g .

Notre première contribution est de donner une expression asymptotique pour le débit optimal de ce problème. Notre deuxième contribution est de formuler une hypothèse que nous appelons “information adjacente partagée deux-à-deux” qui nous permet, lorsqu’elle est satisfaite, d’obtenir une expression à une seule lettre pour le débit optimal. Cette hypothèse est satisfaite si chaque paire de symboles de source “partage” au moins un symbole d’information adjacente pour tout résultat de g . Cette condition a des interprétations en termes de théorie des graphes, car les formules à une seule lettre que nous trouvons émanent de la structure particulière du graphe caractéristique : ce dernier est une union disjointe de produits OU. De plus, ce résultat est d’intérêt pratique car il couvre toutes les instances où la distribution de source est à support plein, sans aucune hypothèse sur f, g . Enfin, nous donnons une interprétation de cette condition, en termes de plus mauvais débit optimal dans un problème auxiliaire de Slepian-Wolf zéro-erreur.

1.4 Organisation du manuscrit

Ce manuscrit est organisé comme suit. Dans le Chapitre 3, nous détaillons la présentation de chacun des problèmes étudiés et l’état de l’art pour chacun d’eux, ainsi que les définitions utilisées dans ce manuscrit. Dans le Chapitre 4, le Chapitre 5, et le Chapitre 6, nous présentons respectivement nos résultats pour les problèmes zéro-erreur *basés sur la*

correction, basés sur les graphes, et pour le calcul de fonction. Les preuves de ces résultats se trouvent respectivement en Appendice **A**, Appendice **B**, et Appendice **C**.

GLOSSARY

For reader's convenience we give here a table of the notations and concepts used throughout this manuscript.

\mathbb{N}^*, \mathbb{R}	set of positive integers, real numbers
$\mathcal{M}_{n,k}(\mathcal{A})$	set of matrices of size $n \times k$ with coefficients in \mathcal{A}
$ \mathcal{S} , \mathcal{S}^c$	cardinality, complement of the set \mathcal{S}
Im	image set of a mapping
Ker	kernel of a matrix
$\{0, 1\}^*$	set of binary words
$\ell(\cdot)$	length of a word
x^n	sequence (x_1, \dots, x_n)
$\Delta(\mathcal{X})$	probability distributions over \mathcal{X}
P_X	probability distribution of the random variable X
supp P_X	support of P_X
Unif(\cdot)	uniform distribution
$P_{X,Y}$	joint distribution of (X, Y)
$P_{X Y}$	conditional distribution of X knowing Y
$P_X \otimes P_Y$	product distribution
$\cdot \rightarrow \cdot \rightarrow \cdot$	Markov chain
T_{x^n}	type of x^n $\left(\frac{1}{n} \left \{t \leq n \mid x_t = x'\} \right \right)_{x' \in \mathcal{X}}$
$\Delta_n(\mathcal{X})$	set of types of sequences from \mathcal{X}^n
$\mathcal{T}_\epsilon^n(P_X)$	typical set $\{x^n \in \mathcal{X}^n \mid \ T_{x^n} - P_X\ _\infty \leq \epsilon\}$
$\mathbb{P}(\cdot)$	probability of an event
$\mathbb{E}[\cdot]$	expected value of a random variable
$H(\cdot), H(\cdot \mid \cdot)$	entropy, conditional entropy
$I(\cdot; \cdot)$	mutual information
$I(\cdot; \cdot \mid \cdot)$	conditional mutual information

C_n, K_n, N_n	cycle, complete, empty graph with n vertices	
ϕ_e, ϕ_d	encoding, decoding function	
\wedge	AND product	Definition 3.2.5
α	independence number	Definition 3.2.6
C_0	zero-error capacity	Definition 3.2.7
$(\mathcal{V}, \mathcal{E}, P_V)$	probabilistic graph	Definition 3.3.4
χ	chromatic number	Definition 3.3.5
\overline{H}	complementary graph entropy	Definition 3.3.6
H_χ	chromatic entropy	Definition 3.3.8
$\mathcal{T}_0^n(Q_X)$	type class for the type Q_X	Definition 4.2.2
$\mathcal{T}_V(x^n)$	V -shell of the sequence x^n	Definition 4.2.2
$G[S]$	subgraph of G induced by \mathcal{S}	Definition 5.1.1
$C(\cdot, P_V)$	zero-error capacity relative to P_V	Definition 5.1.2
\sqcup	disjoint union	Definition 5.1.9
\overline{G}	graph complement of G	Definition 5.5.1
ω	clique number	Definition 5.5.1
H_κ	Körner graph entropy	Definition 5.5.3
\vee	OR product	Definition 6.3.3
\simeq	isomorphic (probabilistic) graphs	Definition B.2.6

A set of words $\mathcal{W} \subseteq \{0, 1\}^*$ is *prefix-free* if for all $w, w' \in \mathcal{W}$, w is not a prefix of w' and *vice-versa*.

INTRODUCTION

Zero-error source coding encompasses a variety of problems from information theory, where a source is compressed by an encoder with rate to be minimized, then has to be retrieved at the decoder with probability of error zero. In all the problems studied here, the source distribution is known and the decoder has access to a side information, i.e. a random variable correlated to the source that helps minimizing the rate and decoding. Studying such settings is useful in applications that require a strong guarantee on the quality of reconstruction of the source, such as remote sensing networks and distributed computing. In a zero-error source coding problem, finding a formula for the optimal rate means finding the strongest theoretic limit on all compression schemes. Furthermore, if the formula is single-letter, then it becomes easily computable. One also has the existence of an optimal zero-error compression scheme that asymptotically achieves this optimal rate. Therefore, finding single-letter expressions for optimal rates in zero-error source coding problems is our main objective in this manuscript.

The zero-error constraint differs from the vanishing-error constraint, the latter only requires the probability of error to go to zero when the block length of the code goes to infinity. In the settings where the decoder does not have a side information, it can be easily shown that the optimal rates in the vanishing error and zero-error regimes coincide. However, when imposing the zero-error constraint when the decoder has a side information, as in the settings studied here, many problems change from a statistical nature to a combinatorial one (e.g. the Slepian-Wolf setting). Therefore, the adequate tools for zero-error problems are drawn from both information theory and combinatorics; and vary from a correcting device that strengthens vanishing-error codes into zero-error ones, to zero-error codes entirely built with graph-theoretic constraints.

The combinatorial nature of zero-error problems is tied to the encoder's lack of knowledge of the decoder's side information. In vanishing-error problems one can assume that the decoder's side information sequence will be typical w.r.t. the source distribution with high probability; even if the encoder does not have access to it. However, the zero-error

constraint makes the encoder’s uncertainty on decoder’s side information harder to be dealt with; in particular it is required to build codes that enable for a perfect reconstruction of the source for *any* realization of the decoder’s side information.

For this reason, we classify the zero-error problems studied here into three categories, depending on what the encoder knows about the decoder’s side information and what the decoder has to retrieve:

- The *correction-based* zero-error problems, in which the encoder has access to all side information and can simulate every decoding.
- The *graph-based* zero-error problems, in which the encoder is “less informed” about the realizations of decoder’s side information.
- The *coding for computing* zero-error problems, in which the decoder has to compute a function of the source and its side information; instead of only recovering the source.

2.1 Correction-based zero-error problems

The *correction-based* problems can be tackled with a correcting device that guarantees the zero-error property. It consists in using a vanishing-error codebook, simulating each decoding, and sending all the source sequences if any decoding fails. This correcting device has a negligible impact on the rate because of the low probability of error, and allows for the conversion of vanishing error coding schemes into zero-error ones. For this reason, the tools we use are typicality, random coding and parity check codes, with correcting device.

In the first problem from this class, which is depicted in Figure 4.1, side information may be present at the decoder. In this setting, the encoder has access to a side information but does not know whether the decoder has access to it; and the encoding is divided into two steps. In the first step, the encoder sends information without knowing whether side information is available at the decoder. The message sent in the first step must be exploitable independently from the presence of the side information; furthermore, a zero-error decoding must be done if the decoder has access to the side information. In the second step, an information complement is sent by the encoder in the case where the decoder has no side information, and the decoder must be able to retrieve the source with zero-error. The two steps are modeled by two decoders and two noiseless channels. We aim at answering the following question: what are the feasible pairs of rates for the two encoding steps, that allow for a zero-error decoding? Our first contribution consists

in determining the feasible rate region, thanks to a coset partition of source sequences coupled with a correcting device; therefore showing that it coincides with the feasible rate region in the vanishing error regime.

The second correction-based zero-error problem is depicted in Figure 4.3 and is a generalization of the first one, in which there are two uncertainties: firstly, the encoder has access to two sources and their respective side information but does not know which one will be requested by the decoder; and secondly, the side information may be present at the decoder. The encoding is divided into three steps. In the first step, the encoder sends information without knowing which source is requested; we aim at sending the common information between the two sources, hence the “Gray-Wyner network” name for this problem. In the second step, the encoder knows which source is requested but does not know if the decoder has access to the side information; if the latter is present at the decoder, then a zero-error decoding must be done. The third step is an information complement sent by the encoder in the case where the decoder does not have the side information; and the decoder must be able to retrieve the source with zero-error. In this setting, our second contribution consists in an inner and an outer bound on the feasible rate region. For the inner bound we use the following coding scheme: auxiliary sequences are used to capture the common information between the two sources sequences, then we do a random binning on them; and the sets of source sequences are partitioned into cosets. Decoding algorithms are based on V -shells of side information sequences, and the zero-error property is ensured by a correcting device.

2.2 Graph-based zero-error problems

In *graph-based* problems, zero-error constraints are captured by a characteristic graph, and determining the optimal rate requires to solve hard graph-theoretic questions. The zero-error Slepian-Wolf problem depicted in Figure 3.4 is one of them, its optimal rate is given by the complementary graph entropy \bar{H} of the characteristic graph, and finding a single-letter formula for \bar{H} is an open question. However, in order to understand better the nature of the zero-error Slepian-Wolf problem, we provide new structural results that link it with other problems, such as the zero-error capacity of a channel C_0 (see Figure 3.2).

Our main study concerns the “linearization” problem. In the zero-error Slepian-Wolf setting, if the encoder is doing several independent compression tasks with their respective

independent side information, then the optimal rate is given by $\overline{H}(\wedge \cdot)$, i.e. \overline{H} of a graph with an AND product structure. As shown by Tuncel et al. in [71], “separating” the independent tasks yields an achievability scheme, hence $\overline{H}(\wedge \cdot) \leq \sum \overline{H}(\cdot)$. Another interesting case of zero-error Slepian-Wolf problem is when the encoder has a partial information $g(Y)$ on the decoder’s side information. The optimal rate is given by $\overline{H}(\sqcup \cdot)$, i.e. \overline{H} of a graph with a disjoint union structure. “Separating” the source realizations w.r.t. the encoder’s side information also yields an achievability scheme, hence $\overline{H}(\sqcup^{P_{g(Y)}} \cdot) \leq \sum_z P_{g(Y)}(z) \overline{H}(\cdot)$. When equality holds in either case we say that “linearization” of \overline{H} holds.

Our contribution is to prove the equivalences of linearizations between $C_0(\wedge \cdot)$, $C_0(\sqcup \cdot)$, $C(\wedge \cdot, P_V)$, $C(\sqcup \cdot, P_V)$, $\overline{H}(\wedge \cdot)$, and $\overline{H}(\sqcup \cdot)$; where $C(\cdot, P_V)$ is the optimal rate in the zero-error channel coding problem where the encoder has to use codewords that are typical w.r.t. P_V . Therefore, the “separated” coding scheme is optimal in the zero-error channel coding problem, if and only if it is optimal in the zero-error Slepian-Wolf setting with the same characteristic graph. Furthermore, in each of these settings, the optimalities of the respective “separated” coding strategies for the product \wedge and the disjoint union \sqcup are equivalent. In order to prove that, we define the zero-error capacity achieving distributions of a channel, and derive several results on them; which enables us to link C_0 with $C(\cdot, P_V)$ and \overline{H} . Furthermore, we give examples and counterexamples of linearization for all these equivalences. Firstly, this yields a single-letter formula for \overline{H} in cases where it was unknown: product of perfect graphs, which are not necessarily perfect. Secondly, the counterexamples developed illustrate that the “separated” coding schemes are not always optimal: in the zero-error Slepian-Wolf setting, the rate may strictly decrease when compressing together independent sources.

