



**HAL**  
open science

## Reconstruction en grandes dimensions

David Salinas

► **To cite this version:**

David Salinas. Reconstruction en grandes dimensions. Autre. Université de Grenoble, 2013. Français.  
NNT : 2013GRENT036 . tel-00947303

**HAL Id: tel-00947303**

**<https://theses.hal.science/tel-00947303>**

Submitted on 20 Feb 2014

**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.

## THÈSE

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE**

Spécialité : **signal, image, parole, télécommunications (SIPT)**

Arrêté ministériel : 7 août 2006

Présentée par

**David Salinas**

Thèse dirigée par **Dominique Attali**

préparée au sein du **Gipsa-lab**

dans l'école doctorale **Electrotechnique Automatique**

**Traitement du Signal (EEATS)**

## Reconstruction en grandes dimensions

Thèse soutenue publiquement le 11 septembre 2013,  
devant le jury composé de :

**M. Jean-Daniel BOISSONNAT**

INRIA Sophia-Antipolis, Rapporteur

**M. Bruno LEVY**

INRIA Nancy, Rapporteur

**Mme Dominique ATTALI**

Gipsa-Lab, Directeur de thèse

**M. Michel DESVIGNES**

Gipsa-Lab, Examineur

**M. André LIEUTIER**

Dassault-Systèmes, Examineur

**Mme Annick MONTANVERT**

Gipsa-Lab, Examineur

**M. Steve OUDOT**

INRIA Saclay, Examineur

**M. Uli WAGNER**

IST Austria, Examineur





UNIVERSITÉ DE GRENOBLE  
ÉCOLE DOCTORALE EEATS  
Electrotechnique Automatique Traitement du Signal

# THÈSE

pour obtenir le titre de

**docteur en sciences**

de l'Université de Grenoble

Présentée et soutenue par

David Salinas

**Reconstruction en grandes dimensions**

Thèse dirigée par Dominique Attali

préparée au laboratoire Gipsa-lab

**Jury :**

<i>Directeur :</i>	Mme Dominique ATTALI	- Gipsa-Lab
<i>Rapporteurs :</i>	M. Jean-Daniel BOISSONNAT	- INRIA Sophia-Antipolis
	M. Bruno LEVY	- INRIA Nancy
<i>Examineurs :</i>	M. André LIEUTIER	- Dassault-Systèmes
	M. Michel DESVIGNES	- Gipsa-Lab
	Mme Annick MONTANVERT	- Gipsa-Lab
	M. Steve OUDOT	- INRIA Saclay
	M. Uli WAGNER	- IST Austria



# Remerciements

Je tiens à remercier mes encadrants et coauteurs Dominique Attali et André Lieutier qui m'ont introduit avec patience et passion au domaine fascinant qu'est la géométrie en grande dimension.

Je remercie également Jean-Daniel Boissonnat et Bruno Levy d'avoir accepté de rapporter mon manuscrit. Je leur suis très reconnaissant d'avoir pris ce temps pour évaluer mes travaux et je les remercie pour toutes leurs indications.

Merci aux autres membres du jury Michel Desvignes, Annick Montanvert, Steve Oudot et Uli Wagner pour leurs questions et leurs suggestions qui m'ont offert un éclairage nouveau sur mon travail.

Je souhaitais faire une thèse depuis la classe préparatoire mais mon parcours universitaire m'a grandement aidé. Je suis particulièrement reconnaissant envers l'ENS Lyon ainsi qu'au Master associé de Nice dirigé par Enrico Formenti. J'ai eu la chance de bénéficier de cours de très grande qualité qui se sont avérés très précieux dans mes recherches.

Je remercie tous mes proches qui sont venus à ma soutenance, parfois de contrées lointaines. Vous voir tous fut pour moi une grande émotion et restera un très beau souvenir.

Merci à toute ma famille de m'avoir tant soutenu. Je remercie tout particulièrement mes grand-parents Rose, René et Marie, Jean-Claude et Jennifer, Bernard et Mamichou<sup>1</sup>. Ma passion pour la science vient incontestablement de leur héritage, transmis avec beaucoup d'affection et de pédagogie. Cette liste semble très incomplète sans Josette et André que je considère comme mes grand-parents « adoptifs » : merci à vous pour tout ce temps passé dans mon enfance à m'éveiller à des choses formidables comme le go ou la physique.

Merci à mes amis. Ceux de Grenoble de longue date avec qui j'ai adoré jouer au foot ou perdre au colon de Catane dans la bonne humeur. Merci aussi à tous ceux du Gipsa-Lab : Farouk, Florent, Benjamin, Laurent, François et Cyrille. En particulier, je remercie Damien qui m'a fourni, en plus du modèle  $\text{\LaTeX}$  de sa thèse, une aide technique et une amitié quotidienne précieuse. Je remercie également tous mes colocataires et voisins de Brossolette avec qui j'ai passé deux très belles années.

Enfin, je remercie Chloé pour son attention, son soutien et sa générosité. J'espère pouvoir l'aider pour sa rédaction comme elle m'a aidé pour la mienne, ce qui n'est pas peu dire.

---

1. Le hasard des choses les a voulu nombreux mais abondance de bien ne nuit pas.



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Préliminaires</b>	<b>7</b>
<b>2 Type d'homotopie du complexe de Rips</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.2 Mesures de défaut de convexité . . . . .	15
2.2.1 Définitions et premières propriétés . . . . .	15
2.2.2 Caractérisation des valeurs critiques de la fonction de distance . . . . .	17
2.2.3 Stabilité . . . . .	20
2.3 Du complexe de Čech au complexe de Rips . . . . .	21
2.3.1 Quasi-complexes de Rips et énoncé des résultats . . . . .	21
2.3.2 Preuve des théorèmes 2.2 et 2.3 . . . . .	22
2.3.3 Optimalité de la condition du théorème 2.3 . . . . .	24
2.4 Reconstruction de forme . . . . .	26
2.4.1 Conditions d'échantillonnage . . . . .	26
2.4.2 Liens avec la fonction critique . . . . .	28
2.4.3 Reconstruction de formes à $\mu$ -reach positif . . . . .	29
2.4.4 Reconstruction avec le complexe de Rips « intrinsèque » . . . . .	33
<b>3 Structure de données complexes simpliciaux</b>	<b>35</b>
3.1 Introduction . . . . .	35
3.2 Structure de données squelette/bloqueurs . . . . .	37
3.3 Taille du complexe de Rips . . . . .	39
3.3.1 Bornes théoriques . . . . .	39



3.3.2	Evaluation expérimentale . . . . .	41
3.4	Opérations élémentaires sur la structure de données squelette/bloqueurs . . . . .	41
3.4.1	Algorithmes de complexité polynomiale en la taille du graphe et le nombre de bloqueurs . . . . .	42
3.4.2	Algorithmes de conversion entre structures de données . . . . .	48
3.5	Conclusion et perspectives . . . . .	49
<b>4</b>	<b>Simplification d'un complexe simplicial par contraction d'arête</b>	<b>51</b>
4.1	Introduction . . . . .	51
4.2	Une condition sur le <i>link</i> pour qu'une contraction d'arête préserve le type d'homotopie . . . . .	52
4.2.1	Contraction d'arête . . . . .	52
4.2.2	Condition du <i>link</i> . . . . .	53
4.2.3	Tester la condition du <i>link</i> efficacement . . . . .	56
4.2.4	Réciproque de la condition du <i>link</i> . . . . .	57
4.3	Simplification par contraction d'arêtes . . . . .	58
4.3.1	Un algorithme . . . . .	58
4.3.2	Priorité sur les arêtes . . . . .	59
4.4	Résultats expérimentaux . . . . .	62
4.4.1	Jeu de données . . . . .	62
4.4.2	Taille de la structure de données au cours des contractions . . . . .	63
4.4.3	Taille des complexes après contraction sur le jeu de données . . . . .	64
<b>5</b>	<b>Simplification de <i>flag</i>-complexes par collapses</b>	<b>69</b>
5.1	Introduction . . . . .	69
5.2	Collapser un complexe de Rips dont les sommets approximent un convexe . . . . .	71
5.2.1	En passant par le complexe de Čech . . . . .	72
5.2.2	Préserver un <i>flag</i> -complexe durant la simplification . . . . .	74

5.2.3	Augmenter le paramètre d'échelle . . . . .	76
5.2.4	Algorithmes pour trouver une séquence de réduction. . . . .	77
5.3	Simplifier un complexe de Rips dont les sommets échantillonnent une variété . .	78
5.3.1	Extraire un sous-ensemble de points sans changer le type d'homotopie .	78
5.3.2	Algorithme de simplification. . . . .	81
5.3.3	Preuve de correction de l'algorithme de reconstruction dans un cas simple	82
5.4	Résultats expérimentaux . . . . .	84
5.4.1	Comparaison des différentes stratégies pour REDUCTIBLE . . . . .	84
5.4.2	Résultats des simplifications sur le jeu de données . . . . .	85
<b>Conclusion</b>		<b>87</b>
<b>Bibliographie</b>		<b>89</b>
<b>Index</b>		<b>95</b>
<b>A Lemmes supplémentaires pour le chapitre 2</b>		<b>101</b>
A.1	Stabilité de la plus petite boule englobante . . . . .	101
A.2	Stabilité des hypothèses du théorème 2.3 . . . . .	102
A.3	Preuve du lemme 2.6 . . . . .	103
<b>B Optimalité de l'hypothèse du théorème 2.3</b>		<b>105</b>
<b>C Constantes pour la reconstruction de formes à <i>reach</i> positif</b>		<b>107</b>
<b>D Pseudo-codes</b>		<b>109</b>
D.1	Calcul du <i>link</i> d'un simplexe . . . . .	109
D.2	Tester l'appartenance d'un simplexe au complexe . . . . .	109
D.3	Calculer les bloqueurs créés par une contraction d'arête . . . . .	111
D.4	Tester si un complexe est un cône . . . . .	111

---

D.5	Tester la condition du <i>link</i> . . . . .	111
D.6	Enumérer les simplexes du $k$ -squelette . . . . .	112
D.7	Pseudo-codes pour réduire un complexe en un point . . . . .	114
<b>E</b>	<b>Lemmes géométriques du chapitre 5</b>	<b>117</b>

# Introduction

L'ère de la numérisation engendre un flux de données sans précédent. Un enjeu scientifique et technologique majeur est d'arriver à extraire de l'information de données toujours plus nombreuses et complexes. Bien souvent, les données peuvent être représentées sous forme de points. Malheureusement, les points ont souvent une dimension élevée ce qui ôte plusieurs outils classiques aux informaticiens cherchant à les analyser en raison du fléau de la dimension. Une bonne nouvelle est que ces points sont souvent situés aux alentours d'une variété de dimension bien inférieure à la dimension des points. En effet, le nombre de degrés de liberté du procédé ayant engendré les points est souvent largement inférieur à la dimension des points générés. Par exemple, de tels points peuvent être une collection d'images d'un objet prises en tournant autour d'un axe. La dimension des points est alors le nombre de pixels des images tandis que la dimension de la variété échantillonnée est de un (puisqu'elle est paramétrée uniquement par l'angle de la pose). Dans ce cas, comme dans beaucoup d'applications, la variété échantillonnée est hautement non linéaire.

Le problème qui se pose alors est de pouvoir extraire des informations de la variété, voire même de pouvoir construire une approximation de celle-ci, à partir d'un ensemble de points l'échantillonnant<sup>2</sup>. Au vu de ces nombreuses applications, beaucoup d'auteurs se sont penchés sur la reconstruction de variétés. Plusieurs approches permettent de reconstruire la variété comme Isomap [TSL00], Locally-linear embedding [RS00] ou encore Laplacian eigenmaps [BN01]. Ces méthodes demandent souvent de calculer la décomposition en valeurs et vecteurs propres d'une matrice de distance obtenue avec les points. Cette décomposition permet ensuite de construire une approximation de la variété échantillonnée. Toutes ces méthodes présentent néanmoins l'inconvénient d'être rarement pourvues d'analyse théorique certifiant leur bon comportement. En particulier, aucune d'entre elles ne garantit que l'approximation construite possède la même topologie que la variété échantillonnée.

L'utilisation de la topologie pour l'analyse de données a connu un essor important ces dernières années dans la recherche académique et dans l'ingénierie<sup>3</sup>. Récemment, de nombreux progrès ont été accomplis pour extraire de l'information d'un nuage de points échantillonnant une variété. Des méthodes ont par exemple été proposées pour calculer efficacement la dimension d'une variété [GW03] ou encore l'ensemble de ses groupes d'homologie, un invariant topologique [CO08]. Des travaux récents ont également permis de résoudre le problème de *reconstruction topologique* sous certaines conditions. Ce problème consiste à reconstruire une approximation ayant la même topologie que la variété tout en respectant les deux contraintes :

1. Trouver des conditions géométriques sur la densité du nuage de points et la courbure de

---

2. Ce problème est souvent intitulé réduction de dimension non-linéaire ou *non-linear dimensionality reduction* en anglais.

3. La startup Ayasdi a récemment levé 10 millions de dollars afin de développer des méthodes fondées sur la topologie et la géométrie pour l'analyse et la visualisation de données complexes. Un de ses cofondateurs est Gunnar Carlsson, auteur d'une revue de l'état de l'art sur l'utilisation de la topologie pour l'extraction d'information dans les données complexes [Car08].

la variété assurant que l'approximation construite ait la même topologie que la variété échantillonnée.