2.3 Coding for computing zero-error problems

The *coding for computing* problems are a generalization of graph-based problems, and also make use of characteristic graphs. However, the function to be retrieved also impacts the adequate graph-theoretic tools to be used. In the setting depicted in Figure 6.2, the decoder has to retrieve a function f of its side information and of the source; the encoder does not know the realizations of the decoder’s side information but observes a deterministically degraded version of them, represented by the function g .

Our first contribution is to give an asymptotic formula for the optimal rate in this setting. Our second contribution is to formulate an hypothesis that we call “pairwise

shared side information” that allows us to derive a single-letter characterization of the optimal rate. This hypothesis is satisfied if every pair of source symbols “share” at least one side information symbol for all output of g . It has graph-theoretic interpretations, as the single-letter formula stems from the particular structure of the characteristic graph: a disjoint union of OR products. Moreover, this result is of practical interest as it covers the cases where the source distribution is full-support, without any assumption on f, g . Finally, we give an interpretation of this condition, in terms of the worst optimal rate in an auxiliary zero-error Slepian-Wolf problem.

2.4 Organization of the manuscript

This manuscript is organized as follows. In Chapter 3 we provide a detailed presentation of all the problems studied and the definitions used in this manuscript, and we describe the state of the art for each one of them. In Chapter 4, Chapter 5, and Chapter 6, we present our results for the *correction-based*, *graph-based*, and *coding for computing* zero-error problems, respectively. The proofs of these results can be found in Appendix A, Appendix B, and Appendix C, respectively.

STATE OF THE ART

In this Chapter we give a detailed presentation of the different problems studied in this manuscript, in both zero-error and vanishing error regime. As illustrated in the following, the zero-error and the vanishing error regimes may lead to different optimal rates in several settings; in particular in the channel coding problem, and in the Slepian-Wolf problem.

3.1 Source coding

The source coding problem is one of the fundamental settings that appear in data compression; it is introduced by Shannon in [59].

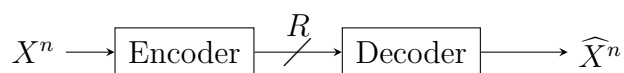


Figure 3.1 – The source coding problem.

Definition 3.1.1 (Source coding problem) *The source coding problem of Figure 3.1 is described by:*

- A finite set \mathcal{X} , and a source distribution $P_X \in \Delta(\mathcal{X})$;
- For all $n \in \mathbb{N}^*$, X^n is the random sequence of n copies of X , drawn in an i.i.d. fashion using P_X .
- An encoder that knows X^n sends binary strings over a noiseless channel to a decoder that wants to retrieve X^n without error.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$, and an encoding function $\phi_e : \mathcal{X}^n \rightarrow \{0, 1\}^*$ such that $\text{Im } \phi_e$ is prefix-free;
- A decoding function $\phi_d : \{0, 1\}^* \rightarrow \mathcal{X}^n$;

- The rate is the average length of the codeword per source symbol, i.e. $R \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]$.

Depending whether a perfect source reconstruction is required, two different regimes can be considered.

Definition 3.1.2 (Zero-error regime, vanishing error regime) *In the zero-error regime, the coding schemes (n, ϕ_e, ϕ_d) must satisfy the zero-error property:*

$$\mathbb{P}(\widehat{X}^n \neq X^n) = 0, \quad (3.1)$$

where $\widehat{X}^n = \phi_d(\phi_e(X^n))$. The objective is to find the minimal rate among all coding schemes under the zero-error constraint:

$$R_{SC0}^* \doteq \inf_{\substack{n, \phi_e, \phi_d \\ \text{zero-error}}} \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]. \quad (3.2)$$

In the vanishing error regime, the coding schemes (n, ϕ_e, ϕ_d) must satisfy the ϵ -error property:

$$\mathbb{P}(\widehat{X}^n \neq X^n) \leq \epsilon, \quad (3.3)$$

where $\widehat{X}^n = \phi_d(\phi_e(X^n))$. The objective is to find the minimal rate among all coding schemes under the ϵ -error constraint, with $\epsilon \rightarrow 0$:

$$R_{SC}^* \doteq \lim_{\epsilon \rightarrow 0} \inf_{\substack{n, \phi_e, \phi_d \\ \epsilon\text{-error}}} \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]. \quad (3.4)$$

Determining the optimal rate means determining theoretic limit of all possible coding schemes in the source coding problem. The zero-error regime induces stronger constraints on the coding schemes than the vanishing error regime, thus leading to a higher optimal rate.

Theorem 3.1.3 (from [59])

$$R_{SC}^* = H(X). \quad (3.5)$$

Theorem 3.1.4 (from [34])

$$R_{SC0}^* = H(X). \quad (3.6)$$

The optimal rate is characterized by the entropy of the information source: Theorem 3.1.3 and Theorem 3.1.4 show that the optimal rate in this setting equals $H(X)$, for both vanishing error and zero-error regimes. This stems from the fact that the optimal rate in the vanishing error regime can be achieved with a Huffman algorithm, which also satisfies the zero-error property.

Alternatively by Remark 3.1.5, a possible zero-error coding strategy consists in indexing the typical set w.r.t. P_X following the approach from [59], and using the correcting device.

Remark 3.1.5 (Correcting device) *In this setting, since the encoder knows the realizations of every random variable in the problem, it can simulate every decoding. Therefore, every coding scheme in the vanishing error regime can be turned into zero-error ones with negligible impact on the rate in the following way: the encoder uses the coding scheme in the vanishing error regime, along with a bit of error. The latter equals 0 in case no decoding error occurs and 1 if any decoding error happens, and all source sequences are sent to the decoder in that case. The correcting device has a negligible additional cost on the rate, due to the negligible cost of the flag bit and the low probability of error of a coding scheme in the vanishing error regime. As a consequence, both optimal rates of this problem in the vanishing error and in the zero-error regime are equal.*

3.2 Channel coding

The channel coding problem is introduced in [59] in the vanishing error regime; and in [58] in the zero-error regime. It is a well-known example where the respective optimal rates in these regimes are different. In particular, a full-support distribution $P_{Y|X}$ for the channel may yield a positive channel capacity (i.e. optimal rate) in the vanishing error regime, and a zero-error capacity equal to 0.

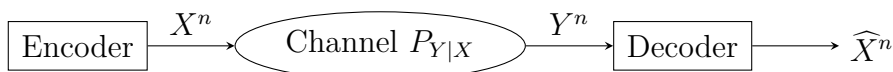


Figure 3.2 – The channel coding problem.

Definition 3.2.1 (Channel coding problem) *The channel coding problem of Figure 3.2 is described by:*

- Two finite sets \mathcal{X}, \mathcal{Y} ;
- A Discrete Memoryless Channel (DMC) with input alphabet \mathcal{X} , output alphabet \mathcal{Y} and transition probability $P_{Y|X} \in \Delta(\mathcal{Y})^{|\mathcal{X}|}$;
- An encoder that sends inputs over the DMC, and a decoder that receives the DMC's outputs.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$, and a codebook $\mathcal{C}_n \subseteq \mathcal{X}^n$;
- A decoding function $\phi_d : \mathcal{Y}^n \rightarrow \mathcal{X}^n$;
- The rate is the average number of messages transmitted per channel use, i.e. $\frac{1}{n} \log |\mathcal{C}_n|$.

The channel coding problem is formally defined for two different regimes, depending whether a positive probability of error is allowed when communicating through the DMC.

Definition 3.2.2 (Zero-error regime, vanishing error regime) *In the zero-error regime, the coding scheme $(n, \mathcal{C}_n, \phi_d)$ must satisfy the zero-error property:*

$$\mathbb{P}(\widehat{X}^n \neq X^n) = 0; \quad (3.7)$$

where $X^n \sim \text{Unif}(\mathcal{C}_n)$, $\widehat{X}^n = \phi_d(Y^n)$, and Y^n is drawn conditionally w.r.t. X^n using $P_{Y|X}^n$. The objective is to find the maximal rate among all coding schemes that satisfy the zero-error property:

$$R_{C0}^* \doteq \sup_{\substack{n, \mathcal{C}_n, \phi_d \\ \text{zero-error}}} \frac{1}{n} \log |\mathcal{C}_n|. \quad (3.8)$$

In the **vanishing error regime**, the coding schemes $(n, \mathcal{C}_n, \phi_d)$ must satisfy the ϵ -error property:

$$\mathbb{P}(\widehat{X}^n \neq X^n) \leq \epsilon, \quad (3.9)$$

where $X^n \sim \text{Unif}(\mathcal{C}_n)$, $\widehat{X}^n = \phi_d(Y^n)$, and Y^n is drawn conditionally w.r.t. X^n using $P_{Y|X}^n$. The objective is to find the maximal rate among all coding schemes that satisfy the ϵ -error

property, with $\epsilon \rightarrow 0$:

$$R_C^* \doteq \lim_{\epsilon \rightarrow 0} \sup_{\substack{n, \mathcal{C}_n, \phi_d \\ \epsilon\text{-error}}} \frac{1}{n} \log |\mathcal{C}_n|. \quad (3.10)$$

3.2.1 Channel coding in the vanishing error regime

The channel capacity in the vanishing error regime is defined by $\max_{P_X \in \Delta(\mathcal{X})} I(X; Y)$; and as shown by Shannon in Theorem 3.2.3, this quantity characterizes the optimal rate of communication over a DMC in the vanishing error regime. This comes from the fact that for all $P_X \in \Delta(\mathcal{X})$, the encoder may choose $2^{nI(X; Y)}$ random codewords from \mathcal{X}^n by using P_X , and expect a correct decoding with high probability by typicality arguments. Taking the maximum over P_X yields the highest achievable rate.

Theorem 3.2.3 (from [59])

$$R_C^* = \max_{P_X \in \Delta(\mathcal{X})} I(X; Y). \quad (3.11)$$

3.2.2 Zero-error channel coding

In the zero-error regime, an adequate graph G_C can be associated to a given instance of channel coding problem in Figure 3.2. This graph is called “characteristic graph” of the problem, as it encompasses the problem data in its structure: the vertices are the source alphabet, and two channel input symbols x, x' are adjacent if they are “confusable”, i.e. $P_{Y|X}(y|x)P_{Y|X}(y|x') > 0$ for some channel output symbol y . In other words, when x and x' are adjacent, knowing the realization y does not allow to distinguish between the realizations x and x' with probability of error 0.

Definition 3.2.4 (Characteristic graph) *Let \mathcal{X}, \mathcal{Y} be two finite sets and $P_{Y|X}$ be a conditional distribution from $\Delta(\mathcal{Y})^{|\mathcal{X}|}$. The characteristic graph associated to $P_{Y|X}$ is defined by:*

- \mathcal{X} as set of vertices,
- $x, x' \in \mathcal{X}$ are adjacent if $P_{Y|X}(y|x)P_{Y|X}(y|x') > 0$ for some $y \in \mathcal{Y}$.

The AND product \wedge is a binary operator on graphs, and is used to build the characteristic graph for more than one channel use. More precisely, two sequences of channel

inputs x^n, x^m are adjacent in the n -th AND power $G_C^{\wedge n}$ (i.e. iterated AND product of G_C) if $P_{Y|X}^n(y^n|x^n)P_{Y|X}^n(y^n|x^m) > 0$ for some sequence of channel outputs y^n .

In Definition 3.2.5, the AND product of graphs with an underlying distribution is defined, as it will be useful for the zero-error source coding problem. In the case of zero-error channel coding, graphs without distribution are considered, but the structure of the AND product is the same regardless of the distribution. The AND product is also called “strong product” or “normal product” in the literature (for example in [44, 47]).

Definition 3.2.5 (AND product \wedge) Let $G_1 = (\mathcal{V}_1, \mathcal{E}_1, P_{V_1})$, $G_2 = (\mathcal{V}_2, \mathcal{E}_2, P_{V_2})$ be two probabilistic graphs, their AND product $G_1 \wedge G_2$ is a probabilistic graph defined by:

- $\mathcal{V}_1 \times \mathcal{V}_2$ as set of vertices,
- $P_{V_1} \otimes P_{V_2}$ as probability distribution on the vertices,
- $(v_1v_2), (v'_1v'_2)$ are adjacent if

$$v_1v'_1 \in \mathcal{E}_1 \text{ AND } v_2v'_2 \in \mathcal{E}_2, \quad (3.12)$$

with the convention of self-adjacency for all vertices.

We denote by $G_1^{\wedge n}$ the n -th AND power:

$$G_1^{\wedge n} = G_1 \wedge \dots \wedge G_1 \quad (n \text{ times}). \quad (3.13)$$

The AND product of graphs without probability distribution has the vertex set and edges defined above, without underlying probability distribution.

If there is a pair of codewords x^n, x^m adjacent in $G_C^{\wedge n}$ in the codebook used by the encoder, then upon receiving the y^n such that $P_{X|Y}^n(x^n|y^n)P_{X|Y}^n(x^m|y^n) > 0$, the decoder is unable to determine whether the encoder sent x^n or x^m , which prevents zero-error decoding. Therefore, a zero-error decoding at the end of the time horizon n is possible if and only if the encoder uses a codebook formed of pairwise non-adjacent symbols in $G_C^{\wedge n}$, i.e. an independent set, which is formally defined below.

Definition 3.2.6 (Independent subset, independence number α) Let $G = (\mathcal{V}, \mathcal{E})$ be a graph. A subset $\mathcal{S} \subseteq \mathcal{V}$ is independent in G if $xx' \notin \mathcal{E}$ for all $x, x' \in \mathcal{S}$. The independence number is the maximal size of an independent set in G , and is denoted by $\alpha(G)$.

Definition 3.2.7 (Zero-error capacity of a graph C_0) Let G be a graph, its zero-error capacity is defined by

$$C_0(G) \doteq \lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha(G^{\wedge n}). \quad (3.14)$$

By construction, the optimal rate for a coding scheme with the zero-error property and time horizon n is $\frac{1}{n} \log \alpha(G^{\wedge n})$. As shown in Proposition 3.2.8, the asymptotic optimal rate is $\lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha(G^{\wedge n})$; this quantity is called the zero-error capacity of a channel (or the zero-error capacity of its characteristic graph), and represents the best zero-error communication rate through a DMC.

Note that, by convention, we define the zero-error capacity with the logarithm. Another existing convention (for example in [44]) for the zero-error capacity is $\Theta(G) \doteq \lim_{n \rightarrow \infty} \sqrt[n]{\alpha(G^{\wedge n})}$; which is equivalent in the sense that $C_0 = \log \Theta$.

Theorem 3.2.8 (from [58]) The optimal rate in the zero-error channel coding setting writes

$$R_{C_0}^* = C_0(G_C), \quad (3.15)$$

where G_C is the characteristic graph associated to the distribution $P_{Y|X}$.

Determining a single-letter expression for C_0 is a wide open problem. We present in Section 5.5 some examples from the literature where C_0 is known, in particular perfect graphs. The Lovász θ function, introduced in [44], is an upper bound on the zero-error capacity. This function is used to show that $C_0(C_5) = \frac{\log 5}{2}$, which makes C_5 the minimally non-perfect graph for which C_0 is known. Further observations on the θ function are derived by Sason in [54]. The zero-error capacity of C_7 is still unknown. Several existing lower bounds on $C_0(C_7)$ result from an independent set found by a computer program; in particular by Vesel and Žerovnik in [72], by Mathew and Östergård in [48], and by Polak and Schrijver in [51].



Figure 3.3 – The cycle graphs C_5 , C_7 .

3.2.3 Related works

The computability of C_0 is investigated in [9] by Boche and Deppe, and they prove that C_0 is not computable. An asymptotic expression for C_0 using semiring homomorphisms is given by Zuydam et al. in [78]. In [31], Gu and Shayevitz study the two-way channel case. An extension of C_0 for secure communication is developed in [74] by Wiese et al.

Another related setting is zero-error transmission over a discrete channel with memory. The case of binary channels with one memory is studied by Ahlswede et al. in [1], followed by Cohen et al. in [17], and the remaining unsolved cases are solved by Cao et al. in [12]. The case of binary channels with two memories is studied by Zhang et al. in [77].

Finally, the Sperner capacity is an extension of the zero-error capacity to digraphs introduced in [28] by Gargano et al. Upper bounds on the Sperner capacity are developed by Alon in [3] based on the maximum outdegree of the digraph; and by Körner et al. in [40] based on an adaptation to digraphs of the local chromatic number.

The interested reader may refer to Körner and Orlicsky’s survey in [38], and Simonyi’s survey in [63].

3.3 Slepian-Wolf problem

The Slepian-Wolf problem is introduced in [64] by Slepian and Wolf, in the vanishing error regime; and the zero-error variant of this problem is presented by Alon and Orlicsky, in [5]. This problem corresponds to a situation in data compression where the decoder has a side-information Y about the source X that has to be retrieved. Its optimal rate in the zero-error regime is given by the complementary graph entropy \overline{H} of the characteristic graph, and finding a single-letter formula for \overline{H} is an open question. As presented in Theorem 3.3.7 and Theorem 3.3.3, the respective optimal rates in the zero-error regime and in the vanishing error regime are different, in particular when the source distribution is full-support. Note that the correcting device described in Remark 3.1.5 cannot be used here for the zero-error regime, as the encoder does not know the realizations of Y .

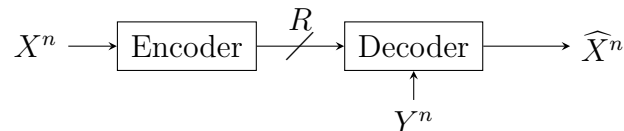


Figure 3.4 – The zero-error Slepian-Wolf problem.

Definition 3.3.1 (Zero-error Slepian-Wolf problem) *The zero-error source coding problem of Figure 3.4 is described by:*

- Two finite sets \mathcal{X} , \mathcal{Y} , and a source distribution $P_{X,Y} \in \Delta(\mathcal{X} \times \mathcal{Y})$;
- For all $n \in \mathbb{N}^*$, (X^n, Y^n) is the random sequence of n copies of (X, Y) , drawn in an i.i.d. fashion using $P_{X,Y}$.
- An encoder that knows X^n sends binary strings over a noiseless channel to a decoder that knows Y^n , and that wants to retrieve X^n without error.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$, and an encoding function $\phi_e : \mathcal{X}^n \rightarrow \{0,1\}^*$ such that $\text{Im } \phi_e$ is prefix-free;
- A decoding function $\phi_d : \mathcal{Y}^n \times \{0,1\}^* \rightarrow \mathcal{X}^n$;
- The rate is the average length of the codeword per source symbol, i.e. $R \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]$.

Definition 3.3.2 (Zero-error regime, vanishing error regime) *In the **zero-error regime**, the coding schemes (n, ϕ_e, ϕ_d) must satisfy the **zero-error property**:*

$$\mathbb{P}(\widehat{X}^n \neq X^n) = 0, \quad (3.16)$$

where $\widehat{X}^n = \phi_d(Y^n, \phi_e(X^n))$. The objective is to find the minimal rate among all coding schemes under the zero-error constraint:

$$R_{SW0}^* \doteq \inf_{\substack{n, \phi_e, \phi_d \\ \text{zero-error}}} \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]. \quad (3.17)$$

In the **vanishing error regime**, the coding schemes (n, ϕ_e, ϕ_d) must satisfy the **ϵ -error property**:

$$\mathbb{P}(\widehat{X}^n \neq X^n) \leq \epsilon, \quad (3.18)$$

where $\widehat{X}^n = \phi_d(Y^n, \phi_e(X^n))$. The objective is to find the minimal rate among all coding schemes under the ϵ -error constraint, with $\epsilon \rightarrow 0$:

$$R_{SW}^* \doteq \lim_{\epsilon \rightarrow 0} \inf_{\substack{n, \phi_e, \phi_d \\ \epsilon\text{-error}}} \frac{1}{n} \mathbb{E}[\ell \circ \phi_e(X^n)]. \quad (3.19)$$

Theorem 3.3.3 (from [64]) *The optimal rate in the Slepian-Wolf setting in the vanishing error regime writes*

$$R_{SW}^* = H(X|Y). \quad (3.20)$$

In the vanishing error regime, the optimal rate $H(X|Y)$ in Theorem 3.3.3 remains the same if the encoder also observes the sequence Y^n .

3.3.1 Zero-error Slepian-Wolf problem

In a similar way to zero-error channel coding, the characteristic graph G_{SW} of an instance of zero-error Slepian-Wolf problem in Figure 3.4 is defined in Definition 3.2.4, with the conditional source distribution $P_{Y|X}$. This graph was first used by Witsenhausen in [75]. However in the zero-error Slepian-Wolf setting it is a probabilistic graph, as it has the underlying distribution P_X on its vertices.

Definition 3.3.4 (Probabilistic graph) *A probabilistic graph G is a tuple $(\mathcal{V}, \mathcal{E}, P_V)$, where \mathcal{V} and \mathcal{E} are respectively the sets of vertices and edges; and $P_V \in \Delta(\mathcal{V})$ is a probability distribution on the vertices.*

Assume that $\phi_e(x^n) = \phi_e(x'^n)$ for some x^n, x'^n adjacent in $G_{SW}^{\wedge n}$, then upon receiving the y^n such that $P_{X|Y}^{\otimes n}(x^n|y^n)P_{X|Y}^{\otimes n}(x'^n|y^n) > 0$, the decoder is unable to determine whether the encoder sent x^n or x'^n , which prevents zero-error decoding. Therefore, the source sequences that are adjacent in the characteristic graph must be mapped to different codewords; hence the use of graph colorings, which map adjacent vertices to different colors.

Definition 3.3.5 (Coloring, chromatic number χ) *Let $G = (\mathcal{V}, \mathcal{E})$ be a graph. A mapping $c : \mathcal{V} \rightarrow \mathcal{C}$ is a coloring if $c^{-1}(i)$ is independent for all $i \in \mathcal{C}$. The chromatic number $\chi(G)$ is the smallest $|\mathcal{C}|$ such that there exists a coloring $c : \mathcal{V} \rightarrow \mathcal{C}$ of G .*

With high probability, the source sequence X^n will be typical w.r.t. P_X , therefore one can consider only the subgraph of $G_{SW}^{\wedge n}$ induced by the set of typical sequences $\mathcal{T}_\epsilon^n(P_X)$. Then the encoder colors this induced subgraph $G_{SW}^{\wedge n}[\mathcal{T}_\epsilon^n(P_X)]$ with a coloring that has the minimum number of colors, and sends the color index to the decoder if X^n is typical,

or the index of X^n in \mathcal{X}^n otherwise. This coding strategy has a rate upper-bounded by

$$\frac{1}{n} + \mathbb{P}(X^n \notin \mathcal{T}_\epsilon^n(P_X)) \log |\mathcal{X}| + \frac{1}{n} \log \chi(G_{SW}^{\wedge n}[\mathcal{T}_\epsilon^n(P_X)]); \quad (3.21)$$

where χ is the chromatic number and the decoder is able to retrieve X^n thanks to the color and Y^n . Koulgi et al. have shown in [41, Theorem 1] that taking the limit when n goes to infinity and ϵ goes to 0 yields the best achievable rate in the zero-error Slepian Wolf problem. This quantity is called complementary graph entropy, and is defined by Körner and Longo in [37].