2. Construire l'approximation efficacement. En particulier, on souhaite que le temps de reconstruction ne dépende pas exponentiellement de la dimension des points.

Une manière de représenter une telle approximation est d'encoder un *complexe simplicial*. Il s'agit d'un ensemble de simplexes ou de faces. Cette notion ainsi que des notions usuelles de topologie seront définies dans le premier chapitre. Notons enfin que par "même topologie", on entend que l'approximation doit être *homéomorphe* à la variété ou avoir son *type d'homotopie* (ce qui est une condition plus faible). Ces deux notions seront également définies dans le premier chapitre.

## Conditions sur l'échantillonnage et sur la variété

Plusieurs hypothèses plus ou moins restrictives sont souvent supposées pour la reconstruction topologique. Concernant l'échantillonnage, l'hypothèse la plus restrictive est de supposer que les points ne sont pas bruités et proches de la variété. Cette hypothèse est souvent relâchée en supposant seulement que les points sont à une distance de Hausdorff bornée de la variété. Néanmoins cette hypothèse est encore restrictive puisqu'elle ne tolère ni les points aberrants ni les échantillonnages de densité non uniforme. Plusieurs conditions d'échantillonnage ont été proposées permettant de pallier ces deux problèmes. Afin de tolérer une densité de points non uniforme, plusieurs auteurs proposent d'imposer à ce que l'échantillonnage soit localement aussi dense qu'une mesure locale de la courbure de la variété, appelée *local feature size* [AB99, CDR05, BG10]. Enfin, la distance à la mesure permet de spécifier des conditions d'échantillonnage et des méthodes de reconstruction robustes à la présence de points aberrants [GMM11]. Quant à la variété échantillonnée, l'hypothèse la plus simple est de supposer que son *reach* est strictement positif. Le *reach* d'une variété (ou portée en français) est la plus petite distance entre son *axe médian* et elle-même; l'axe médian est l'ensemble des points ayant au moins deux plus proches points sur la variété. Cette condition sur le *reach* exclut les variétés "anguleuses" possédant par exemple des arêtes vives puisque cette mesure s'annule dans ce cas. Des auteurs ont introduit le  $\mu$ -*reach* qui permet de prendre en compte une classe plus large de variétés, éventuellement non lisses [CCSL09].

## Algorithmes de reconstruction topologique

Les premières applications de la reconstruction topologique se situaient principalement en l'informatique graphique, en informatique médicale, en ingénierie inverse ou encore en cartographie [AB99]. Une application typique est la modélisation d'un solide. Etant donné un ensemble de points de la surface d'un solide obtenus grâce à un scanner 3D, on cherche à modéliser la surface. Ainsi, la plupart des algorithmes de reconstruction topologique considéraient initialement des points de dimension 3 et utilisaient la triangulation de Delaunay pour construire une approximation homéomorphe comme le montre la revue de l'état de l'art [CG06]. Des travaux ont ensuite permis des résultats similaires pour des points de dimension

arbitraire en utilisant par exemple l'offset des points [NSW08] ou encore la triangulation de Delaunay [CDR05]. Malheureusement, construire la triangulation de Delaunay est difficile en raison de sa complexité élevée en grandes dimensions [AAD07]. Les méthodes nécessitant de construire l'offset des points souffrent du même écueil. En effet, pour construire celui-ci, il faut soit construire un complexe d' $\alpha$ -*shape* nécessitant la construction d'une triangulation de Delaunay, soit construire un complexe de Čech dont la complexité rend également difficile toute construction efficace [DI12]. Les complexes de témoins, introduits par Silva [Sil], ont permis une approche différente pour construire une approximation homéomorphe à la variété mais celle-ci est toujours exponentielle en la dimension des points [BGO07]. La seule méthode efficace (i.e. polynomiale avec la dimension des points) pour construire une approximation homéomorphe connue à ce jour a été introduite par Boissonnat et Ghosh [BG10]. Elle calcule la triangulation de Delaunay en calculant une triangulation locale de celle-ci sur les plans tangents et en la recollant ensuite globalement. Néanmoins la preuve suppose que les points ne soient pas bruités contrairement à d'autres approches précédemment citées. Toutes les méthodes que nous venons d'évoquer considèrent des variétés à *reach* strictement positif interdisant les arêtes vives ou les points anguleux. Des travaux ont permis de donner un algorithme de reconstruction pour des variétés éventuellement non lisses dès lors que leur  $\mu$ -*reach* est positif [CCSL09]. Néanmoins cette approche impose de construire l'union des boules qui est, comme nous l'avons vu précédemment, difficile à calculer en grande dimension.

## Contributions

Le fil conducteur de cette thèse est l'analyse et l'utilisation de *flag-complexes* pour la reconstruction topologique. Les *flag-complexes* sont des complexes qui jouissent de la propriété d'être entièrement définis par le graphe de leurs sommets et arêtes. En effet, pour ces complexes, chaque simplexe correspond à une clique dans ce graphe. Un *flag-complexe* retiendra particulièrement notre attention : le complexe de Rips. Ce dernier a été utilisé en topologie algorithmique ainsi que pour l'analyse de données [CCS03, CS03, DSC04, dSG07]. Plusieurs raisons expliquent la popularité de ce complexe simplicial. La première est bien évidemment liée à sa nature de *flag-complexe*. Cette propriété permet un stockage efficace [Zom10a] et l'extraction efficace d'informations topologiques [CO08, Zom10b]. La seconde raison tient à la proximité du complexe de Rips à l'union des boules centrées en les points. Nous verrons que cette proximité peut s'exprimer à l'aide d'un entrelacement avec le complexe de Čech. Néanmoins, malgré cette proximité, le complexe de Rips n'a pas en général la même topologie que le complexe de Čech. Nous donnons des conditions géométriques assurant que le complexe de Rips partage le même type d'homotopie que le complexe de Čech. Nous dérivons alors de ce résultat des conditions d'échantillonnage sous lesquelles le complexe de Rips possède le même type d'homotopie qu'une variété. Les conditions sur la variété prennent en compte des formes à  $\mu$ -*reach* positif pour  $\mu$  suffisamment grand. Ainsi, le complexe de Rips permet de répondre au problème de reconstruction topologique, pour des variétés éventuellement non lisses, à condition de pouvoir représenter efficacement le complexe de Rips. Dans ce but, nous introduisons ensuite une structure de données permettant d'encoder le complexe de Rips de manière compacte. Celle-ci permet également de représenter tout complexe simplicial mais

elle est particulièrement adaptée aux complexes proches de *flag-complexes*. Cette structure nous permet ensuite d'expérimenter et analyser plusieurs opérations de simplifications de complexes simpliciaux. En particulier, nous étudions les *contractions d'arêtes* et les *collapses de simplexes*. La contraction d'arête est une opération utilisée initialement en informatique graphique permettant la simplification de maillage et l'opération de collapse est une opération classique de topologie algébrique.

Nos principales contributions sont les suivantes, à chacune d'entre elles est dédié un chapitre de ce document :

- Une condition géométrique permettant d'assurer que le complexe de Rips a le même type d'homotopie que le complexe de Čech voire même, en renforçant la condition, qu'une variété à  $\mu$ -reach positif.
- Une structure de données pour complexes simpliciaux particulièrement adaptée aux complexes proches de *flag-complexes*.
- Une condition permettant d'assurer qu'une contraction d'arête préserve le type d'homotopie ainsi qu'un algorithme de simplification par contraction d'arêtes.
- Une condition géométrique permettant d'assurer qu'on peut réduire un complexe de Rips en un point par collapses et un algorithme de simplification par collapse permettant de réduire la dimension d'un *flag-complexe*.

## Chapitre 1

Ce chapitre préliminaire introduit les principales définitions géométriques et topologiques qui sont utilisées tout au long de cette thèse.

## Chapitre 2

Dans ce chapitre, nous introduisons deux mesures de *feature size*, toutes deux appelées *défaut de convexité*. Elles mesurent à quel point un objet diffère localement d'un convexe, de la même manière que la courbure mesure à quel point un objet diffère localement d'un plan. Ces mesures nous permettent de formuler des conditions d'échantillonnage pour le complexe de Čech et pour le complexe de Rips. Dans un premier temps, ces conditions nous permettent de donner des conditions sous lesquelles les complexes de Čech et de Rips partagent le même type d'homotopie. On donne également des conditions d'échantillonnage qui garantissent que ces complexes ont le même type d'homotopie que la variété échantillonnée, en supposant que celle-ci ait un  $\mu$ -reach positif pour  $\mu$  suffisamment grand.

## Chapitre 3

Nous introduisons une structure de données pour complexes simpliciaux. Elle encode un complexe dans une paire que l'on nomme *squelette/bloqueurs*. La première composante de cette paire est encodée dans un graphe et représente le 1-*squelette* du complexe (i.e. son

ensemble de sommets et d'arêtes). La deuxième composante est appelée bloqueurs. Elle permet de stocker les différences avec un *flag*-complexe. Intuitivement, les bloqueurs représentent les non-faces minimales du complexe. Nous montrons comment réaliser les principales opérations simpliciales dans notre structure. Lorsque le nombre de bloqueurs est constant, la plupart de ces opérations peuvent être réalisées efficacement (i.e. en temps polynomial en la taille du 1-squelette du complexe).

## Chapitre 4

Une condition locale permet d'assurer que la contraction d'une arête préserve l'homéomorphisme pour des complexes homéomorphes à des variétés de dimension 2 ou 3 [DEGN98]. Nous montrons que cette condition permet également de garantir qu'une contraction d'arête préserve le type d'homotopie pour tout complexe simplicial. Cette condition peut être encodée efficacement dans la structure de données squelette/bloqueur ce qui nous permet d'introduire un algorithme de simplification par contractions d'arêtes. Lorsqu'on contracte des arêtes d'un complexe, la propriété de *flag*-complexe peut être perdue. Cela arrive lorsqu'une contraction d'arête crée un ou plusieurs bloqueurs. Avec la structure de données du chapitre 2, nous étudions la contraction d'arête sur des complexes de Rips qui sont initialement des *flag*-complexes.

## Chapitre 5

Nous introduisons une opération appelée *collapse généralisé* qui étend la notion de collapse classique. Cette opération consiste à retirer d'un complexe une face et toutes les faces la contenant lorsque le *link* de cette face (i.e. son voisinage) a la topologie d'un point. Le problème qui se pose est alors de savoir quand est-ce qu'un complexe a la topologie d'un point. Sous certaines hypothèses de convexité, nous montrons qu'on peut trouver une séquence de réduction d'un complexe de Čech ou de Rips vers un point. Cela nous permet d'introduire un algorithme de simplification de complexes simpliciaux par collapses généralisés. Cet algorithme se prête particulièrement à la simplification de *flag*-complexes puisque cette propriété est maintenue durant la simplification. En implémentant cet algorithme dans la structure de données du chapitre 2, nous constatons son utilité pour réduire la dimension de complexes de Rips.

## Publications

Les résultats présentés dans ce mémoire ont fait l'objet des publications suivantes, co-écrites avec Dominique Attali et André Lieutier :

- Dominique Attali, André Lieutier, and David Salinas. Collapsing rips complexes. *29th European Workshop on Computational Geometry (EuroCG)*, 2013
- Dominique Attali, André Lieutier, and David Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. *Int. J. Comput. Geometry Appl.*, 22(4):279–304, 2012



- Dominique Attali, André Lieutier, and David Salinas. Vietoris-Rips Complexes also Provide Topologically Correct Reconstructions of Sampled Shapes. *Computational Geometry*, 46(4, special issue):448–465, 2012
- Dominique Attali, André Lieutier, and David Salinas. Vietoris-rips complexes also provide topologically correct reconstructions of sampled shapes. In *Proceedings of the 27th annual ACM symposium on Computational geometry*, pages 491–500, 2011
- Dominique Attali, André Lieutier, and David Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. In *Proceedings of the 27th annual ACM symposium on Computational geometry*, SoCG '11, pages 501–509, 2011

# Préliminaires

Dans ce chapitre, nous rappelons plusieurs définitions et propriétés classiques qui seront utilisées tout au long de cette thèse. La seule définition non-classique de ce chapitre est la définition de l'*expansion* qui interviendra principalement dans les chapitres 3 et 4. Cette partie se termine avec la définition des complexes de Čech et de Rips qui seront utilisés intensément par la suite.

**Espace métrique, distances et plus petite boule englobante.** Tout au long de cette thèse, nous considérerons des sous-ensembles d'un espace euclidien  $\mathbb{R}^D$  avec  $D \geq 1$ . La distance euclidienne entre deux points  $x$  et  $y$  de  $\mathbb{R}^D$  est notée  $\|x - y\|$ . Etant donnés deux sous-ensembles  $X$  et  $Y$  de  $\mathbb{R}^D$ , on note  $d_H(Y | X) = \sup_{y \in Y} d(y, X)$  pour la distance partielle de Hausdorff de  $Y$  à  $X$ , où  $d(y, X)$  est l'infimum des distances euclidiennes entre  $y$  et les points  $x$  de  $X$ . Observons que  $d_H(Y | X) \leq \varepsilon$  si et seulement si  $Y$  est contenu dans le  $\varepsilon$ -*offset*  $X^\varepsilon = \{y \in \mathbb{R}^D \mid d(y, X) \leq \varepsilon\}$ . La *distance de Hausdorff* entre  $X$  et  $Y$  est  $d_H(X, Y) = \max\{d_H(X | Y), d_H(Y | X)\}$ . Rappelons que le *diamètre* d'un sous-ensemble  $\sigma$  de  $\mathbb{R}^D$  est le supremum des distances entre les paires de points dans  $\sigma$ , que nous notons  $\text{Diam } \sigma = \sup_{p, q \in \sigma} \|p - q\|$ . Un sous-ensemble  $\sigma$  est dit *borné* si son diamètre est fini.

La boule fermée de centre  $z$  et de rayon  $r$  est notée  $B(z, r)$ . Les boules seront toujours supposées fermées sauf si le contraire est précisé. Il est connu que la plus petite boule englobante d'un sous-ensemble non-vide et borné de  $\mathbb{R}^D$  est bien définie. Nous notons son centre  $\text{Center}(\sigma)$  et son rayon  $\text{Rad}(\sigma)$ . En notant  $\text{Hull}(X)$  l'enveloppe convexe d'un sous-ensemble  $X \subset \mathbb{R}^D$  et  $\text{Cl}(X)$  pour la fermeture de  $X$ , il est facile de montrer (par contradiction) que  $\text{Center}(\sigma) \in \text{Cl}(\text{Hull}(\sigma))$ . Les inégalités suivantes seront utiles pour comparer les définitions des complexes de Čech et de Rips :

$$\frac{1}{2} \text{Diam}(\sigma) \leq \text{Rad}(\sigma) \leq \frac{\vartheta_D}{2} \text{Diam}(\sigma) \quad \text{où } \vartheta_D = \sqrt{\frac{2D}{D+1}}. \quad (1.1)$$

L'inégalité de droite est aussi connue sous le nom de théorème de Jung et une preuve courte peut être trouvée dans [dSG07]. En particulier, nous avons  $\text{Rad}(\sigma) < \frac{1}{\sqrt{2}} \text{Diam}(\sigma)$  pour toute dimension  $D$ .

**Complexe simplicial abstrait.** Lorsque l'on cherche à représenter une approximation d'une forme, on considère souvent un complexe simplicial. Il s'agit d'un ensemble de simplexes (*i.e.* un ensemble de sommets, d'arêtes, de triangles, de tétraèdres, ...) vérifiant une

propriété de "complétion" : dès lors qu'un simplexe  $\sigma$  appartient au complexe simplicial, tous les simplexes inclus dans  $\sigma$  doivent également appartenir au complexe. Par exemple, si un triangle appartient à un complexe simplicial, le complexe doit aussi contenir les 3 points et les 3 arêtes inclus dans le triangle. Nous rappelons la définition de complexe simplicial ainsi que plusieurs définitions de topologie classique qui lui sont généralement associées [Mun93].

Soit  $P$  un ensemble fini de points de  $\mathbb{R}^D$ . Un sous-ensemble non vide  $\sigma \subset P$  est appelé un *simplexe abstrait*. Sa dimension est son nombre d'éléments moins un. Un  $i$ -simplexe est un simplexe abstrait de dimension  $i$ . Si  $\tau \subset \sigma$  est sous-ensemble non vide,  $\tau$  est une *face* de  $\sigma$  et  $\sigma$  est une *coface* de  $\tau$ . Un *complexe simplicial abstrait*  $K$  est une collection non vide de simplexes abstraits qui contient, en plus des simplexes abstraits, toutes les faces de ces derniers. L'ensemble des sommets de  $K$  est l'union de ses éléments,  $\text{Vert}(K) = \bigcup_{\sigma \in K} \sigma$ . Un sous-complexe de  $K$  est un complexe simplicial  $L \subset K$ . Un exemple particulier de sous-complexe de  $K$  est le  $i$ -squelette constitué de tous les simplexes de dimension inférieure ou égale à  $i$ , que l'on dénote  $K^{(i)}$ .

Comme le suggère son appellation, un complexe simplicial abstrait ne prend pas en compte la géométrie des points (on aurait pu prendre  $P$  un ensemble fini de points de n'importe quel ensemble). Les deux prochaines définitions permettent de donner une interprétation géométrique au complexe simplicial. Le *shadow* de  $K$  est le sous-ensemble de  $\mathbb{R}^D$  couvert par l'enveloppe convexe des simplexes de  $K$ ,  $\text{Shd } K = \bigcup_{\sigma \in K} \text{Hull}(\sigma)$ , à ne pas confondre avec  $|K|$ , l'espace sous-jacent d'une réalisation géométrique de  $K$ . Soit  $n$  le nombre de sommets de  $K$ . L'espace sous-jacent  $|K|$  de  $K$  peut être défini (à un homéomorphisme près) en considérant une fonction  $f : \text{Vert}(K) \rightarrow \mathbb{R}^{n-1}$  qui envoie les  $n$  sommets de  $K$  sur  $n$  points affinement indépendants dans  $\mathbb{R}^{n-1}$  et en posant  $|K| = \bigcup_{\sigma \in K} \text{Hull}(f(\sigma))$ . Les deux espaces  $|K|$  et  $\text{Shd } K$  ne sont pas homéomorphes en général comme le montre la figure 1.2.

Toutes les définitions de cette partie sont classiques sauf la définition de l'*expansion* d'un complexe que l'on donne maintenant.

**Définition 1.1.** *Soit  $K$  un complexe simplicial de dimension  $k$ . L'expansion de  $K$ , notée  $\text{Expand}(K)$ , est le plus grand complexe simplicial ayant  $K$  comme  $k$ -squelette.*

Durant cette thèse, une attention particulière sera apportée à l'étude des *flag-complexes* (aussi appelés complexes de cliques). Un *flag-complexe* est un complexe simplicial qui est l'expansion d'un graphe  $G$  et on le dénote alors  $\text{Flag}(G)$ . Ce type de complexe possède l'avantage de pouvoir être défini implicitement par la donnée d'un graphe puisque tout simplexe du complexe est représenté par une clique dans son graphe  $G$ . Ainsi, ce type de complexe peut être représenté avec une structure de données compacte, un graphe, plutôt que par l'ensemble de ces simplexes qui est souvent très volumineux.

Dans un complexe simplicial, lorsque l'on souhaite tester si une simplification préserve la topologie, on est souvent amené à tester une condition locale. Les notions de *Star* et de *Link* d'un simplexe permettent d'exprimer la notion de voisinage dans un complexe simplicial. Soit  $\sigma$  un simplexe d'un complexe simplicial  $K$ . Le *star* de  $\sigma$  dans  $K$ , noté  $\text{St}_K(\sigma)$ , est l'ensemble des simplexes de  $K$  ayant  $\sigma$  comme face. La fermeture de  $\text{St}_K(\sigma)$  est notée  $\overline{\text{St}_K(\sigma)}$ ; il s'agit

du plus petit complexe simplicial contenant  $\overline{\text{St}_K(\sigma)}$ . Le *link* de  $\sigma$  dans  $K$ , noté  $\text{Lk}_K(\sigma)$ , est l'ensemble des simplexes de  $K$  dans  $\overline{\text{St}_K(\sigma)}$  disjoints de  $\sigma$ . Nous aurons également besoin de l'opération de *join*. Etant donnés deux complexes simpliciaux  $K$  et  $L$ , le plus petit complexe simplicial contenant les simplexes de la forme  $\kappa \cup \lambda$  où  $\kappa \in K$  et  $\lambda \in L$  est appelé le join de  $K$  et  $L$  et il sera noté  $K \star L$ .

Dans cette thèse, nous chercherons à montrer qu'une approximation construite à partir d'un nuage de points échantillonnant une forme possède la même topologie que la forme en question. Nous définissons maintenant formellement ce que l'on entend par "même topologie".

**Homéomorphisme et type d'homotopie.** Deux espaces topologiques  $X$  et  $Y$  sont dits *homéomorphes* s'il existe un *homéomorphisme* de  $X$  vers  $Y$ . Un homéomorphisme est une bijection  $f : X \rightarrow Y$  telle que  $f$  et  $f^{-1}$  soient continues. On utilise la notation  $X \approx Y$  pour dénoter le fait que  $X$  soit homéomorphe à  $Y$ . Intuitivement,  $X$  et  $Y$  sont homéomorphes si on peut déformer continuellement  $X$  en  $Y$  sans écraser d'épaisseur. Par exemple, sur la figure 1.1, les deux premiers espaces sont homéomorphes alors que les deux derniers ne le sont pas.

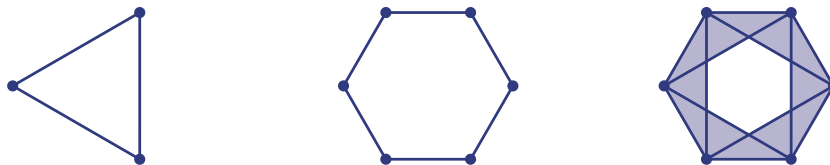


FIGURE 1.1 – Les deux premiers espaces topologiques sont homéomorphes contrairement aux deux derniers. Les trois espaces ont le même type d'homotopie.

Une relation plus faible entre deux espaces topologiques est la relation d'*équivalence d'homotopie*. Cette relation est plus faible car deux espaces homéomorphes ont toujours le même type d'homotopie. Pour définir le type d'homotopie, nous avons besoin de plusieurs définitions. Etant donnés  $X$  et  $Y$  deux espaces topologiques, ainsi que  $f, g : X \rightarrow Y$  deux fonctions continues de  $X$  dans  $Y$ , on dit que  $f$  est homotope à  $g$  s'il existe une fonction  $H : X \times [0, 1] \rightarrow Y$  continue telle que :

$$\forall x \in X, H(x, 0) = f(x) \text{ et } H(x, 1) = g(x)$$

Deux espaces topologiques  $X$  et  $Y$  ont le même *type d'homotopie* s'il existe deux fonctions continues  $f : X \rightarrow Y$  et  $g : X \rightarrow Y$  telles que  $f \circ g$  et  $g \circ f$  soient homotopes à l'identité. Dans ce cas, on note  $X \simeq Y$ . Intuitivement, l'équivalence d'homotopie entre deux espaces  $X$  et  $Y$  signifie qu'on peut déformer  $X$  dans  $Y$  en écrasant éventuellement de l'épaisseur ce qui est interdit par l'homéomorphisme. Par exemple, tous les espaces de la figure 1.1 ont le même type d'homotopie. Un complexe simplicial ayant le type d'homotopie d'un point est dit *contractile*.

Nous rappelons maintenant deux manières de construire un complexe simplicial abstrait étant donné un ensemble de points dans  $\mathbb{R}^D$  et un paramètre d'échelle  $t \geq 0$ . Elles sont sujettes à quelques variations d'un auteur à un autre dans la littérature.

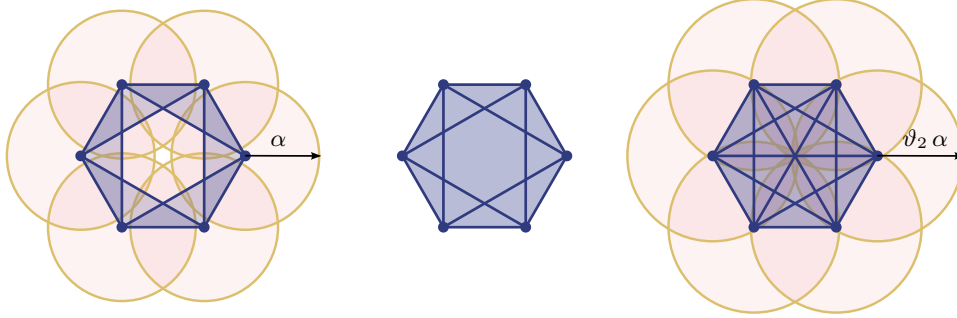


FIGURE 1.2 – A gauche : le complexe de Čech avec le paramètre  $\alpha$  comprend six triangles et a le type d'homotopie d'un cercle. Au milieu : le complexe de Rips de paramètre  $\alpha$  contient deux triangles en plus et est homéomorphe à une 2-sphère. Son shadow est un disque topologique. A droite : le complexe de Čech de paramètre  $\vartheta_2 \alpha$  contient toutes les faces d'un 5-simplexe et est homéomorphe à une 5-boule.

**Complexe de Čech.** Le *complexe de Čech*,  $\text{Cech}(P, \alpha)$ , est le complexe simplicial abstrait dont les  $k$ -simplexes correspondent aux sous-ensembles de  $k + 1$  points pouvant être contenus dans une boule de rayon  $\alpha$ ,

$$\text{Cech}(P, \alpha) = \{\sigma \mid \emptyset \neq \sigma \subset P, \text{Rad}(\sigma) \leq \alpha\}.$$

De manière équivalente, un  $k$ -simplexe  $\{p_0, \dots, p_k\}$  appartient au complexe de Čech si et seulement si les  $k + 1$  boules euclidiennes  $B(p_i, \alpha)$  ont une intersection commune non-vide. Notons  $\text{Nrv } F = \{G \subset F \mid \bigcap G \neq \emptyset\}$  le *nerf* de la collection  $F$ . Le complexe de Čech est le nerf de la collection de boules  $\{B(p, t) \mid p \in P\}$ . Puisque les intersections des boules sont convexes, le théorème du nerf [Bjo95, ES97] implique que le complexe  $\text{Cech}(P, \alpha)$  a le même type d'homotopie que l'union des boules, c'est-à-dire,  $|\text{Cech}(P, \alpha)| \simeq P^\alpha = \bigcup_{p \in P} B(p, \alpha)$ .