Definition 3.3.6 (Complementary graph entropy \bar{H}) For all probabilistic graph $G = (\mathcal{V}, \mathcal{E}, P_V)$, the quantity $\bar{H}(G)$ is defined by:

$$\bar{H}(G) = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \chi(G^{\wedge n}[\mathcal{T}_\epsilon^n(P_V)]). \quad (3.22)$$

Theorem 3.3.7 (from [41, Theorem 1]) The optimal rate in the zero-error Slepian-Wolf setting writes

$$R_{SW0}^* = \bar{H}(G_{SW}), \quad (3.23)$$

where G_{SW} is the probabilistic graph formed of the characteristic graph associated to the distribution $P_{Y|X}$, with the underlying distribution P_X on its vertices.

The zero-error Slepian Wolf setting that we study is called “restricted inputs” by Alon and Orlitsky in [5] where they give another characterization of the optimal rate with chromatic entropies. The chromatic entropy $H_\chi(G_{SW}^{\wedge n})$ corresponds to the minimal entropy of a coloring of $G_{SW}^{\wedge n}$. Therefore, after normalization, it characterizes the best rate at a fixed number n of source uses with a perfect compression of the color. As stated in Theorem 3.3.9, by taking the limit when n goes to infinity, one obtains the optimal rate; and also, another expression for \bar{H} .

A stronger notion of zero-error coding, called “unrestricted inputs”, is also introduced in [5]. It requires the zero-error property to be satisfied even for the sequences of symbols (X^n, Y^n) that take values out of the support of $P_{X,Y}^n$. Alon and Orlitsky provide in [5], a single-letter formula for the optimal rate.

Definition 3.3.8 (Chromatic entropy H_χ) Let $G = (\mathcal{V}, \mathcal{E}, P_V)$ be a probabilistic graph,

its chromatic entropy is defined by

$$H_\chi(G) = \inf\{H(c(V)) \mid c \text{ is a coloring of } G\}. \quad (3.24)$$

Theorem 3.3.9 (from [5, Lemma 6]) For all probabilistic graph G ,

$$R_{SW0}^* = \lim_{n \rightarrow \infty} \frac{1}{n} H_\chi(G_{SW}^{\wedge n}). \quad (3.25)$$

The two distinct asymptotic expressions are equal to the optimal rate.

$$R_{SW0}^* = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \chi(G^{\wedge n}[\mathcal{T}_\epsilon^n(P_V)]) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\chi(G_{SW}^{\wedge n}). \quad (3.26)$$

Proposition 3.3.10 (Full support) If the distribution $P_{X,Y}$ has full support, then $R_{SW0}^* = H(X)$.

The case of $P_{X,Y}$ with full support is a worst case for the zero-error regime, for which the optimal rate is $H(X)$, instead of $H(X|Y)$ in the vanishing error regime.

There is no known single letter expression for the optimal rate R_{SW0}^* except for a few special cases; in particular for perfect graphs and $(C_5, \text{Unif}(\{1, \dots, 5\}))$.

CONTRIBUTIONS FOR CORRECTION-BASED ZERO-ERROR PROBLEMS

In Section 4.1, we present the source coding problems considered in this chapter. We call these problems “correction-based” as they can be tackled with a correcting device which gives the zero-error property to coding schemes designed for the vanishing error regime. As a result, for this class of problems, the optimal rate in the zero-error regime coincides with the optimal rate in the vanishing error regime.

We determine in Section 4.2 the feasible rate region of the zero-error source coding problem when side information may be present at the decoder, presented in Section 4.1.1. In Section 4.3 we give an inner and an outer bound on the feasible rate region of the zero-error source coding problem for a Gray-Wyner network, presented in Section 4.1.3. We show that these bounds coincide in several cases of interest.

4.1 Correction-based zero-error problems

4.1.1 Zero-error source coding when side information may be present at the decoder

This scenario arises in interactive compression, where the user can randomly access part of the data directly in the compressed domain. A source sequence X^n models the smallest entity that can be requested, for instance a file of a database, a frame of a video, or a block of an omnidirectional image [8, 49, 53].

In this setting, the encoder has access to a side information but does not know whether the decoder has access to it; and the encoding is divided into two steps. In the first step, the encoder sends information without knowing whether side information is available at

the decoder. The message sent in the first step must be exploitable independently from the presence of the side information; furthermore, a zero-error decoding must be done if the decoder has access to the side information. In the second step, an information complement is sent by the encoder in the case where the decoder has no side information, and the decoder must be able to retrieve the source with zero-error. The two steps are modeled by two decoders and two noiseless channels.

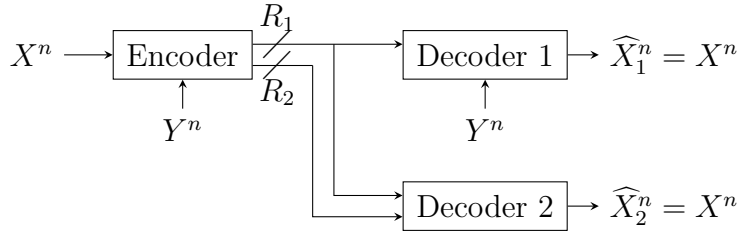


Figure 4.1 – Source coding when side information may be present.

Definition 4.1.1 *The setting of Figure 4.1 is described by:*

- Two finite sets \mathcal{X} , \mathcal{Y} , and a source distribution $P_{X,Y} \in \Delta(\mathcal{X} \times \mathcal{Y})$;
- For all $n \in \mathbb{N}^*$, (X^n, Y^n) is the random sequence of n copies of (X, Y) , drawn in an i.i.d. fashion using $P_{X,Y}$.
- An encoder that knows (X^n, Y^n) sends binary strings over two noiseless channels to two decoders that want to retrieve X^n without error. Decoder 1 has Y^n as side information.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$,
- Two encoding functions and two decoding functions

$$\phi_e^{(1)} : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{0, 1\}^*, \quad \phi_e^{(2)} : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{0, 1\}^*, \quad (4.1)$$

$$\phi_d^{(1)} : \{0, 1\}^* \times \mathcal{Y}^n \rightarrow \mathcal{X}^n, \quad \phi_d^{(2)} : (\{0, 1\}^*)^2 \rightarrow \mathcal{X}^n, \quad (4.2)$$

such that $\text{Im } \phi_e^{(1)}$ and $\text{Im } \phi_e^{(2)}$ are prefix-free;

- *The rates over each channel are the average length of the codeword per source symbol:*

$$R_1 \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e^{(1)}(X^n, Y^n)] \quad R_2 \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e^{(2)}(X^n, Y^n)]; \quad (4.3)$$

- $n, \phi_e^{(1)}, \phi_e^{(2)}, \phi_d^{(1)}, \phi_d^{(2)}$ must satisfy the zero-error property:

$$\mathbb{P}(\widehat{X}_1^n \neq X^n) = \mathbb{P}(\widehat{X}_2^n \neq X^n) = 0; \quad (4.4)$$

where

$$\widehat{X}_1^n = \phi_d^{(1)}(\phi_e^{(1)}(X^n, Y^n), Y^n), \quad (4.5)$$

$$\widehat{X}_2^n = \phi_d^{(2)}(\phi_e^{(1)}(X^n, Y^n), \phi_e^{(2)}(X^n, Y^n)). \quad (4.6)$$

The objective is to find the feasible rate region \mathcal{R}_{SI} , which is the closure of the set

$$\{(R_1, R_2) \in \mathbb{R}^2 \mid \exists(n, \phi_e^{(1)}, \phi_e^{(2)}, \phi_d^{(1)}, \phi_d^{(2)}) \text{ zero-error, with rates } (R_1, R_2)\}. \quad (4.7)$$

A way to achieve zero-error coding in the setting of Figure 4.1 is to use conditional Huffman coding of X knowing Y ; and send the source X to decoder 1 at rate $R_1 = H(X|Y)$. Then, to recover the source X , decoder 2 needs to obtain the side information Y , which requires a rate of $R_2 = H(Y) \geq I(X; Y)$.

In order to be exploitable by both decoders, part of the information sent through the common channel must be independent from Y . For this reason the setting of Figure 4.1 is closely related to the Slepian and Wolf (SW) problem in [64], seen as lossless source coding with side information at the decoder only. In this work, it is shown that the corner point $(R_1, R_2) = (H(X|Y), I(X; Y))$ is achievable in the vanishing probability of error regime. In [19], Csiszar proved in that linear codes achieve the optimal SW rate region. Several works in [13, 14, 73] investigate the duality between SW setting and channel coding using linear codes, as the side information Y can be seen as the input of a virtual channel with input X . However these tools cannot be straightforwardly adapted to the zero-error setting, as the linear codes proposed also present a vanishing probability of error.

The setting of Figure 4.1 can be seen as a zero-error variant with side informations known at the encoder of the successive refinement problem proposed by Kaspi in [35]; later generalized by Timo et al. in [67] for more than two decoders. Even if the lossy reconstruction of the source makes it fundamentally different from the zero-error setting, there are notable examples that present the same tools as in SW. The side information scalable source coding (i.e. the decoder 2 has a side information Y' s.t. $X \rightarrow Y \rightarrow Y'$) in [66] for instance uses nested random binning. This random binning approach was further developed in [2] to give a unified coding scheme that works for both scalable source coding

and Wyner-Ziv successive refinement in [65] (i.e. the decoder 2 has a side information Y' s.t. $X \rightarrow Y' \rightarrow Y$).

In [45], Ma and Cheng use linear codes in a zero-error SW restriction, under symmetry assumptions on the source. However, a zero-error SW coding scheme in our setting does not use at all the side information knowledge at the encoder. A more in-depth review of the literature for the zero-error SW setting can be found in Section 3.3.1.

In Section 4.2, we characterize the feasible rate region. More precisely, we show that the pair of rates $(R_1, R_2) = (H(X|Y), I(X; Y))$ is achievable in the zero-error regime and moreover, it is the corner-point of the feasible rate region.

4.1.2 Gray-Wyner problem

The problem built by Gray and Wyner in [30], aims at capturing the common information between two random variables X and Y .

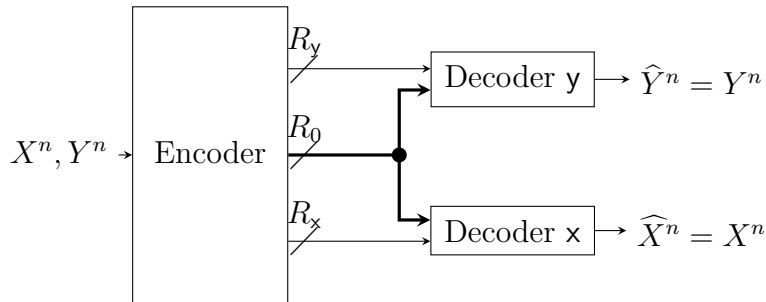


Figure 4.2 – The Gray-Wyner source coding problem.

Definition 4.1.2 *The setting of Figure 4.2 is described by:*

- Two finite sets \mathcal{X} , \mathcal{Y} , and a source distribution $P_{X,Y} \in \Delta(\mathcal{X} \times \mathcal{Y})$;
- For all $n \in \mathbb{N}^*$, (X^n, Y^n) is the random sequence of n copies of (X, Y) , drawn in an i.i.d. fashion using $P_{X,Y}$.
- An encoder that knows (X^n, Y^n) sends binary strings over three noiseless channels to the decoder x (resp. y) that wants to retrieve X^n (resp. Y^n) without error.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$,

- Three encoding functions and two decoding functions

$$\phi_e^{(0)}, \phi_e^{(x)}, \phi_e^{(y)} : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{0, 1\}^*, \quad (4.8)$$

$$\phi_d^{(x)} : (\{0, 1\}^*)^2 \rightarrow \mathcal{X}^n, \quad (4.9)$$

$$\phi_d^{(y)} : (\{0, 1\}^*)^2 \rightarrow \mathcal{Y}^n, \quad (4.10)$$

such that $\text{Im } \phi_e^{(i)}$ is prefix-free for all $i \in \{0, x, y\}$;

- The rates over each channel are the average length of the codeword per source symbol:

$$R_i \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e^{(i)}(X^n, Y^n)] \quad (4.11)$$

for all $i \in \{0, x, y\}$;

- $n, (\phi_e^{(i)})_{i \in \{0, x, y\}}, (\phi_d^{(j)})_{j \in \{x, y\}}$ must satisfy the ϵ -error property:

$$\mathbb{P}(\widehat{X}^n \neq X^n) \leq \epsilon \quad (4.12)$$

$$\mathbb{P}(\widehat{Y}^n \neq Y^n) \leq \epsilon; \quad (4.13)$$

where

$$\widehat{X}^n = \phi_d^{(x)}(\phi_e^{(0)}(X^n, Y^n), \phi_e^{(x)}(X^n, Y^n)), \quad (4.14)$$

$$\widehat{Y}^n = \phi_d^{(y)}(\phi_e^{(0)}(X^n, Y^n), \phi_e^{(y)}(X^n, Y^n)). \quad (4.15)$$

The objective is to find the feasible rate region \mathcal{R} , which is the closure of the set

$$\bigcap_{\epsilon > 0} \left\{ (R_0, R_x, R_y) \in \mathbb{R}^3 \mid \exists n, (\phi_e^{(i)})_{i \in \{0, x, y\}}, (\phi_d^{(j)})_{j \in \{x, y\}} \epsilon\text{-error, with rates } (R_0, R_x, R_y) \right\}. \quad (4.16)$$

In [30], Gray and Wyner have determined the feasible rate region in the vanishing error regime, as illustrated in Theorem 4.1.3.

Theorem 4.1.3 *The feasible rate region is the closure of the following set:*

$$\bigcup_{\substack{\mathcal{W} \text{ finite set,} \\ P_{W|X,Y} \in \Delta(\mathcal{W})^{|\mathcal{X}||\mathcal{Y}|}}} \left\{ (R_0, R_x, R_y) \in \mathbb{R}^3 \mid R_0 \geq I(W; X, Y), R_x \geq H(X|W), R_y \geq H(Y|W) \right\}. \quad (4.17)$$

Following the same idea as in Remark 3.1.5, this region is also the feasible rate region in the zero-error regime (i.e. $\mathbb{P}(\widehat{X}^n \neq X^n) = \mathbb{P}(\widehat{Y}^n \neq Y^n) = 0$).

4.1.3 Zero-error source coding for a Gray-Wyner network when side information may be present at the decoder

The following setting is an extension of both problems presented in Section 4.1.2 and Section 4.1.1, in which there are two uncertainties: firstly, the encoder has access to two sources and their respective side information but does not know which one will be requested by the decoder; and secondly, the side information may be present at the decoder. The encoding is divided into three steps. In the first step, the encoder sends information without knowing which source is requested; we aim at sending the common information between the two sources, hence the “Gray-Wyner network” name for this problem. In the second step, the encoder knows which source is requested but does not know if the decoder has access to the side information; if the latter is present at the decoder, then a zero-error decoding must be done. The third step is an information complement sent by the encoder in the case where the decoder does not have the side information; and the decoder must be able to retrieve the source with zero-error.

Definition 4.1.4 *The setting of Figure 4.3 is described by:*

- Four finite sets $\mathcal{U}, \mathcal{V}, \mathcal{X}, \mathcal{Y}$, and a source distribution $P_{U,V,X,Y} \in \Delta(\mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y})$;
- For all $n \in \mathbb{N}^*$, (U^n, V^n, X^n, Y^n) is the random sequence of n copies of (U, V, X, Y) , drawn in an i.i.d. fashion using $P_{U,V,X,Y}$.
- An encoder that knows (U^n, V^n, X^n, Y^n) sends binary strings over four noiseless channels to three decoders that want to retrieve X^n or Y^n without error. Decoder x_1 (resp. y) has U^n (resp. V^n) as side information.

A coding scheme in this setting is described by:

- A time horizon $n \in \mathbb{N}^*$,

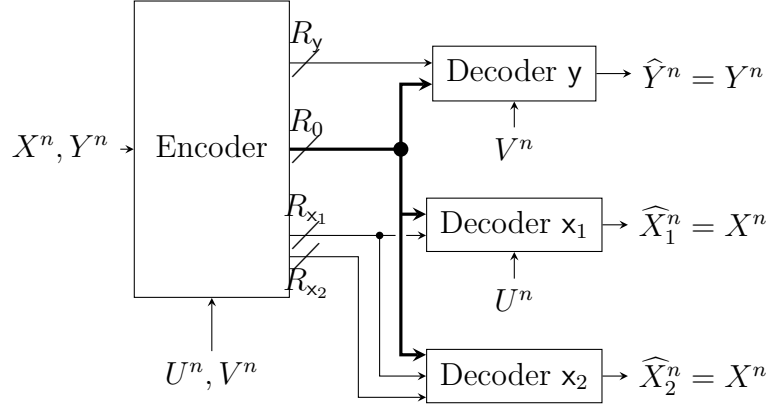


Figure 4.3 – Zero-error source coding for a Gray-Wyner network when side information may be present.

- Four encoding functions and three decoding functions

$$\phi_e^{(0)}, \phi_e^{(x_1)}, \phi_e^{(x_2)}, \phi_e^{(y)} : \mathcal{U}^n \times \mathcal{V}^n \times \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{0, 1\}^*, \quad (4.18)$$

$$\phi_d^{(x_1)} : (\{0, 1\}^*)^2 \times \mathcal{U}^n \rightarrow \mathcal{X}^n, \quad \phi_d^{(x_2)} : (\{0, 1\}^*)^3 \rightarrow \mathcal{X}^n, \quad (4.19)$$

$$\phi_d^{(y)} : (\{0, 1\}^*)^2 \times \mathcal{V}^n \rightarrow \mathcal{Y}^n, \quad (4.20)$$

such that $\text{Im } \phi_e^{(i)}$ is prefix-free for all $i \in \{0, x_1, x_2, y\}$;

- The rates over each channel are the average length of the codeword per source symbol:

$$R_i \doteq \frac{1}{n} \mathbb{E}[\ell \circ \phi_e^{(i)}(U^n, V^n, X^n, Y^n)] \quad (4.21)$$

for all $i \in \{0, x_1, x_2, y\}$;

- $n, (\phi_e^{(i)})_{i \in \{0, x_1, x_2, y\}}, (\phi_d^{(j)})_{j \in \{x_1, x_2, y\}}$ must satisfy the zero-error property:

$$\mathbb{P}(\widehat{X}_1^n \neq X^n) = \mathbb{P}(\widehat{X}_2^n \neq X^n) = \mathbb{P}(\widehat{Y}^n \neq Y^n) = 0; \quad (4.22)$$

where

$$\widehat{X}_1^n = \phi_d^{(1)}(\phi_e^{(0)}(S^n), \phi_e^{(x_1)}(S^n), U^n), \quad (4.23)$$

$$\widehat{X}_2^n = \phi_d^{(2)}(\phi_e^{(0)}(S^n), \phi_e^{(x_1)}(S^n), \phi_e^{(x_2)}(S^n)), \quad (4.24)$$

$$\widehat{Y}^n = \phi_d^{(y)}(\phi_e^{(0)}(S^n), \phi_e^{(y)}(S^n), V^n), \quad (4.25)$$

and $S^n = (U^n, V^n, X^n, Y^n)$.

The objective is to find the feasible rate region \mathcal{R}_{GW} , which is the closure of the set

$$\left\{ (R_0, R_{x_1}, R_{x_2}, R_y) \in \mathbb{R}^4 \mid \right. \quad (4.26)$$

$$\left. \exists n, \left(\phi_e^{(i)} \right)_{i \in \{0, x_1, x_2, y\}}, \left(\phi_d^{(j)} \right)_{j \in \{x_1, x_2, y\}} \text{ zero-error, with rates } (R_0, R_{x_1}, R_{x_2}, R_y) \right\}. \quad (4.27)$$

This scenario can be likened to zero-error source coding problems with side information at the decoders, but not at the encoder. In such problems, the encoder has partial or no knowledge of the side information, and zero-error and vanishing error constraints do not lead to the same rate, in general. The zero-error constraint leads to the use of graph theoretical concepts for the characterization of the optimal rate. Witsenhausen first studied in [75] whether Slepian and Wolf’s results can be extended to the zero-error case, using the chromatic number of the characteristic graph of the source distribution. In [62], Simonyi studied the generalization to more than one decoder with different side information, and proved that the optimal rate is the one associated to the worst side information (in the sense of Witsenhausen). Tuncel and Rose later extended this result to variable-length codes in [70]. A more in-depth review of the literature for the zero-error Slepian-Wolf setting can be found in Section 3.3.1.

A second class is the zero-error source coding problems with an encoder that has access to all side information, and decoders have access to the different side information. In these problems, the characterization of the optimal rates can be done using a packing lemma in type classes and the following correcting argument: the source sequence can be fully transmitted to each decoder whenever a decoding fails. This correcting argument allows for the extension of lossless results into zero-error results with same rates. A particular case of Sgarro’s problem in [57] when side information are fully transmitted by the helpers can be extended to the zero-error case when the side information is also available at the encoder. In Section 4.2, we characterize the feasible rate region of a problem in this class presented in Section 4.1.1, where conditional coding does not achieve optimal rates.

Finally, the setting of Figure 4.3 is closely related to several problems in the literature, with different side information at the decoders and different desired source random variable, which have been studied under a vanishing error hypothesis. A covering lemma in type classes can be used to characterize the optimal rates in such problems, for example the lossless Gray and Wyner’s (GW) problem [30]. In [68], Timo et al. study the GW setting with side information at the decoders only and give an inner bound on the

feasible rate region. In [7], Benammar and Zaidi consider a lossy GW setting with side information at the decoders only and with an hypothesis on the variables requested at the decoders. Laich and Wigger study in [43] the influence of the side information at the encoder in the lossless Kaspi/Heegard-Berger problem. They show that, for some source distributions, adding side information at the encoder lowers the minimal required rate to decode without loss.

In Section 4.3 we give an inner bound and an outer bound on the feasible rate region of the setting in Figure 4.3.

4.2 Main results for the zero-error source coding problem when side information may be present at the decoder

We now determine the feasible rate region of the zero-error source coding problem when side information may be present at the decoder, presented in Section 4.1.1

Theorem 4.2.1

$$\mathcal{R}_{SI} = \left\{ (R_1, R_2) \in \mathbb{R}^2 \mid R_1 \geq H(X|Y), R_1 + R_2 \geq H(X) \right\}. \quad (4.28)$$

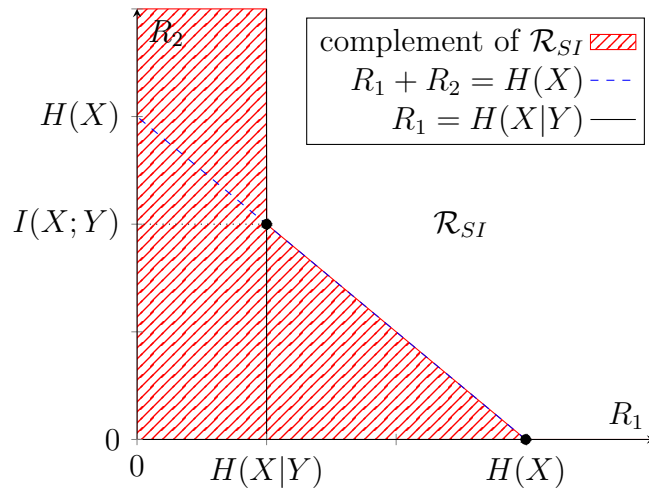


Figure 4.4 – Zero-error achievable rate region \mathcal{R}_{SI} .