**Le complexe de Rips.** Le *complexe de Vietoris-Rips* est une variante du complexe de Čech qui est plus facile à calculer. Le complexe de Vietoris-Rips,  $\text{Rips}(P, \alpha)$ , est le complexe simplicial abstrait dont les  $k$ -simplexes correspondent aux sous-ensembles de  $k + 1$  points de  $P$  dont le diamètre est au plus  $2\alpha$ ,

$$\text{Rips}(P, \alpha) = \{\sigma \mid \emptyset \neq \sigma \subset P, \text{Diam } \sigma \leq 2\alpha\}.$$

Par souci de concision,  $\text{Rips}(P, \alpha)$  sera appelé simplement complexe de Rips par la suite. Rappelons que le *flag-complexe* d'un graphe  $G$ , dénoté par  $\text{Flag}(G)$ , est le complexe simplicial maximal dont le 1-squelette est  $G$ . Le complexe de Rips est un exemple de *flag-complexe*. Plus précisément, ce complexe est le plus grand complexe simplicial partageant avec le complexe de Čech le même 1-squelette,  $\text{Rips}(P, \alpha) = \text{Flag}((\text{Cech}(P, \alpha))^{(1)})$ . En général,  $\text{Rips}(P, \alpha)$  et

$\text{Cech}(P, \alpha)$  ne partagent pas la même topologie comme le montre la figure 1.2. Le complexe  $\text{Rips}(P, \alpha)$  ne partage donc pas toujours le même type d'homotopie que le  $\alpha$ -offset  $P^\alpha$ . Néanmoins, les complexes de Čech et de Rips sont assez proches puisque la chaîne d'inclusion suivante, découlant de l'équation (1.1), est vérifiée [dSG07] :

$$\text{Cech}(P, \alpha) \subset \text{Rips}(P, \alpha) \subset \text{Cech}(P, \vartheta_D \alpha) \quad \text{où } \vartheta_D = \sqrt{\frac{2D}{D+1}}. \quad (1.2)$$

Le but du prochain chapitre sera d'établir une condition sur l'ensemble de points  $P$  sous laquelle  $|\text{Rips}(P, \alpha)| \simeq |\text{Cech}(P, \alpha)|$  et donc  $|\text{Rips}(P, \alpha)| \simeq P^\alpha$ . Pour cela on montrera que l'on peut simplifier, sous certaines conditions, le complexe  $\text{Rips}(P, \alpha)$  vers le complexe  $\text{Cech}(P, \alpha)$  grâce à une série de *collapses*.

**Collapses.** L'opération de collapse permet de simplifier un complexe simplicial. Il s'agit d'une opération classique en topologie algorithmique et en topologie algébrique. Nous donnons la définition de collapse élémentaire et nous étendons ensuite cette définition au collapse classique, au collapse étendu et au collapse généralisé. Chacune de ces opérations préserve le type d'homotopie et consiste à retirer le *star* d'un simplexe lorsqu'une certaine condition sur son *link* est vérifiée.

Un *collapse élémentaire* est l'opération qui enlève une paire de simplexes  $(\sigma_{\min}, \sigma_{\max})$  lorsque  $\sigma_{\max}$  est l'unique coface propre de  $\sigma_{\min}$  [DEG99]. De manière équivalente, un collapse élémentaire est l'opération qui ôte  $\text{St}_K(\sigma_{\min})$  à  $K$  lorsque  $\text{Lk}_K(\sigma_{\min})$  est un point. Le résultat de cette opération est un complexe simplicial  $K \setminus \text{St}_K(\sigma_{\min}) = K \setminus \{\sigma_{\min}, \sigma_{\max}\}$  qui se déforme en se rétractant sur  $K$ . L'opération inverse qui ajoute  $\text{St}_K(\sigma_{\min})$  à  $K$  est un *anti-collapse élémentaire* et il s'agit également d'une opération préservant le type d'homotopie. Un complexe simplicial est dit *collapsible* s'il peut être réduit à un sommet par une séquence finie de collapses élémentaires. Par exemple, la fermeture d'un simplexe  $\sigma$ ,  $\text{Cl}(\sigma) = \bigcup_{\emptyset \neq \tau \subset \sigma} \{\tau\}$ , est collapsible.  $\text{Cl}(\sigma)$  est un exemple de cône. Un *cône* est un complexe simplicial  $K$  contenant un sommet  $o$  tel que l'implication suivante est vérifiée :  $\sigma \in K \implies \sigma \cup \{o\} \in K$ . Les cônes sont également collapsibles. Clairement, tout complexe collapsible est contractile<sup>1</sup>.

Nous listons maintenant plusieurs généralisations possibles de l'opération de collapse élémentaire. L'opération qui ôte  $\text{St}_K(\sigma_{\min})$  à  $K$  est appelée :

- un *collapse (classique)* si le *link* de  $\sigma_{\min}$  est la fermeture d'un simplexe [DEG99]; l'opération inverse est appelée un *anti-collapse (classique)*.
- un *collapse (étendu)* si le *link* de  $\sigma_{\min}$  est un cône; l'opération inverse est appelée un *anti-collapse (étendu)* [BM12a];
- un *collapse (généralisé)* si le *link* de  $\sigma_{\min}$  est contractile; l'opération inverse est appelée *anti-collapse (généralisé)*.

Les trois collapses (classiques, étendus, généralisés) préservent le type d'homotopie. Les deux premiers peuvent être exprimés comme une composition de collapses élémentaires (voir le lemme 1.1). Le dernier type de collapse peut lui être exprimé comme une séquence de collapses et d'anti-collapses élémentaires comme le montre la proposition 1.2.

1. La réciproque n'est pas vraie. Un fameux contre-exemple est la maison de Bing à deux chambres [Hat02].

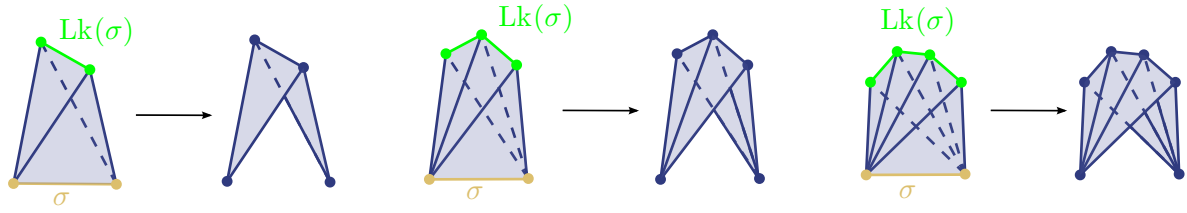


FIGURE 1.3 – Illustration d'un collapse classique (gauche), étendu (milieu) et généralisé (droite). Le *link* de  $\sigma$ , représenté en vert, est respectivement la fermeture d'un simplexe, un cône et un complexe contractile.

**Lemme 1.1.** *Soit  $K$  un complexe simplicial et soit  $\sigma$  un simplexe de  $K$ . Si le link de  $\sigma$  est collapsable, alors il existe une séquence de collapses élémentaires de  $K$  vers  $K \setminus \text{St}_K(\sigma)$ .*

*Démonstration.* Notons  $L$  le *link* de  $\sigma$  dans  $K$ . Comme  $L$  est collapsable, il peut être réduit à un point  $o$  par une séquence de collapses élémentaires. Nous déduisons de cette séquence de collapse une séquence de collapse qui réduit  $K$  vers  $K \setminus \text{St}_K(\sigma)$ . Pour chaque collapse élémentaire enlevant une paire  $(\alpha, \alpha \cup \beta)$  dans  $L$ , on lui associe le collapse enlevant la paire  $(\sigma \cup \alpha, \sigma \cup \alpha \cup \beta)$  dans  $K$ . Il s'agit bien d'un collapse puisque, comme  $\alpha$  a une unique coface propre  $\alpha \cup \beta$  dans  $L$ ,  $\sigma \cup \alpha$  a une unique coface propre  $\sigma \cup \alpha \cup \beta$  dans  $K$ . A la fin de de cette séquence de collapses élémentaires, le *link* de  $\sigma$  est réduit au point  $\{o\}$ . Après un dernier collapse ôtant la paire  $(\sigma, \sigma \cup \{o\})$  nous obtenons le complexe  $K \setminus \text{St}_K(\sigma)$ .  $\square$

En réadaptant la preuve du lemme précédent, il est facile de montrer que si le *link* de  $\sigma$  peut être réduit à un point par une séquence de collapses et d'anti-collapses élémentaires, alors  $K$  peut également être réduit par une séquence de collapses et d'anti-collapses élémentaires vers  $K \setminus \text{St}_K(\sigma)$ . En utilisant le lemme suivant dont une preuve peut être trouvée dans [Smi11], on obtient la proposition 1.2 qui permet de garantir que l'opération de collapse généralisé préserve le type d'homotopie.

**Proposition 1.1.** *Soit  $K$  un complexe simplicial.  $K$  est contractile si et seulement si il peut être réduit à un point par une séquence de collapses et d'anti-collapses élémentaires.*

**Proposition 1.2.** *Soit  $K$  un complexe simplicial et soit  $\sigma$  un simplexe de  $K$ . Si le link de  $\sigma$  est contractile, alors il existe une séquence de collapses/anti-collapses élémentaires de  $K$  vers  $K \setminus \text{St}_K(\sigma)$ .*

# Type d'homotopie du complexe de Rips

## Sommaire

<b>2.1</b>	<b>Introduction</b>	<b>13</b>
<b>2.2</b>	<b>Mesures de défaut de convexité</b>	<b>15</b>
2.2.1	Définitions et premières propriétés	15
2.2.2	Caractérisation des valeurs critiques de la fonction de distance	17
2.2.3	Stabilité	20
<b>2.3</b>	<b>Du complexe de Čech au complexe de Rips</b>	<b>21</b>
2.3.1	Quasi-complexes de Rips et énoncé des résultats	21
2.3.2	Preuve des théorèmes 2.2 et 2.3	22
2.3.3	Optimalité de la condition du théorème 2.3	24
<b>2.4</b>	<b>Reconstruction de forme</b>	<b>26</b>
2.4.1	Conditions d'échantillonnage	26
2.4.2	Liens avec la fonction critique	28
2.4.3	Reconstruction de formes à $\mu$ -reach positif	29
2.4.4	Reconstruction avec le complexe de Rips « intrinsèque »	33

## 2.1 Introduction

Le complexe de Rips est un exemple de *flag complexe* et bénéficie donc de la propriété que tout sous-ensemble de  $P$  appartient au complexe si et seulement si toutes ses arêtes appartiennent au complexe. Autrement dit, le complexe de Rips est complètement déterminé par le graphe de ses sommets et de ses arêtes. Cette forme de stockage compacte justifie l'intérêt pour un tel complexe, du moins en grandes dimensions. Des travaux récents ont porté sur la construction [Zom10a] ou la simplification de ce complexe [ALS12a, Zom10b]. Cependant, pour pouvoir utiliser le complexe de Rips, on voudrait pouvoir s'assurer que sa topologie reflète bien la topologie d'une forme échantillonnée avec les sommets du complexe. Un complexe simplicial similaire est le complexe de Čech. Le complexe de Čech de  $P$  à l'échelle  $\alpha$  a le même type d'homotopie que le  $\alpha$ -offset de  $P$  et permet donc de reproduire la topologie d'une forme échantillonnée par  $P$  [NSW08]. Cette propriété a été utilisée par Chazal et Oudot



[CO08] afin d'extraire des informations topologiques d'une forme à partir de la filtration du complexe de Rips, en l'emboîtant avec la filtration du complexe de Čech et en utilisant la persistance topologique. Cependant le complexe de Čech est difficile à calculer en raison de son nombre important de simplexes et de l'absence de représentation implicite compacte comme pour le complexe de Rips [DI12].

Dans ce chapitre, nous mettrons en évidence une relation plus directe entre le complexe de Rips et la forme échantillonnée. Nous donnons des conditions sous lesquelles le complexe de Rips capture la topologie de la forme. Dans un autre contexte, il a été prouvé dans [Hau95, Lat01] que le complexe de Rips d'un ensemble de points possède le même type d'homotopie qu'une forme échantillonnée pour peu qu'il soit suffisamment proche d'une variété Riemannienne pour la distance de Gromov-Hausdorff. Cependant, ces résultats concernent les variétés lisses, considèrent la métrique Riemannienne intrinsèque plutôt que la métrique euclidienne ambiante. Néanmoins, ils suggèrent qu'on puisse utiliser les complexes de Rips pour produire des approximations de formes ayant la bonne topologie. Lorsque la distance utilisée est la distance associée à la norme infinie  $\ell^\infty$ , le complexe de Rips des points  $P$  avec comme paramètre  $\alpha$  est identique au complexe de Čech de  $P$  à l'échelle  $\alpha$  et possède le même type d'homotopie que l'union des hypercubes de côté  $2\alpha$  centrés en chacun des points de  $P$  [CS03]. Dans ce cas, les auteurs de [AL10] donnent des conditions sous lesquelles le complexe de Rips de  $P$  reproduit le type d'homotopie de la forme échantillonnée par  $P$ . Dans ce chapitre, nous supposons que la distance utilisée est la distance euclidienne.