Proof. [Converse of Theorem 4.2.1] In this setting, each decoder must retrieve X with zero-error. Using Shannon lossless source coding result [18, Theorem 5.3.1] and Slepian-Wolf Theorem [64, Theorem 2] on each decoder, we have $R_1 \geq H(X|Y)$ and $R_1 + R_2 \geq H(X)$, as the zero-error source codes are a subclass of lossless codes considered for these converses. \square

Our achievability result relies on a random coding argument. We use Csiszar and Körner’s method of types [20, Chapter 2] in order to calibrate a linear code which is used to partition the set of source sequences. The encoder sends the coset of the source sequence to all decoders and the index of the source sequence in its coset to decoder 2. We show that the zero-error property is satisfied and the corresponding rates converge to the pair of target rates $(H(X|Y), I(X; Y))$.

In order to prove Theorem 4.2.1, we show that

$$(H(X|Y), I(X; Y)) \in \mathcal{R}_{SI}. \quad (4.29)$$

In order to complete the achievability result we use a time sharing with the point $(H(X), 0)$, which is known to be achievable by compressing X using a Huffman code and sending the resulting binary sequence via $\phi_e^{(1)}$.

4.2.1 Definitions for the achievability proof

Definition 4.2.2 (Type class, V -shell) Let $Q_{X,Y} \in \Delta_n(\mathcal{X} \times \mathcal{Y})$, its type class is the set $\mathcal{T}_0^n(Q_{X,Y}) = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n \mid T_{x^n, y^n} = Q_{X,Y}\}$. The $Q_{X|Y}$ -shell of a sequence $y^n \in \mathcal{Y}^n$ is the set $\mathcal{T}_{Q_{X|Y}}(y^n) = \{x^n \in \mathcal{X}^n \mid T_{x^n, y^n} = Q_{X,Y}\}$.

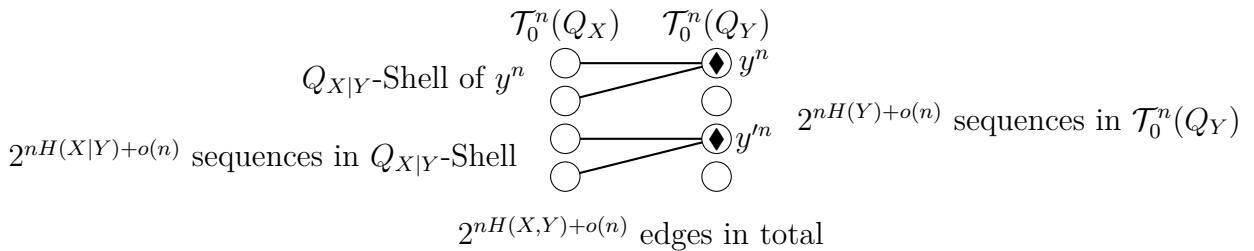


Figure 4.5 – An illustration of type classes and $Q_{X|Y}$ -shells. Edges correspond to jointly typical sequences with $\epsilon = 0$, i.e. $T_{x^n, y^n} = Q_{X,Y}$. At most $2^{nI(X;Y)+o(n)}$ disjoint $Q_{X,Y}$ -shells can be packed in $\mathcal{T}_0^n(Q_X)$, as $\frac{|\mathcal{T}_0^n(Q_Y)|}{|\mathcal{T}_{Q_{X|Y}}(y^n)|} = \frac{2^{nH(X)+o(n)}}{2^{nH(X|Y)+o(n)}} = 2^{nI(X;Y)+o(n)}$.

Definition 4.2.3 (Generator/parity matrix, syndrome, coset) Let \mathcal{A} be a finite set such that $|\mathcal{A}|$ is prime, so we can give $\mathcal{A} \simeq \mathbb{Z}/|\mathcal{A}|\mathbb{Z}$ a field structure. For all $n, k \in \mathbb{N}^*$, we denote by $\mathcal{M}_{n,k}(\mathcal{A})$ the set of $n \times k$ matrices over the finite field \mathcal{A} .

Let $k \in \mathbb{N}^*$, a generator matrix is a matrix $\mathbf{G} \in \mathcal{M}_{n,k}(\mathcal{A})$. An associated parity matrix is a matrix $\mathbf{H} \in \mathcal{M}_{n-k,n}(\mathcal{A})$ such that $\text{Im } \mathbf{G} = \text{Ker } \mathbf{H}$, where Im and Ker denote the image and the kernel, respectively.

The syndrome of a sequence $a^n \in \mathcal{A}^n$ is $\mathbf{H}a^n$. The coset associated to the syndrome $\mathbf{H}a^n$ is the set $\text{Im } \mathbf{G} + a^n = \{a'^n \in \mathcal{A}^n \mid \mathbf{H}a'^n = \mathbf{H}a^n\}$.

4.2.2 Coding scheme

For all $n \in \mathbb{N}^*$, we show the existence of a sequence of $(n, R_1^{(n)}, R_2^{(n)})$ -zero-error source codes that achieves the corner-point $(H(X|Y), I(X;Y))$ of the zero-error rate region \mathcal{R}_{SI} . Our proof is based on a linear code adjusted depending on the random type T_{X^n, Y^n} , and coset partitioning of the Hamming space.

We assume w.l.o.g. that $P_{X,Y} \neq P_X \otimes P_Y$. We also assume w.l.o.g. that $|\mathcal{X}|$ is prime number by padding (i.e. extending with zeros) $P_{X,Y}$ if necessary. We fix the block-length n and a constant parameter $\delta \in (0; \log |\mathcal{X}| - H(X|Y))$ that will represent a rate penalty.

- Random code generation: For each pair of sequences (x^n, y^n) , we define the parameter

$$k \doteq \left\lceil n - n \frac{H(T_{x^n, y^n}) - H(T_{y^n}) + \delta}{\log |\mathcal{X}|} \right\rceil^+. \quad (4.30)$$

where $\lceil \cdot \rceil$ denotes the ceiling function and $(\cdot)^+$ denotes $\max(\cdot, 0)$. We denote by K the random variable induced by k defined in (4.30), for the random sequences (X^n, Y^n) . A generator matrix $\mathbf{G} \in \mathcal{M}_{n,n}(\mathcal{X})$ is randomly drawn, with i.i.d. entries drawn according to the uniform distribution on \mathcal{X} . If $k \neq 0$, let \mathbf{G}_k be the matrix obtained by extracting the k first lines of \mathbf{G} , and \mathbf{H}_k a parity matrix associated to \mathbf{G}_k .

The random code \mathcal{C} consists of the set of random matrices $\mathcal{C} = (\mathbf{G}_k, \mathbf{H}_k)_{k \leq n}$. Before the transmission starts, a code realization is chosen and revealed to the encoder and both decoders.

- Encoding function $\phi_e^{(1)}$: Let $E \in \{0, 1\}$ be such that $E = 0$ if $K \neq 0$ and $(\text{Im } \mathbf{G}_K + X^n) \cap \mathcal{T}_{T_{X^n|Y^n}}(Y^n) = \{X^n\}$, $E = 1$ otherwise; where $T_{X^n|Y^n}$ is the conditional

distribution obtained from T_{X^n, Y^n} . Then we define

$$\phi_e^{(1)}(X^n, Y^n) = \begin{cases} b(T_{X^n, Y^n}, E, \mathbf{H}_K X^n) & \text{if } E = 0, \\ b(T_{X^n, Y^n}, E, X^n) & \text{if } E = 1, \end{cases} \quad (4.31)$$

where $b(\cdot)$ denotes the binary expansion.

- Encoding function $\phi_e^{(2)}$: If $E = 0$, the index of X^n in its coset $\text{Im } \mathbf{G}_K + X^n$ is compressed using a Huffman code with the distribution P_{X^n} . Let $\iota(\mathbf{G}_K, X^n, Y^n)$ be the resulting binary sequence, then we set

$$\phi_e^{(2)}(X^n, Y^n) = \iota(\mathbf{G}_K, X^n, Y^n). \quad (4.32)$$

Otherwise, $\phi_e^{(2)}(X^n, Y^n) = 0$.

- Decoding function $\phi_d^{(1)}$: It observes $\phi_e^{(1)}(X^n, Y^n)$ and extracts E and T_{X^n, Y^n} . If $E = 1$,

$$\phi_d^{(1)}(\phi_e^{(1)}(X^n, Y^n), Y^n) = X^n. \quad (4.33)$$

Otherwise $E = 0$, it extracts $\mathbf{H}_K X^n$ and determines the coset $\text{Im } \mathbf{G}_K + X^n$. Moreover, by using T_{X^n, Y^n} and Y^n it determines the $T_{X^n|Y^n}$ -shell $\mathcal{T}_{T_{X^n|Y^n}}(Y^n)$, and therefore returns an element

$$\phi_d^{(1)}(\phi_e^{(1)}(X^n, Y^n), Y^n) \in (\text{Im } \mathbf{G}_K + X^n) \cap \mathcal{T}_{T_{X^n|Y^n}}(Y^n).$$

- Decoding function $\phi_d^{(2)}$: It observes $\phi_e^{(1)}(X^n, Y^n)$ and extracts E and T_{X^n, Y^n} . If $E = 0$, it extracts $\mathbf{H}_K X^n$ and determines the coset $\text{Im } \mathbf{G}_K + X^n$, and it returns $\phi_d^{(2)}(\phi_e^{(1)}(X^n, Y^n), \phi_e^{(2)}(X^n, Y^n))$, the element of $\text{Im } \mathbf{G}_K + X^n$ with index $\phi_e^{(2)}(X^n, Y^n)$. If $E = 1$, it returns

$$\phi_d^{(2)}(\phi_e^{(1)}(X^n, Y^n), \phi_e^{(2)}(X^n, Y^n)) = X^n.$$

Remark 4.2.4 *The parameter K is selected so that when $K > 0$, the number of parity bits of the linear code asymptotically matches the conditional entropy.*

4.2.3 Zero-error property and rate analysis

We now prove that the code built in Section 4.2.2 satisfies the zero-error property. It is clear that both decoders retrieve X^n with zero-error when $E = 1$.

If $E = 0$, then by definition of E we have $(\text{Im } \mathbf{G}_K + X^n) \cap \mathcal{T}_{T_{X^n|Y^n}}(Y^n) = \{X^n\}$, hence $\phi_d^{(1)}(\phi_e^{(1)}(X^n, Y^n), Y^n) = X^n$ with probability 1. On the other hand, $\phi_e^{(2)}(X^n, Y^n) = \iota(\mathbf{G}_K, X^n, Y^n)$, so the element of $\text{Im } \mathbf{G}_K + X^n$ with index $\phi_e^{(2)}(X^n, Y^n)$ is X^n . Thus, $\phi_d^{(2)}(\phi_e^{(1)}(X^n, Y^n), \phi_e^{(2)}(X^n, Y^n)) = X^n$ with probability 1.

Lemma 4.2.5 (Rate analysis) *For all parameter $\delta > 0$, the sequence of rates of the codes built in Section 4.2.2 satisfy*

$$R_1^{(n)} \xrightarrow{n \rightarrow \infty} H(X|Y) + \delta, \quad R_2^{(n)} \xrightarrow{n \rightarrow \infty} I(X; Y). \quad (4.34)$$

The proof can be found in Appendix A.1.

4.3 Main results for the zero-error source coding problem for a Gray-Wyner network when side information may be present at the decoder

We first give an inner bound on the feasible rate region in Theorem 4.3.1, based on the following coding strategy. For all realization (x^n, y^n) of the source, an auxiliary sequence w^n is used to capture the common information between the source sequences, and the sets of possible sequences x^n and y^n are partitioned into cosets. As the side information is available at some decoders, random binning is done so that less information is transmitted on w^n through the common channel. Decoding algorithms are based on V -shells of side information sequences, and the zero-error property is ensured by an error bit that is accompanied by all source sequences if set to 1. We also give an outer bound on the feasible rate region in Theorem 4.3.2, and we show in Section 4.3.1 that the inner and outer bound coincide in several cases of interest.

Theorem 4.3.1 (Inner bound) *The rate tuples $(R_0, R_{x_1}, R_{x_2}, R_y)$ that satisfy*

$$R_0 \geq I(U, V, X, Y; W) - \min(I(U; W), I(V; W)), \quad (4.35)$$

$$R_{x_1} \geq H(X|U, W), \quad (4.36)$$

$$R_{x_2} \geq I(X; U|W) + \min(I(U; W), I(V; W)), \quad (4.37)$$

$$R_y \geq H(Y|V, W), \quad (4.38)$$

for some finite set \mathcal{W} and distribution $P_{U,V,W,X,Y} = P_{U,V,X,Y}P_{W|U,V,X,Y}$ are achievable.

Theorem 4.3.2 (Outer bound) *The rate tuples $(R_0, R_{x_1}, R_{x_2}, R_y)$ that are achievable must satisfy*

$$R_0 \geq I(X, Y; W|U, V), \quad (4.39)$$

$$R_{x_1} \geq H(X|U, W), \quad (4.40)$$

$$R_y \geq H(Y|V, W), \quad (4.41)$$

$$R_0 + R_{x_1} + R_{x_2} \geq H(X), \quad (4.42)$$

for some finite set \mathcal{W} and distribution $P_{U,V,W,X,Y} = P_{U,V,X,Y}P_{W|U,V,X,Y}$.

The proofs of Theorem 4.3.1 and Theorem 4.3.2 are respectively given in Appendix A.2 and in Appendix A.3. These proofs are based on the following Lemma 4.3.3 and Lemma 4.3.5. Lemma 4.3.3 is a covering/packing lemma for type classes. It is different from the packing lemma in [20, Lemma 10.1], as the latter states the existence of a family of codewords that satisfy the packing property, instead of drawing them at random following a given distribution.

Lemma 4.3.3 (Covering/Packing lemma in type classes) *Let $k \in \mathbb{N}^*$,*

$P_{A,B} \in \Delta_k(\mathcal{A} \times \mathcal{B})$, $P_{\bar{A}} \in \Delta(\mathcal{A})$, and

$$R_c \doteq I(A; B) + D(P_A \| P_{\bar{A}}). \quad (4.43)$$

Let $R > 0$ such that $R \neq R_c$. For all $n \in k\mathbb{N}^$, let $\mathcal{C}^{(n)} \doteq (\bar{A}_{[1]}^n, \dots, \bar{A}_{[2^{nR}]}^n)$ be a codebook of random sequences, drawn with a joint distribution that satisfies the marginal condition $\bar{A}_{[i]}^n \sim P_{\bar{A}}^{\otimes n}$ for all $i \in \{1, \dots, 2^{nR}\}$.*

4.3. Main results for the zero-error source coding problem for a Gray-Wyner network when side information may be present at the decoder

- If $R < R_c$, then we have for all $b^n \in \mathcal{T}_0^n(P_B)$:

$$\mathbb{P}\left(\exists a^n \in \mathcal{C}^{(n)}, T_{a^n, b^n} = P_{A,B}\right) \leq 2^{-n(R_c - R) + o(n)}. \quad (4.44)$$

- If $R > R_c$, assuming that the sequences in $\mathcal{C}^{(n)}$ are iid, we have for all $b^n \in \mathcal{T}_0^n(P_B)$:

$$\begin{aligned} \mathbb{P}\left(\exists a^n \in \mathcal{C}^{(n)}, \text{ s.t. } T_{a^n, b^n} = P_{A,B}\right) \\ = 1 - \exp\left[-2^{n(R - R_c) + o(n)}\right]. \end{aligned} \quad (4.45)$$

The proof of Lemma 4.3.3 can be found in Appendix A.4.

Remark 4.3.4 For $P_{\bar{A}} = P_A$, we have $R_c = I(A; B)$. For $P_{\bar{A}} = \text{Unif}(\mathcal{X})$, we have $R_c = \log |\mathcal{A}| - H(A|B)$.

Lemma 4.3.5 (Coset partition) Assume that $|\mathcal{A}|$ is prime. Let $k \in \mathbb{N}$, $P_{A,B} \in \Delta_k(\mathcal{A} \times \mathcal{B})$, $R_c = \log |\mathcal{A}| - H(A|B)$ and $R < R_c$ such that $kR / \log |\mathcal{A}| \in \mathbb{N}$. For all $n \in k\mathbb{N}$, let $\mathbf{G}^{(n)} \in \mathcal{M}_{n, nR / \log |\mathcal{A}|}(\mathcal{A})$ be a generator matrix whose entries are iid random variables drawn with the distribution $\text{Unif}(\mathcal{A})$. Then we have for all $(a^n, b^n) \in \mathcal{T}_0^n(P_{A,B})$:

$$\mathbb{P}\left(\exists \tilde{a}^n \in \text{Im } \mathbf{G}^{(n)} + a^n \setminus \{a^n\}, T_{\tilde{a}^n, b^n} = P_{A,B}\right) \leq 2^{-n(R_c - R) + o(n)}. \quad (4.46)$$

Proof. For all a^n , we have by construction that the coset $\text{Ker } \mathbf{H}^{(n)} + a^n \setminus \{a^n\}$ is formed of $|\mathcal{A}|^{nR / \log |\mathcal{A}|} - 1 = 2^{nR} - 1$ random codewords, which are pairwise independent and identically distributed: they follow the distribution $\text{Unif}(\mathcal{A})^{\otimes n}$. The result follows from Lemma 4.3.3. \square

4.3.1 Comparison with previous results

In this section, we derive the zero-error achievable rate region for several special cases, in which our inner bound is optimal.

First, consider the zero-error variant of Gray-Wyner problem [30] by setting U and V constant, and removing the decoder x_2 . Our inner bound allows to derive the zero-error rate achievable region for this problem:

$$R_0 \geq I(X, Y; W), \quad R_{x_1} \geq H(X|W), \quad (4.47)$$

$$R_y \geq H(Y|W), \quad R_{x_2} \geq 0, \quad (4.48)$$

for some distribution $P_{U,V,W,X,Y} = P_{U,V,X,Y}P_{W|U,V,X,Y}$. Due to the converse from [30, Theorem 4], this bound is optimal.

Another scheme of interest is when Y and V are constant, which gives the problem presented in Section 4.1.1. We obtain the following

$$\begin{array}{ll} \text{Inner bound:} & \text{Outer bound:} \\ R_0 \geq I(X; W|U), & R_0 \geq I(X; W|U), \end{array} \quad (4.49)$$

$$R_{x_1} \geq H(X|U, W), \quad R_{x_1} \geq H(X|U, W), \quad (4.50)$$

$$R_{x_2} \geq I(U; X, W), \quad R_0 + R_{x_1} + R_{x_2} \geq H(X), \quad (4.51)$$

$$R_y \geq 0, \quad R_y \geq 0, \quad (4.52)$$

for some distribution $P_{U,V,W,X,Y} = P_{U,V,X,Y}P_{W|U,V,X,Y}$. By choosing W constant (resp. $W = X$) we retrieve the achievability of the tuple $(R_0, R_{x_1}, R_{x_2}, R_y) = (0, H(X|U), I(U; X), 0)$ (resp. $(H(X|U), 0, I(U; X), 0)$). It proves that the bound is optimal as the outer bound gives $R_0 + R_{x_1} \geq H(X|U)$ and $R_0 + R_{x_1} + R_{x_2} \geq H(X)$.

We obtain the problem of Timo et al. [68] by removing the side information at the encoder, and removing the decoder x_2 . Then the possible distributions $P_{U,V,W,X,Y}$ must satisfy the Markov chain $(U, V) \rightarrow (X, Y) \rightarrow W$, and we obtain the same inner bound as them:

$$R_0 \geq \max(I(X, Y; W|U), I(X, Y; W|V)), \quad (4.53)$$

$$R_{x_1} \geq H(X|U, W), \quad (4.54)$$

$$R_{x_2} \geq 0, \quad (4.55)$$

$$R_y \geq H(Y|V, W), \quad (4.56)$$

for some distribution $P_{U,V,W,X,Y} = P_{U,V,X,Y}P_{W|X,Y}$.

BIBLIOGRAPHY

- [1] Rudolf Ahlswede, Ning Cai, and Zhen Zhang, « Zero-error capacity for models with memory and the enlightened dictator channel », *in: IEEE Transactions on Information Theory* 44.3 (1998), pp. 1250–1252.
- [2] Emrah Akyol et al., « On scalable coding in the presence of decoder side information », *in: 2014 IEEE International Symposium on Information Theory (ISIT)*.
- [3] Noga Alon, « On the capacity of digraphs », *in: European Journal of Combinatorics* 19.1 (1998), pp. 1–5.
- [4] Noga Alon, « The Shannon capacity of a union », *in: Combinatorica* 18.3 (1998), pp. 301–310.
- [5] Noga Alon and Alon Orlitsky, « Source coding and graph entropies », *in: IEEE Transactions on Information Theory* 42.5 (1996), pp. 1329–1339.
- [6] Sourya Basu, Daewon Seo, and Lav R Varshney, « Hypergraph-based Coding Schemes for Two Source Coding Problems under Maximal Distortion », *in: IEEE International Symposium on Information Theory (ISIT)*, 2020.
- [7] Meryem Benammar and Abdellatif Zaidi, « Rate-Distortion Region of a Gray–Wyner Model with Side Information », *in: Entropy* 20.1 (2018), p. 2.
- [8] Navid Mahmoudian Bidgoli, Thomas Maugey, and Aline Roumy, « Fine granularity access in interactive compression of 360-degree images based on rate-adaptive channel codes », *in: IEEE Transactions on Multimedia* 23 (2020), pp. 2868–2882.
- [9] Holger Boche and Christian Deppe, « Computability of the zero-error capacity of noisy channels », *in: 2021 IEEE Information Theory Workshop (ITW)*, IEEE, 2021, pp. 1–6.
- [10] Boris Bukh and Christopher Cox, « On a fractional version of Haemers’ bound », *in: IEEE Transactions on Information Theory* 65.6 (2018), pp. 3340–3348.
- [11] Peter J Cameron, « 6-transitive graphs », *in: Journal of Combinatorial Theory, Series B* 28.2 (1980), pp. 168–179.

- [12] Qi Cao et al., « On zero-error capacity of binary channels with one memory », *in: IEEE Transactions on Information Theory* 64.10 (2018), pp. 6771–6778.
- [13] Jun Chen, Da-Ke He, and Ashish Jagmohan, « The equivalence between Slepian–Wolf coding and channel coding under density evolution », *in: IEEE Transactions on Communications* 57.9 (2009), pp. 2534–2540.
- [14] Jun Chen et al., « On the linear codebook-level duality between Slepian–Wolf coding and channel coding », *in: IEEE Transactions on Information Theory* 55.12 (2009), pp. 5575–5590.
- [15] Maria Chudnovsky and Paul D Seymour, « The structure of claw-free graphs. », *in: BCC*, 2005, pp. 153–171.
- [16] Maria Chudnovsky et al., « The strong perfect graph theorem », *in: Annals of mathematics* (2006), pp. 51–229.
- [17] Gérard Cohen, Emanuela Fachini, and János Körner, « Zero-error capacity of binary channels with memory », *in: arXiv preprint arXiv:1402.0349* (2014).
- [18] Thomas M Cover, *Elements of information theory*, John Wiley & Sons, 1999.
- [19] Imre Csiszar, « Linear codes for sources and source networks: Error exponents, universal coding », *in: IEEE Transactions on Information Theory* 28.4 (1982), pp. 585–592.
- [20] Imre Csiszár and János Körner, *Information theory: coding theorems for discrete memoryless systems*, Cambridge University Press, 2011.
- [21] Imre Csiszár and János Körner, « On the capacity of the arbitrarily varying channel for maximum probability of error », *in: Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 57.1 (1981), pp. 87–101.
- [22] Imre Csiszár et al., « Entropy splitting for antiblocking corners and perfect graphs », *in: Combinatorica* 10.1 (1990), pp. 27–40.
- [23] Lingyu Duan et al., « Video coding for machines: A paradigm of collaborative compression and intelligent analytics », *in: IEEE Transactions on Image Processing* 29 (2020), pp. 8680–8695.
- [24] Soheil Feizi and Muriel Médard, « On network functional compression », *in: IEEE transactions on information theory* 60.9 (2014), pp. 5387–5401.