Il est nécessaire pour un nuage de points d'être suffisamment précis et dense afin de pouvoir refléter la topologie de la forme qu'il échantillonne. La précision de l'échantillonnage est souvent exprimée grâce à la distance de Hausdorff. Les résultats de reconstructions certifiés sont souvent de la forme suivante : si la distance de Hausdorff est plus petite qu'une certaine notion de caractéristique topologique de la forme, alors la sortie possède la même topologie que la forme. Les premières conditions d'échantillonnage ont été exprimées grâce au *reach* qui est l'infimum des distances entre les points de la forme et entre les points de son axe médian [Att98, AB99, BC02, NSW08, CL08]. Malheureusement, le *reach* s'annule dans le cas des formes possédant des arêtes vives concaves et ne permet donc pas d'exprimer des conditions d'échantillonnage pour des variétés non lisses ou des objets stratifiés. Pour palier à ce problème, les auteurs de [CCSL09] ont introduit une nouvelle caractéristique topologique, le  $\mu$ -*reach*, qui leur a permis de formuler des conditions d'échantillonnage pour reconstruire une large classe de compacts non-lisses.

Dans ce chapitre, nous introduisons deux fonctions, toutes deux appelées *défaut de convexité* qui mesurent à quel point un objet diffère localement d'un convexe. Nous utilisons ensuite ces mesures afin de formuler des conditions d'échantillonnage pour le complexe de Čech et pour le complexe de Rips. Ces conditions garantissent que ces complexes ont le même type d'homotopie que la forme échantillonnée, en supposant que celle-ci ait un  $\mu$ -*reach* positif pour  $\mu$  suffisamment grand.

*Ce chapitre reprend des résultats co-écrits avec Dominique Attali et André Lieutier et publiés dans les articles [ALS12b, ALS11b].*

## 2.2 Mesures de défaut de convexité

Dans cette section, nous introduisons deux fonctions qui, étant donné un sous-ensemble non vide et borné  $X \subset \mathbb{R}^D$ , offrent deux manières de mesurer le défaut de convexité de  $X$ . A partir de la première fonction, nous donnons dans la section 2.3 une condition qui garantit que le complexe de Rips d'un ensemble de points  $P$  de paramètre  $\alpha$  se rétracte par déformation sur le complexe de Čech des points  $P$  avec pour paramètre  $\alpha$ . La deuxième fonction permettra quant-à-elle de formuler dans la section 2.4 des conditions échantillonnage sous lesquelles le complexe de Čech et de Rips d'un ensemble de points  $P$  sont des reconstructions topologiquement correctes d'une forme  $X$  échantillonnée par les points de  $P$ .

### 2.2.1 Définitions et premières propriétés

Nous supposons dans cette section que le sous-ensemble  $X \subset \mathbb{R}^D$  est toujours non-vidé et borné. En particulier, tout sous-ensemble  $\sigma \subset X$  est aussi borné et sa plus petite boule englobante de centre  $\text{Center}(\sigma)$  et de rayon  $\text{Rad}(\sigma)$  est donc bien définie. Nous commençons par définir l'ensemble des centres de  $X$  à l'échelle  $t$  comme le sous-ensemble (voir la partie gauche de la figure 2.1 ainsi que la figure 2.4) :

$$\text{Centers}(X, t) = \bigcup_{\substack{\emptyset \neq \sigma \subset X \\ \text{Rad}(\sigma) \leq t}} \{\text{Center}(\sigma)\}.$$

Rappelons que l'enveloppe convexe de  $X$  est notée  $\text{Hull}(X)$ . Nous étendons maintenant la notion d'enveloppe convexe à une notion locale. Précisément, nous définissons l'enveloppe convexe de  $X$  à l'échelle  $t$  comme le sous-ensemble (voir la partie droite de la figure 2.1) :

$$\text{Hull}(X, t) = \bigcup_{\substack{\emptyset \neq \sigma \subset X \\ \text{Rad}(\sigma) \leq t}} \text{Hull}(\sigma).$$

Si  $X$  est compact alors  $\text{Hull}(X, t)$  contient  $\text{Centers}(X, t)$ . Lorsque  $P$  est fini alors  $\text{Hull}(P, t)$  est le *shadow* du complexe de Čech,  $\text{Cech}(P, t)$ .

**Définition 2.1** (Fonctions de défaut de convexité). *Étant donné un sous-ensemble  $X \subset \mathbb{R}^D$ , on associe à  $X$  deux fonctions à valeurs dans  $\mathbb{R}^+$ . La première  $c_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  est définie par  $c_X(t) = d_H(\text{Centers}(X, t) | X)$  et la seconde  $h_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  est définie par  $h_X(t) = d_H(\text{Hull}(X, t) | X)$ .*

Intuitivement,  $c_X$  et  $h_X$  peuvent être vues comme des fonctions mesurant le défaut de convexité de  $X$  à une certaine échelle. Plus précisément, notons que si  $X \subset \mathbb{R}^D$  est compact, alors les trois équivalences suivantes sont vérifiées :  $X$  convexe si et seulement si  $c_X = 0$  si et seulement si  $h_X = 0$ . Les deux fonctions de convexité  $c_X$  et  $h_X$  nous serviront de deux

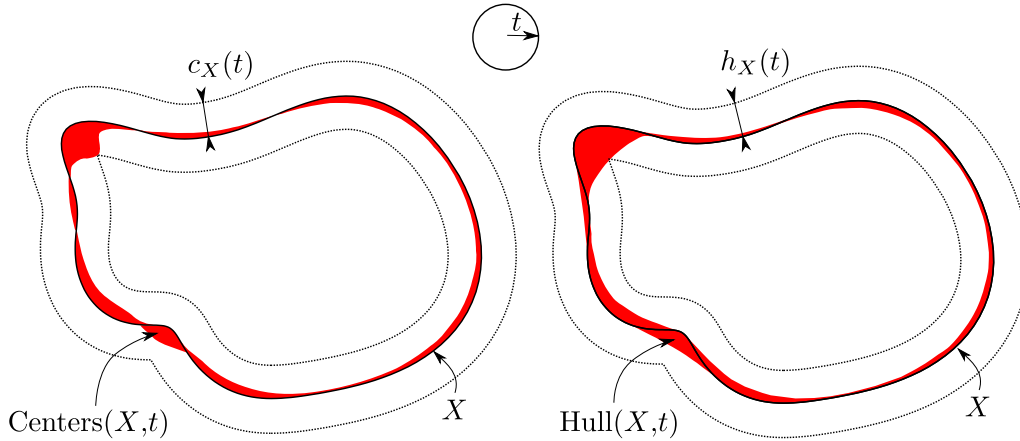


FIGURE 2.1 – A gauche, plus petit *offset* de  $X$  contenant  $\text{Centers}(X, t)$ . A droite, plus petit *offset* de  $X$  contenant  $\text{Hull}(X, t)$ .

manières. La fonction  $c_P$  nous permettra de préciser des conditions géométriques assurant que les complexes de Čech et de Rips partagent le même type d'homotopie tandis qu'on tirera profit de la fonction  $h_P$ , qui se révélera plus stable que  $c_P$ , pour exprimer des conditions d'échantillonnage dans des théorèmes de reconstruction dans la section 2.4.1. Le graphe de la fonction  $c_P$  est tracé pour plusieurs ensembles de points finis dans la figure 2.3.

Avant d'étudier plus en détail les fonctions  $c_X$  et  $h_X$  dans les deux prochaines sections, remarquons quelques propriétés simples. Puisque  $X$  est un sous-ensemble de  $\text{Centers}(X, t)$  et de  $\text{Hull}(X, t)$ , les deux distances de Hausdorff partielles  $d_H(X | \text{Centers}(X, t))$  et  $d_H(X | \text{Hull}(X, t))$  sont nulles. Ainsi, nous aurions également pu utiliser la distance de Hausdorff à la place de la distance de Hausdorff partielle dans notre définition. Les deux fonctions  $c_X$  et  $h_X$  sont nulles en 0, croissantes sur l'intervalle  $[0, \text{Rad}(X)]$  et constantes au-delà de  $\text{Rad}(X)$ . Puisque  $\text{Center}(\sigma) \in \text{Cl}(\text{Hull}(\sigma))$ , nous avons que  $c_X \leq h_X$ . On peut facilement vérifier que pour un sous-ensemble  $X \subset \mathbb{R}^D$  et deux réels positifs  $t$  et  $\alpha$ , les trois énoncés suivants sont équivalents : (1)  $h_X(t) \leq \alpha$ ; (2)  $\text{Hull}(X, t) \subset X^\alpha$ ; (3)  $[\text{Rad}(\sigma) \leq t \implies \text{Hull}(\sigma) \subset X^\alpha]$  pour tout  $\sigma \subset X$ . En particulier, on obtient que  $h_X(t) \leq t$  pour tout  $t \geq 0$  puisque  $\text{Rad}(\sigma) \leq t \implies \text{Hull}(\sigma) \subset \sigma^t$  (il s'agit d'une conséquence directe du lemme 2.1 ci-dessous appliqué pour  $x = y$ .)

**Lemme 2.1.** *Pour tout sous-ensemble non vide et borné  $\sigma \subset \mathbb{R}^D$ , tout point  $x \in \mathbb{R}^D$  et tout point  $y \in \text{Hull}(\sigma)$ , l'inégalité suivante est vérifiée :*

$$d(x, \sigma)^2 \leq \|x - y\|^2 + \text{Rad}(\sigma)^2 - \|y - \text{Center}(\sigma)\|^2$$

*Démonstration.* Le résultat est clair lorsque  $d(x, \sigma) \leq \|x - y\|$ . Supposons que  $d(x, \sigma) > \|x - y\|$ . Soit  $B_0$  la plus petite boule englobante de  $\sigma$  et soit  $B_1$  la plus grande boule centrée en  $x$  dont l'intérieur n'intersecte pas  $\sigma$ ; voir la figure 2.2. Par construction,  $\sigma \subset B_0 \setminus B_1$ . Rappelons que la puissance d'un point  $y$  à une boule  $B$  est  $\pi_B(y) = \|y - z\|^2 - r^2$ , où  $z$  est le centre de  $B$  et  $r$  son rayon. Soit  $H_{01}$  l'ensemble de points dont la puissance à  $B_0$  est inférieure à la puissance

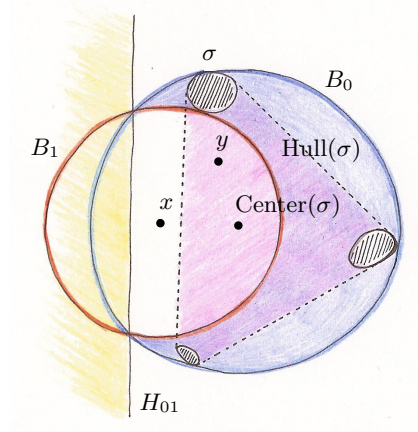


FIGURE 2.2 – Notations pour la preuve du lemme 2.1.

à  $B_1$ .  $H_{01}$  est un demi-espace fermé contenant l'ensemble  $B_0 \setminus B_1$ . En particulier, il contient  $\sigma$  et tout point  $y \in \text{Hull}(\sigma)$ . Ainsi,  $\pi_{B_0}(y) \leq \pi_{B_1}(y)$ , ce qui montre le résultat.  $\square$

### 2.2.2 Caractérisation des valeurs critiques de la fonction de distance

Nous avons vu dans la section précédente que  $c_X(t) \leq h_X(t) \leq t$  pour tout  $t$ . Le but de cette section est d'établir que l'égalité est atteinte si et seulement si  $t$  est une valeur critique de la fonction de distance à  $X$ . Cette propriété ne sera pas utilisée avant la section 2.4 mais permet de mieux comprendre le résultat de la section 2.3.

Définissons maintenant les points et les valeurs critiques de la fonction de distance. La fonction de distance  $d(\cdot, X)$  du compact  $X \subset \mathbb{R}^D$  associe à tout point  $y \in \mathbb{R}^D$  sa distance euclidienne à  $X$ ,  $d(y, X) = \min_{x \in X} \|x - y\|$ . Bien que la fonction de distance ne soit pas différentiable, on peut définir une notion de points critiques de manière analogue à la notion classique des fonctions différentiables comme le montre la figure 2.5. Précisément, Grove définit dans [Gro93, page 360] les points critiques de la fonction de distance d'un sous-ensemble fermé d'une variété Riemannienne. Nous rappelons ici cette définition dans notre contexte. Soit  $\Gamma_X(y) = \{x \in X \mid d(y, X) = \|x - y\|\}$  l'ensemble des points de  $X$  qui sont les points les plus proches de  $y$ .

**Définition 2.2.** *On dit que  $y \in \mathbb{R}^D$  est un point critique de la fonction de distance  $d(\cdot, X)$  si  $y \in \text{Hull}(\Gamma_X(y))$ . Les valeurs critiques de  $d(\cdot, X)$  sont les images par  $d(\cdot, X)$  des points critiques.*

Le lemme suivant fournit deux caractérisations des valeurs critiques de la fonction de distance d'un compact  $X \subset \mathbb{R}^D$ , fondées respectivement sur les fonctions de défaut de convexité  $c_X$  et  $h_X$ .