- [25] Michael Fekete, « Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten », *in: Mathematische Zeitschrift* 17.1 (1923), pp. 228–249.
- [26] Li Gao, Sander Gribling, and Yinan Li, « On a tracial version of Haemers bound », *in: IEEE Transactions on Information Theory* (2022).
- [27] Wen Gao et al., « Recent standard development activities on video coding for machines », *in: arXiv preprint arXiv:2105.12653* (2021).
- [28] L Gargano, J Körner, and U Vaccaro, « Sperner theorems on directed graphs and qualitative independence », *in: J. Combin. Theory Ser. A* 61 (1992), pp. 173–192.
- [29] Luisa Gargano, János Körner, and Ugo Vaccaro, « Capacities: from information theory to extremal set theory », *in: Journal of Combinatorial Theory, Series A* 68.2 (1994), pp. 296–316.
- [30] RM Gray and AD Wyner, « Source coding for a simple network », *in: Bell System Technical Journal* 53.9 (1974), pp. 1681–1721.
- [31] Yujie Gu and Ofer Shayevitz, « On the non-adaptive zero-error capacity of the discrete memoryless two-way channel », *in: Entropy* 23.11 (2021), p. 1518.
- [32] Xuan Guang et al., « Improved upper bound on the network function computing capacity », *in: IEEE Transactions on Information Theory* 65.6 (2019), pp. 3790–3811.
- [33] Willem Haemers et al., « On some problems of Lovász concerning the Shannon capacity of a graph », *in: IEEE Transactions on Information Theory* 25.2 (1979), pp. 231–232.
- [34] David A Huffman, « A method for the construction of minimum-redundancy codes », *in: Proceedings of the IRE* 40.9 (1952), pp. 1098–1101.
- [35] Amiram H. Kaspi, « Rate-distortion function when side-information may be present at the decoder », *in: IEEE Transactions on Information Theory* 40.6 (1994), pp. 2031–2034.
- [36] Peter Keevash and Eoin Long, « On the normalized Shannon capacity of a union », *in: Combinatorics, Probability and Computing* 25.5 (2016), pp. 766–767.
- [37] J Korner and Giuseppe Longo, « Two-step encoding for finite sources », *in: IEEE Transactions on Information Theory* 19.6 (1973), pp. 778–782.

- [38] Janos Korner and Alon Orlitsky, « Zero-error information theory », *in: IEEE Transactions on Information Theory* 44.6 (1998), pp. 2207–2229.
- [39] János Körner, « Coding of an information source having ambiguous alphabet and the entropy of graphs », *in: 6th Prague conference on information theory*, 1973, pp. 411–425.
- [40] János Körner, Concetta Pilotto, and Gábor Simonyi, « Local chromatic number and Sperner capacity », *in: Journal of Combinatorial Theory, Series B* 95.1 (2005), pp. 101–117.
- [41] Prashant Koulgi et al., « On zero-error source coding with decoder side information », *in: IEEE Transactions on Information Theory* 49.1 (2003), pp. 99–111.
- [42] Dinesh Krithivasan and S Sandeep Pradhan, « Distributed source coding using abelian group codes: A new achievable rate-distortion region », *in: IEEE Transactions on Information Theory* 57.3 (2011).
- [43] Thomas Laich and Michèle Wigger, « Utility of encoder side information for the lossless kaspi/heegard-berger problem », *in: 2013 IEEE International Symposium on Information Theory*, IEEE, 2013, pp. 3065–3069.
- [44] László Lovász, « On the Shannon capacity of a graph », *in: IEEE Transactions on Information theory* 25.1 (1979), pp. 1–7.
- [45] Rick Ma and Samuel Cheng, « Zero-error Slepian–Wolf coding of confined-correlated sources with deviation symmetry », *in: IEEE Transactions on Information Theory* 59.12 (2013), pp. 8195–8209.
- [46] Derya Malak and Muriel Médard, « Hyper Binning for Distributed Function Coding », *in: 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, 2020, pp. 1–5.
- [47] Katalin Marton, « On the Shannon capacity of probabilistic graphs », *in: Journal of Combinatorial Theory, Series B* 57.2 (1993), pp. 183–195.
- [48] K Ashik Mathew and Patric RJ Östergård, « New lower bounds for the Shannon capacity of odd cycles », *in: Designs, Codes and Cryptography* 84 (2017), pp. 13–22.
- [49] Thomas Maugey et al., « Incremental coding for extractable compression in the context of massive random access », *in: IEEE transactions on Signal and Information Processing over Networks* 6 (2020), pp. 251–260.

- [50] Alon Orlitsky and James R Roche, « Coding for computing », *in: Proceedings of IEEE 36th Annual Foundations of Computer Science*, IEEE, 1995, pp. 502–511.
- [51] Sven C Polak and Alexander Schrijver, « New lower bound on the Shannon capacity of C_7 from circular graphs », *in: Information Processing Letters* 143 (2019), pp. 37–40.
- [52] Jithin Ravi and Bikash Kumar Dey, « Function Computation Through a Bidirectional Relay », *in: IEEE Transactions on Information Theory* 65.2 (2018), pp. 902–916.
- [53] Aline Roumy and Thomas Maugey, « Universal lossless coding with random user access: the cost of interactivity », *in: 2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2015, pp. 1870–1874.
- [54] Igal Sason, « Observations on the Lovász θ -Function, Graph Capacity, Eigenvalues, and Strong Products », *in: Entropy* 25.1 (2023), p. 104.
- [55] Alexander Schrijver, « On the Shannon capacity of sums and products of graphs », *in: Indagationes Mathematicae* 34.1 (2023), pp. 37–41.
- [56] Milad Sefidgaran and Aslan Tchamkerten, « Distributed function computation over a rooted directed tree », *in: IEEE Transactions on Information Theory* 62.12 (2016), pp. 7135–7152.
- [57] Andrea Sgarro, « Source coding with side information at several decoders », *in: IEEE Transactions on Information Theory* 23.2 (1977), pp. 179–182.
- [58] Claude Shannon, « The zero error capacity of a noisy channel », *in: IRE Transactions on Information Theory* 2.3 (1956), pp. 8–19.
- [59] Claude Elwood Shannon, « A mathematical theory of communication », *in: The Bell system technical journal* 27.3 (1948), pp. 379–423.
- [60] Ofer Shayevitz, « Distributed computing and the graph entropy region », *in: IEEE transactions on information theory* 60.6 (2014), pp. 3435–3449.
- [61] Gábor Simonyi, « Graph entropy: a survey », *in: Combinatorial Optimization* 20 (1995), pp. 399–441.
- [62] Gábor Simonyi, « On Witsenhausen’s zero-error rate for multiple sources », *in: IEEE Transactions on Information Theory* 49.12 (2003), pp. 3258–3260.

- [63] Gábor Simonyi, « Perfect graphs and graph entropy. An updated survey », *in: Perfect graphs* (2001), pp. 293–328.
- [64] David Slepian and Jack Wolf, « Noiseless coding of correlated information sources », *in: IEEE Transactions on information Theory* 19.4 (1973), pp. 471–480.
- [65] Yossef Steinberg and Neri Merhav, « On successive refinement for the Wyner–Ziv problem », *in: IEEE Transactions on Information Theory* 50.8 (2004), pp. 1636–1654.
- [66] Chao Tian and Suhas N Diggavi, « Side-information scalable source coding », *in: IEEE Transactions on Information Theory* 54.12 (2008), pp. 5591–5608.
- [67] Roy Timo, Terence Chan, and Alex Grant, « Rate distortion with side-information at many decoders », *in: IEEE Transactions on Information Theory* 57.8 (2011), pp. 5240–5257.
- [68] Roy Timo et al., « Source coding for a simple network with receiver side information », *in: 2008 IEEE International Symposium on Information Theory*, IEEE, 2008, pp. 2307–2311.
- [69] Hugo Touchette, « Legendre-Fenchel transforms in a nutshell », *in: URL <http://www.maths.qmul.ac.uk/~ht/archive/lfth2.pdf>* (2005).
- [70] Ertem Tuncel and Kenneth Rose, « On variable-length coding of sources with side information at multiple decoders », *in: IEEE International Symposium on Information Theory, 2003. Proceedings.* IEEE, 2003, p. 140.
- [71] Ertem Tuncel et al., « On complementary graph entropy », *in: IEEE transactions on information theory* 55.6 (2009), pp. 2537–2546.
- [72] Aleksander Vesel and Janez Žerovnik, « Improved lower bound on the Shannon capacity of C_7 », *in: Information processing letters* 81.5 (2002), pp. 277–282.
- [73] Lele Wang and Young-Han Kim, « Linear code duality between channel coding and Slepian-Wolf coding », *in: 2015 53rd IEEE Annual Allerton Conference on Communication, Control, and Computing (Allerton)*.
- [74] Moritz Wiese et al., « Secure estimation and zero-error secrecy capacity », *in: IEEE Transactions on Automatic Control* 64.3 (2018), pp. 1047–1062.

- [75] Hans Witsenhausen, « The zero-error side information problem and chromatic numbers (corresp.) », *in: IEEE Transactions on Information Theory* 22.5 (1976), pp. 592–593.
- [76] Hirosuke Yamamoto, « Wyner-ziv theory for a general function of the correlated sources (corresp.) », *in: IEEE Transactions on Information Theory* 28.5 (1982), pp. 803–807.
- [77] Guofen Zhang et al., « The zero-error capacity of binary channels with 2-memories », *in: Advances in Mathematics of Communications* (2022).
- [78] Jeroen Zuiddam et al., *Algebraic complexity, asymptotic spectra and entanglement polytopes*, Institute for Logic, Language and Computation, 2018.

Titre : Théorie de l'information zéro-erreur dans les réseaux : graphes, codage pour le calcul de fonctions et dualité source-canal

Mot clés : Théorie de l'information zéro-erreur, codage de source, dualité source-canal

Résumé : Cette thèse de doctorat porte sur la théorie de l'information zéro-erreur, notamment sur le codage de source avec information adjacente, le codage de canal et la dualité source-canal. Ces travaux s'articulent autour de trois axes : 1. les problèmes de codage zéro-erreur basés sur la correction, 2. les problèmes de codage zéro-erreur basés sur les graphes de confusion, 3. les problèmes de codage zéro-erreur pour le calcul d'une fonction. Une contribution importante de ce travail concerne le problème de la linéarisation du débit optimal, lorsque l'encodeur traite plusieurs tâches conjointement. Nous démontrons l'équivalence de la linéarisation pour les produits de graphes et pour leurs unions disjointes. Cette observation permet de caractériser les débits optimaux pour une classe de problèmes irrésolus depuis les années 1950. Les contributions de cette thèse ont donné lieu à des publications dans les actes des meilleurs conférences internationales de théorie de l'information et deux articles de revues sont en cours de préparation.

Title: Zero-error network information theory: graphs, coding for computing and source-channel duality

Keywords: Zero-error Information Theory, Source coding, Source-channel duality

Abstract: This doctoral thesis focuses on zero-error information theory, particularly on source coding with side information, channel coding and source-channel duality. These works revolve around three axes: 1. zero-error coding problems based on correction, 2. zero-error coding problems based on confusion graphs, 3. zero-error coding problems for calculating a function. An important contribution of this work concerns the problem of linearizing the optimal bitrate, when the encoder processes several tasks jointly. We demonstrate the equivalence of linearization for graph products and for their disjoint unions. This observation makes it possible to characterize the optimal rates for a class of unsolved problems since the 1950s. The contributions of this thesis have given rise to publications in the proceedings of the best international Information Theory conferences and two journal articles are in preparation.