**Lemme 2.2.** *Pour tout compact  $X \subset \mathbb{R}^D$  et tout réel  $t > 0$ , les trois énoncés suivants sont équivalents : (1)  $t$  est une valeur critique de  $d(\cdot, X)$  ; (2)  $c_X(t) = t$  ; (3)  $h_X(t) = t$ .*

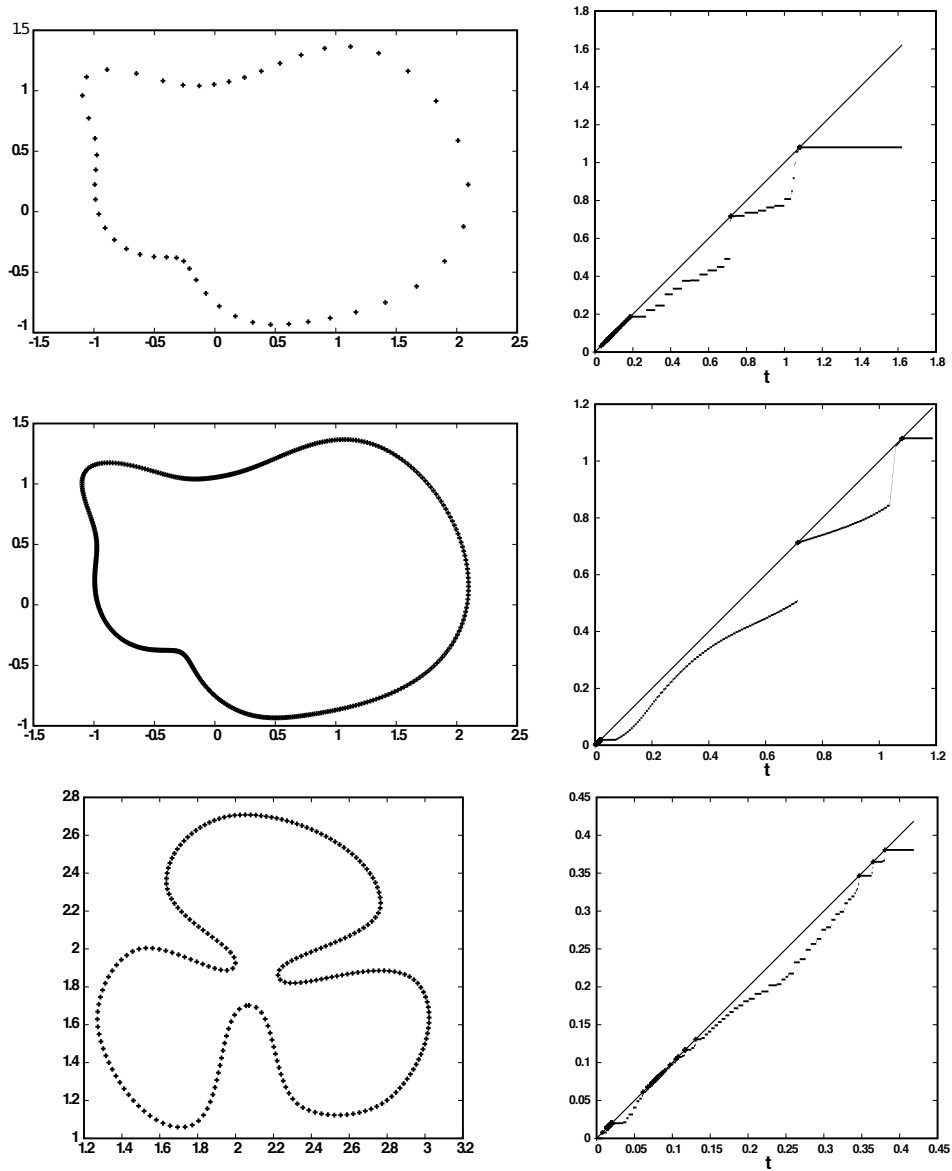


FIGURE 2.3 – Différents ensembles de points  $P$  sur la gauche et tracé des fonctions de défaut de convexité  $c_P$  respectives sur la droite.

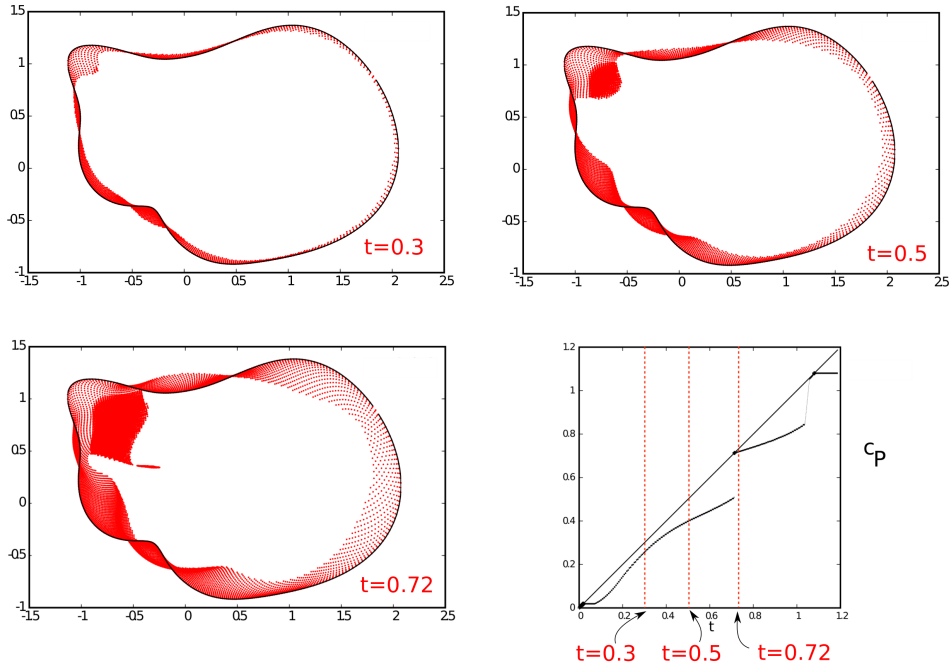


FIGURE 2.4 – Les trois premières figures représentent  $\text{Centers}(P, t)$  pour  $t \in \{0, 3; 0, 5; 0, 72\}$ . La dernière figure donne le graphe de la fonction  $c_P(t)$ .

*Démonstration.* Soit  $\sigma \subset \mathbb{R}^D$  un sous-ensemble fermé et non vide et un point  $y \in \mathbb{R}^D$ . En posant  $x = y$  dans le lemme 2.1, on remarque que si  $y \in \text{Hull}(\sigma)$  vérifie  $d(y, \sigma) \geq t$  et si  $\text{Rad}(\sigma) \leq t$ , alors  $y = \text{Center}(\sigma)$  et  $t = \text{Rad}(\sigma)$ .

Démontrons maintenant que (1)  $\implies$  (2). Soit  $y$  un point critique dont la distance à  $X$  est  $t$  et posons  $\sigma = \Gamma_X(y)$  comme illustré sur la figure 2.5. Par définition,  $y \in \text{Hull}(\sigma)$  et par construction  $d(y, \sigma) = t$  et  $\text{Rad}(\sigma) \leq t$ . D'après notre remarque précédente,  $y = \text{Center}(\sigma)$  et donc  $c_X(t) = t$ . L'implication (2)  $\implies$  (3) découle du fait que  $c_X(t) \leq h_X(t) \leq t$ . Montrons que (3)  $\implies$  (1). Autrement dit, supposons que  $h_X(t) = t$  et montrons que  $t$  est alors une valeur critique de  $d(\cdot, X)$ . Comme  $X$  est compact,  $h_X(t) = t$  implique que l'on peut trouver un compact  $\emptyset \neq \sigma \subset X$  avec  $\text{Rad}(\sigma) \leq t$  et  $y \in \text{Hull}(\sigma)$  tel que  $t = d(y, X) \leq d(y, \sigma)$ . Notre

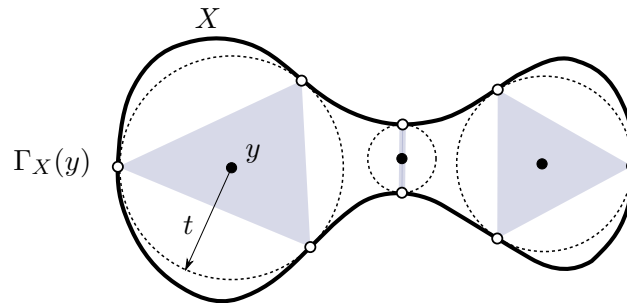


FIGURE 2.5 – Les points noirs sont les points critiques de la fonction de distance à la courbe  $X$ .

remarque implique que  $y = \text{Center}(\sigma)$ ,  $t = \text{Rad}(\sigma)$  et  $\sigma$  est un ensemble de points de  $X$  dont la distance à  $y$  est minimale. Puisque  $y \in \text{Hull}(\sigma) \subset \text{Hull}(\Gamma_X(y))$ , on en déduit que  $y$  est un point critique de la fonction de distance de  $X$  ce qui conclut la preuve.  $\square$

Une adaptation de la théorie de Morse à la fonction de distance nous indique que les changements de topologie des  $t$ -offsets  $X^t$  ont lieu lorsque le paramètre d'offset  $t$  atteint une valeur critique de la fonction de distance de  $X$ . En effet, en modifiant légèrement la proposition 1.8 de [Gro93, page 362], on obtient :

**Théorème 2.1** (Théorème d'isotopie [Gro93]). *Soit  $X \subset \mathbb{R}^D$  un compact et  $\beta \geq \alpha > 0$  deux réels positifs. Si la fonction de distance  $d(\cdot, X)$  n'a pas de valeur critique dans l'intervalle  $[\alpha, \beta]$ , alors  $X^\beta$  se rétracte par déformation sur  $X^\alpha$ .*

Notre caractérisation des valeurs critiques nous permet de ré-exprimer les conditions de ce théorème. En particulier, on obtient que si  $c_X(t) < t$  pour tout  $t \in [\alpha, \beta]$ , alors  $X^\beta$  se rétracte par déformation sur  $X^\alpha$ . En remplaçant  $X$  par un ensemble fini de points  $P$  et en utilisant le fait que le complexe Cech( $P, t$ ) a le même type d'homotopie que le  $t$ -offset  $P^t$ , on obtient que le Cech( $P, \beta$ ) a le même type d'homotopie que Cech( $P, \alpha$ ) dès lors que  $c_P(t) < t$  pour tout  $t \in [\alpha, \beta]$ . Nous verrons dans la section 2.3 que, sous cette même condition, un résultat plus fort est vérifié : il existe une séquence de collapsés de Cech( $P, \beta$ ) vers Cech( $P, \alpha$ ). En renforçant cette condition, il est possible de montrer l'existence d'une séquence de collapsés de Rips( $P, \beta$ ) vers Cech( $P, \alpha$ ). Des variantes de cette condition seront déclinées dans la section 2.4 permettant de certifier la correction topologique de reconstruction de forme par les complexes de Čech ou de Rips.

### 2.2.3 Stabilité

Dans cette section, nous établissons deux résultats sur la stabilité de  $c_X$  et  $h_X$ . Pour des raisons techniques, nous avons besoin de la stabilité de  $c_P$  par rapport aux perturbations de  $P$  à la fin de la preuve du théorème 2.3 dans la section 2.3 afin de relâcher une hypothèse de généricité. La stabilité de  $h_X$  par rapport aux perturbations de  $X$  sera un élément clé dans la section 2.4 et permettra d'établir les constantes dans les théorèmes de reconstruction.

**Lemme 2.3.** *Pour toute paire  $X, P \subset \mathbb{R}^D$  telle que  $d_H(X, P) \leq \varepsilon$  et pour tout  $t \geq 0$ ,*

$$c_P(t) \leq c_X(t + \varepsilon) + \sqrt{2t\varepsilon + \varepsilon^2} + \varepsilon.$$

*Démonstration.* Soit  $\sigma \subset P$  un sous-ensemble non vide de  $P$  avec  $\text{Rad}(\sigma) \leq t$  et posons  $\xi = X \cap \sigma^\varepsilon$ . Par construction,  $\xi$  est non vide et  $d_H(\xi, \sigma) \leq \varepsilon$ . Ainsi, en posant  $\delta = \sqrt{2t\varepsilon + \varepsilon^2}$ , le lemme A.1 implique que  $\text{Rad}(\xi) \leq t + \varepsilon$  et que  $\|\text{Center}(\sigma) - \text{Center}(\xi)\| \leq \delta$ . On obtient alors :

$$\text{Center}(\sigma) \subset \text{Center}(\xi)^\delta \subset X^{c_X(t+\varepsilon)+\delta} \subset P^{c_X(t+\varepsilon)+\delta+\varepsilon},$$

ce qui montre l'inégalité.  $\square$

**Lemme 2.4.** *Pour toute paire  $X, P \subset \mathbb{R}^D$  telle que  $d_H(X, P) \leq \varepsilon$  et pour tout  $t \geq 0$ ,  $h_P(t) \leq h_X(t + \varepsilon) + 2\varepsilon$ .*

*Démonstration.* Soit  $\sigma \subset P$  un sous-ensemble non vide de  $P$  avec  $\text{Rad}(\sigma) \leq t$  et posons  $\xi = X \cap \sigma^\varepsilon$ . Par construction,  $\xi$  est non vide et  $d_H(\xi, \sigma) \leq \varepsilon$ . Ainsi, le lemme A.1 implique que  $\text{Rad}(\xi) \leq t + \varepsilon$ . En utilisant  $\text{Hull}(\xi^\varepsilon) = \text{Hull}(\xi)^\varepsilon$ , on obtient que  $\text{Hull}(\sigma) \subset \text{Hull}(\xi)^\varepsilon \subset X^{h_X(t+\varepsilon)+\varepsilon} \subset P^{h_X(t+\varepsilon)+2\varepsilon}$ , ce qui montre l'inégalité.  $\square$

## 2.3 Du complexe de Čech au complexe de Rips

Dans cette section, nous introduisons une famille de complexes de Rips possédant deux paramètres et nous donnons une condition sur un ensemble de points assurant qu'un complexe de Rips de cette famille se rétracte par déformation sur un complexe de Čech. Nous commençons par définir cette famille de complexes de Rips avant d'établir nos résultats dans la section 2.3.1. Enfin, nous présentons la preuve de ces résultats dans la section 2.3.2. Nous montrons également que la condition de notre résultat est optimale en petite dimension dans la section 2.3.3.

### 2.3.1 Quasi-complexes de Rips et énoncé des résultats

De la même manière que [CdSEG10], nous commençons par définir les quasi-complexes de Rips. La motivation de cette définition est de pouvoir tenir compte d'une incertitude éventuelle sur des mesures en autorisant certaines arêtes à appartenir ou non au quasi complexe de Rips ; voir [CdSEG10].

**Définition 2.3.** *Pour tout ensemble de points  $P \subset \mathbb{R}^D$  et tous réels  $\beta \geq \alpha \geq 0$ , on appelle le flag complexe de tout graphe  $G$  satisfaisant  $\text{Rips}(P, \alpha) \subset \text{Flag}(G) \subset \text{Rips}(P, \beta)$  un  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$ .*

On remarque que le complexe simplicial  $\text{Flag}(G)$  est un  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$  si et seulement si toute paire de points dans  $P$  de distance inférieure à  $2\alpha$  est connectée par une arête dans  $G$  et qu'aucune arête de  $G$  ne possède une longueur strictement plus grande que  $2\beta$ . De manière équivalente, pour toute paire de points  $(p, q) \in P^2$ ,  $\|p - q\| \leq 2\alpha$  implique que  $pq \in G$  et  $\|p - q\| > 2\beta$  implique que  $pq \notin G$ . En particulier,  $K$  est un  $(\alpha, \alpha)$ -quasi complexe de Rips de  $P$  si et seulement si  $K = \text{Rips}(P, \alpha)$ .

Pour établir nos résultats, il est commode de définir une *valeur inerte* de  $P$  comme un réel  $\alpha$  tel que  $\text{Rad}(\sigma) \neq \alpha$  pour tout  $\sigma$  sous-ensemble non vide de  $P$ . L'ensemble  $P$  possède un nombre fini de valeurs inertes puisqu'il est fini. Ainsi, l'hypothèse que  $\alpha$  est une valeur inerte n'est pas trop restrictive.

Dans un premier temps, nous établissons des conditions dans le théorème 2.2 assurant qu'il existe une séquence de collapsés classiques transformant un complexe de Čech en un



autre complexe de Čech. Nous rappelons qu'un *collapse classique* est l'opération qui enlève la paire de simplexes  $(\sigma, \tau)$  d'un complexe simplicial  $K$  en supposant que  $\tau$  est l'unique coface propre de  $\sigma$  dans  $K$ . De cette soustraction résulte un complexe simplicial  $K \setminus \{\sigma, \tau\}$  vers lequel  $K$  se rétracte par déformation.

**Théorème 2.2.** *Soit  $P \subset \mathbb{R}^D$  un ensemble fini de points. Pour tous réels  $\beta \geq \alpha \geq 0$  tels que  $\alpha$  est une valeur inerte de  $P$  et  $c_P(t) < t$  pour tout  $t \in [\alpha, \beta]$ , il existe une séquence de collapses classiques de  $\text{Cech}(P, \beta)$  vers  $\text{Cech}(P, \alpha)$ .*

Ce théorème peut être vu comme une version combinatoire du théorème d'isotopie présenté dans la section 2.2.2. Nous sommes maintenant prêts pour établir le principal résultat de ce chapitre :

**Théorème 2.3.** *Soit  $P \subset \mathbb{R}^D$  un ensemble fini de points. Pour tous réels  $\beta \geq \alpha \geq 0$  tels que  $\alpha$  est une valeur inerte de  $P$  et tels que  $c_P(\vartheta_D\beta) < 2\alpha - \vartheta_D\beta$ , il existe une séquence de collapses classiques de n'importe quel  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$  vers le Čech complexe  $\text{Cech}(P, \alpha)$ .*

Remarquons en particulier qu'en posant  $\alpha = \beta$  dans le théorème précédent, on obtient des conditions pour lesquelles  $|\text{Rips}(P, \alpha)| \simeq |\text{Cech}(P, \alpha)| \simeq P^\alpha$ . Nous verrons dans la section 2.3.3 que la borne sur  $c_P$  est optimale en faible dimension. La partie gauche de la figure 2.8 montre une illustration graphique des hypothèses de ce théorème. Donnons une brève idée de la preuve du théorème 2.3. Observons que la condition  $c_P(\vartheta_D\beta) < 2\alpha - \vartheta_D\beta$  implique que  $c_P(t) < t$  pour tout  $t \in [\alpha, \vartheta_D\beta]$  puisque

$$c_P(t) \leq c_P(\vartheta_D\beta) < 2\alpha - \vartheta_D\beta \leq \alpha \leq t,$$

dès lors que  $\alpha \leq t \leq \vartheta_D\beta$ . Le théorème 2.2 implique alors qu'il existe une séquence de collapses classiques réduisant  $\text{Cech}(P, \vartheta_D\beta)$  vers  $\text{Cech}(P, \alpha)$ . Puisque tout  $(\alpha, \beta)$ -quasi complexe de Rips  $\text{Flag}(G)$  est emboîté entre  $\text{Cech}(P, \alpha)$  et  $\text{Cech}(P, \vartheta_D\beta)$ , l'idée clé dans la preuve de ce théorème est d'inspecter les changements dans le complexe  $\text{Cech}(P, t) \cap \text{Flag}(G)$  lorsque l'on diminue le paramètre  $t$  de  $\vartheta_D\beta$  vers  $\alpha$ , c'est-à-dire lorsque l'on va de  $\text{Cech}(P, \vartheta_D\beta)$  vers  $\text{Flag}(G)$ . La preuve consiste à vérifier que ces changements ne modifient pas le type d'homotopie, en particulier, on vérifie que ces changements sont des collapses étendus.

### 2.3.2 Preuve des théorèmes 2.2 et 2.3

Nous montrons directement le théorème 2.3 ; la preuve du théorème 2.2 peut être facilement adaptée en modifiant légèrement le début de la preuve présentée ci-dessous.

*Preuve du théorème 2.3.* Soit  $G$  un graphe dont le *flag* complexe est un  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$ . Pour  $t \geq 0$ , considérons le complexe simplicial  $\mathcal{F}(t) = \text{Cech}(P, t) \cap \text{Flag}(G)$ . Clairement, nous avons la chaîne d'inclusions :

$$\text{Cech}(P, \alpha) \subset \text{Rips}(P, \alpha) \subset \text{Flag}(G) \subset \text{Rips}(P, \beta) \subset \text{Cech}(P, \vartheta_D\beta)$$

et donc  $\mathcal{F}(\alpha) = \text{Cech}(P, \alpha)$  et  $\mathcal{F}(\vartheta_D\beta) = \text{Flag}(G)$ . En augmentant continûment le paramètre  $t$  de  $\alpha$  vers  $\vartheta_D\beta$ , on obtient une famille de complexes de Čech emboîtés :

$$\text{Cech}(P, \alpha) = \mathcal{C}_0 \subset \mathcal{C}_1 \subset \cdots \subset \mathcal{C}_k = \text{Cech}(P, \vartheta_D\beta).$$

Pour  $0 < i < k$ , soit  $t_i$  la plus petite valeur de  $t$  telle que  $\mathcal{C}_i = \text{Cech}(P, t)$  et soit  $\mathcal{F}_i = \mathcal{F}(t_i)$ . En particulier,  $\mathcal{C}_i = \text{Cech}(P, t_i)$  et  $\mathcal{F}_i = \mathcal{C}_i \cap \text{Flag}(G)$ . En intersectant chaque complexe  $\mathcal{C}_i$  de la séquence ci-dessus avec  $\text{Flag}(G)$ , on obtient une autre famille de complexes simpliciaux emboîtés :

$$\text{Cech}(P, \alpha) = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_k = \text{Flag}(G).$$

En premier lieu, supposons que  $P$  satisfasse les deux hypothèses de généricité  $(\star)$  et  $(\star\star)$  au lieu de l'hypothèse  $\text{Rad}(\sigma) \neq \alpha$  pour tout  $\emptyset \neq \sigma \subset P$  :

- $(\star)$  Pour tous simplexes  $\sigma, \tau \subset P$ , si  $\text{Rad}(\sigma) = \text{Rad}(\tau)$  alors  $\text{Center}(\sigma) = \text{Center}(\tau)$  ;
- $(\star\star)$  Pour toute boule  $B$ , l'ensemble des simplexes dans  $P$  ayant  $B$  comme boule englobante minimale est soit vide soit possède un unique élément minimal pour l'inclusion.

Sous ces deux conditions, nous prouvons le théorème en deux étapes. Dans un premier temps, nous montrons que l'opération réduisant  $\mathcal{C}_i$  en  $\mathcal{C}_{i-1}$  est un collapse pour tout  $0 < i \leq k$ . Dans un deuxième temps, nous montrons que soit  $\mathcal{F}_{i-1}$  est identique à  $\mathcal{F}_i$  soit l'opération transformant  $\mathcal{F}_i$  en  $\mathcal{F}_{i-1}$  est un collapse pour tout  $0 < i \leq k$ .

(a) D'après l'hypothèse  $(\star)$ , tous les simplexes de l'ensemble  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  partagent la même boule englobante minimale  $B(z_i, t_i)$  de centre  $z_i$  et de rayon  $t_i$ . L'hypothèse  $(\star\star)$  implique quant-à-elle que l'ensemble des simplexes partageant la même boule englobante minimale  $B(z_i, t_i)$  possède un élément minimal pour l'inclusion  $\sigma_i$ . Il suffit de montrer que  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  est la *star* de  $\sigma_i$  et possède un unique élément maximal pour l'inclusion  $\tau_i \neq \sigma_i$  ce qui garantit que l'opération transformant  $\mathcal{C}_i$  en  $\mathcal{C}_{i-1}$  est un collapse classique. Pour cela, supposons que  $\eta$  soit une coface de  $\sigma_i$  dans  $\mathcal{C}_i$ . Puisque  $\sigma_i \subset \eta$ , on déduit que  $t_i = \text{Rad}(\sigma_i) \leq \text{Rad}(\eta)$  et ainsi  $\eta \in \mathcal{C}_i \setminus \mathcal{C}_{i-1}$ . Le complexe  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$  est alors la *star* de  $\sigma_i$  dans  $\mathcal{C}_i$ . Notons que le simplexe  $\tau_i = \{p \in P \mid \|z_i - p\| \leq t_i\}$  obtenu en rassemblant tous les points de  $P$  dans  $B(z_i, t_i)$  appartient à  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$ , de plus, il s'agit de l'unique simplexe maximal pour l'inclusion dans cet ensemble (voir la partie gauche de la figure 2.6). Afin de montrer que l'opération réduisant  $\mathcal{C}_i$  à  $\mathcal{C}_{i-1}$  est un collapse, il reste à établir que  $\sigma_i \neq \tau_i$ . Comme  $\sigma_i$  est l'élément minimal pour l'inclusion parmi tous les simplexes dont la plus petite boule englobante est  $B(z_i, t_i)$ , les sommets de  $\sigma_i$  sont tous situés dans la sphère de centre  $z_i$  et de rayon  $t_i$ . D'un autre côté, par définition de  $c_P(t_i)$ , il existe au moins un point  $o$  de  $P$  à distance  $c_P(t_i)$  du centre  $z_i$ . Comme  $c_P(t_i) \leq c_P(\vartheta_D\beta) < \alpha \leq t_i$ , le point  $o$  appartient à l'intérieur de  $B(z_i, t_i)$ . Ainsi,  $o \notin \sigma_i$ ,  $o \in \tau_i$ , et donc  $\sigma_i \neq \tau_i$ , ce qui montre que l'opération réduisant  $\mathcal{C}_i$  vers  $\mathcal{C}_{i-1}$  est un collapse classique.

(b) Penchons-nous à présent sur  $\mathcal{F}_i$  et  $\mathcal{F}_{i-1}$ . Si  $\sigma_i \notin \mathcal{F}_i$ , alors  $\mathcal{F}_i = \mathcal{F}_{i-1}$ . Si  $\sigma_i \in \mathcal{F}_i$ , la *star* de  $\sigma_i$  dans  $\mathcal{F}_i$  est la *star* de  $\sigma_i$  dans  $\mathcal{C}_i$  intersecté avec la *flag* de  $G$  et  $\mathcal{F}_{i-1} = \mathcal{F}_i \setminus \text{St}_{\mathcal{F}_i}(\sigma_i)$  (voir la partie droite de la figure 2.6). Montrons que la *link* de  $\sigma_i$  dans  $\mathcal{F}_i$  est un cône d'apex  $o$ , ce qui garantira que la réduction de  $\mathcal{F}_i$  à  $\mathcal{F}_{i-1}$  est un collapse étendu. Supposons que  $\eta$  soit une coface de  $\sigma_i$  dans  $\mathcal{F}_i$  et montrons que  $\eta \cup \{o\}$  est aussi une coface. Clairement,  $\eta \cup \{o\}$  appartient au complexe de Čech  $\mathcal{C}_i$  puisque pour tous les points  $p \in \eta \cup \{o\}$ ,  $\|z_i - p\| \leq t_i$ . Remarquons que  $\eta \cup \{o\}$  appartient aussi à  $\text{Flag}(G)$ . Pour s'en convaincre, puisque  $\eta$  appartient à  $\text{Flag}(G)$ ,

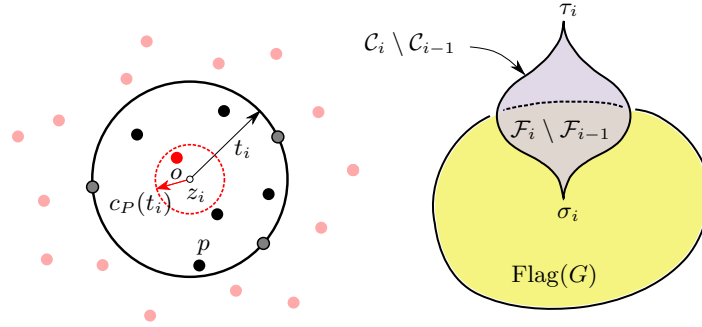


FIGURE 2.6 – Notations pour la preuve du théorème 2.3. A gauche :  $\tau_i$  est le simplexe dont les sommets sont les points de  $P$  dans la boule  $B(z_i, t_i)$ .  $\sigma_i$  est la face obtenue en gardant les sommets sur le bord de  $B(z_i, t_i)$ . A droite : représentation schématique des simplexes de  $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$ .

il suffit de vérifier que toutes les arêtes connectant  $o$  à un sommet  $p$  de  $\eta$  ont une longueur inférieure ou égale à  $2\alpha$ . Pour tous les points  $p \in \eta$ , les inégalités suivantes sont vérifiées :

$$\|p - o\| \leq \|z_i - p\| + \|z_i - o\| \leq t_i + c_P(t_i) \leq 2\alpha$$

ce qui montre que  $\eta \cup \{o\} \in \text{Flag}(G)$ . Ainsi,  $\eta \cup \{o\}$  appartient à  $\mathcal{F}_i$ . En posant  $\eta = \sigma_i$ , on obtient que  $\sigma_i \cup \{o\}$  est une coface de  $\sigma_i$  et puisque  $o \notin \sigma_i$ , il s'ensuit que  $\{o\}$  appartient au *link* de  $\sigma_i$  dans  $\mathcal{F}_i$ . Le *link* de  $\sigma_i$  dans  $\mathcal{F}_i$  est donc un cône, ce qui conclut la preuve du théorème 2.3 en supposant les hypothèses de généricité  $(\star)$  et  $(\star\star)$  à la place de la condition  $\text{Rad}(\sigma) \neq \alpha$  pour tout  $\emptyset \neq \sigma \subset P$ .

Lorsque  $P$  ne satisfait pas les hypothèses de généricité  $(\star)$  et  $(\star\star)$ , nous utilisons le lemme A.2 dans l'annexe A.2 pour trouver une perturbation  $f$  des points telle que  $f(P)$  vérifie  $(\star)$  et  $(\star\star)$  ainsi que les hypothèses  $(i)$ ,  $(ii)$  et  $(iii)$  du lemme A.2 pour un  $\beta' > \beta$ . En utilisant le théorème 2.3 à  $f(P)$  avec les valeurs  $\alpha$  et  $\beta'$ , on montre qu'il existe une séquence de collapsés étendus du  $(\alpha, \beta')$ -quasi complexe de Rips  $\text{Flag}(f(G)) = f(\text{Flag}(G))$  vers le complexe de Čech  $\text{Cech}(f(P), \alpha) = f(\text{Cech}(P, \alpha))$ . Ainsi, le théorème est encore vérifié dans les cas non génériques.  $\square$

### 2.3.3 Optimalité de la condition du théorème 2.3

Nous montrons ici que, lorsque la dimension ambiante  $D$  est égale à 2, la condition du théorème 2.3 sur la fonction de défaut de convexité  $c_P$  est optimale pour les complexes de Rips (i.e. lorsque  $\alpha = \beta$ ); il n'est pas possible d'avoir un résultat similaire au théorème 2.3 avec une borne plus grande sur  $c_P$ . Pour cela, nous construisons un ensemble de points  $P_\gamma$  tel que  $\text{Rips}(P_\gamma, \alpha) \not\cong \text{Cech}(P_\gamma, \alpha)$  et tel que  $c_{P_\gamma}(\vartheta_D \alpha)$  est supérieure mais infinitésimalement proche de  $(2 - \vartheta_D)\alpha$ . Afin de vérifier que les complexes de Rips et de Čech ont un type d'homotopie différent, nous avons besoin du théorème suivant dont une preuve sera donnée dans le chapitre 5.

**Théorème 2.1.** *Si  $\text{Hull}(P) \subset P^\alpha$  alors  $\text{Cech}(P, \alpha)$  est collapsable.*

Afin de construire notre ensemble de points  $P_\gamma$ , notons  $\{v_0, \dots, v_D\}$  les sommets d'un simplexe régulier  $\sigma$  de diamètre  $2\alpha$ . Etant donné  $\gamma > 0$ , notons également  $\bar{v}_i$  le point de la sphère  $S(\text{Center}(\sigma), \gamma)$  le plus éloigné de  $v_i$ . Nous définissons alors l'ensemble des points  $P_\gamma = \{v_0, \dots, v_D, \bar{v}_0, \dots, \bar{v}_D\}$ ; voir la figure 2.7.

Intéressons-nous maintenant aux complexes de Rips et de Čech des points  $P_\gamma$  de paramètres  $\alpha$ . En particulier, remarquons que :

- (i) Le complexe  $\text{Rips}(P_\gamma, \alpha)$  est une boule topologique lorsque  $\gamma = (2 - \vartheta_D)\alpha$  et il s'agit d'une sphère topologique lorsque  $\gamma$  est supérieur mais infinitésimalement proche de  $(2 - \vartheta_D)\alpha$ .
- (ii) Le complexe  $\text{Cech}(P_\gamma, \alpha)$  est collapsable lorsque  $(2 - \vartheta_D)\alpha \leq \gamma \leq \alpha$ .

Lorsque  $\gamma = (2 - \vartheta_D)\alpha$ , le 1-squelette de  $\text{Rips}(P_\gamma, \alpha)$  est un graphe complet (voir la partie gauche de la figure 2.7). En effet, comme  $\sigma$  est un simplexe régulier, on sait que  $\text{Rad}(\sigma) = \frac{\vartheta_D}{2} \text{Diam}(\sigma) = \vartheta_D \alpha$  et ainsi  $\text{Diam}(\sigma) - \text{Rad}(\sigma) = (2 - \vartheta_D)\alpha$ . Ainsi lorsque  $\gamma = (2 - \vartheta_D)\alpha$ , les deux sphères  $S(e_i, \alpha)$  et  $S(e'_i, \alpha)$  sont tangentes pour tout  $i$  et le 1-squelette de  $\text{Rips}(P_\gamma, \alpha)$  est alors un graphe complet. Lorsque  $\gamma$  est infinitésimalement supérieur à  $(2 - \vartheta_D)\alpha$ , les arêtes  $v_i \bar{v}_i$  disparaissent du complexe  $\text{Rips}(P_\gamma, \alpha)$ . Le complexe de Rips est alors la sphère topologique  $S_D$  puisqu'il se décompose en un produit  $S_0 \star \dots \star S_0$  (où  $S_0$  dénote la 0-sphère) [Hat02].

Concernant le complexe de Čech, un calcul de distance donné dans l'annexe B montre que, lorsque  $\frac{2 - \vartheta_D^2}{\vartheta_D} \alpha \leq \gamma \leq \vartheta_D \alpha$ ,  $c_{P_\gamma}(\vartheta_D \alpha) = h_{P_\gamma}(\vartheta_D \alpha) = \gamma$ . Comme  $\frac{2 - \vartheta_D^2}{\vartheta_D} \alpha \leq (2 - \vartheta_D)\alpha$  et  $\alpha \leq \vartheta_D \alpha$ , on a donc que  $c_{P_\gamma}(\vartheta_D \alpha) = h_{P_\gamma}(\vartheta_D \alpha) = \gamma$  lorsque  $(2 - \vartheta_D)\alpha \leq \gamma \leq \alpha$ . Ainsi le théorème 2.1 indique que le complexe  $\text{Cech}(P_\gamma, \alpha)$  est collapsable lorsque  $(2 - \vartheta_D)\alpha \leq \gamma \leq \alpha$ .

On obtient finalement la conclusion souhaitée à savoir que si  $\gamma$  est infinitésimalement proche de  $(2 - \vartheta_D)\alpha$  alors  $\text{Rips}(P_\gamma, \alpha) \not\cong \text{Cech}(P_\gamma, \alpha)$ . On a donc construit un ensemble de points  $P_\gamma$  tel que  $\text{Rips}(P_\gamma, \alpha) \not\cong \text{Cech}(P_\gamma, \alpha)$  et tel que  $c_{P_\gamma}(\vartheta_D \alpha)$  est supérieure mais infinitésimalement proche de  $(2 - \vartheta_D)\alpha$  ce qui montre l'optimalité de la condition sur la fonction de défaut de convexité  $c_P$  du théorème 2.3 lorsque la dimension  $D$  est inférieure à 3. Le problème de l'optimalité de cette condition lorsque  $D > 2$  reste quant-à-lui ouvert.

**Remarque 2.1.** *Remarquons que le contre-exemple donné dans cette section ne s'étend pas en toute dimension. En effet, pour une dimension ambiante  $D$  assez grande, la propriété  $c_{P_\gamma}(\vartheta_D \alpha) = \gamma$  pour  $(2 - \vartheta_D)\alpha \leq \gamma \leq \alpha$  est perdue. Pour le montrer, prenons  $\alpha = 1$  et  $\gamma = (2 - \vartheta_D)$ . On peut alors calculer la limite  $\lim_{D \rightarrow \infty} \|v_0 - \bar{v}_D\| = 2\sqrt{2} - \sqrt{2}$ . Ainsi le milieu  $c$  de  $v_0$  et  $\bar{v}_D$  appartient à  $\text{Centers}(P_\gamma, \vartheta_D)$  et alors  $c_{P_\gamma}(\vartheta_D \alpha) \geq d(c, P_\gamma) = \sqrt{2} - \sqrt{2} > \gamma = 2 - \sqrt{2}$ .*

**Remarque 2.2.** *On pourrait chercher un théorème similaire au théorème 2.3, avec une borne sur  $h_P$  plutôt que  $c_P$ . Comme  $h_P \geq c_P$ , le théorème 2.3 tient toujours si on garde toutes les hypothèses sauf la borne sur  $c_P$  qu'on remplace par la borne  $h_P(\vartheta_D \beta) < 2\alpha - \vartheta_D \beta$ . On pourrait espérer obtenir le même résultat avec une borne plus faible sur  $h_P$  mais c'est impossible. En effet, le même raisonnement utilisé dans l'exemple précédent pour montrer qu'on ne peut pas abaisser la condition sur  $c_P$  permet également de montrer qu'on ne peut pas abaisser la condition sur  $h_P$  du moins lorsque  $D < 3$ .*

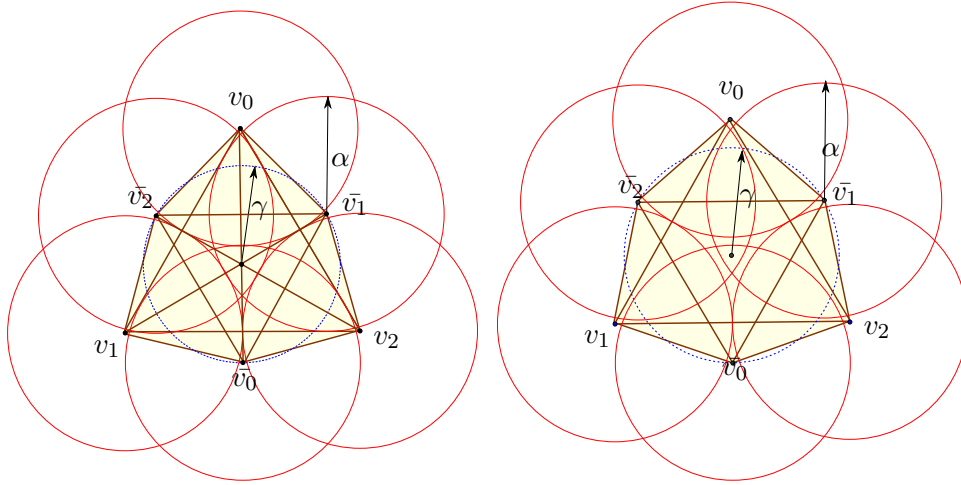


FIGURE 2.7 – A gauche : le complexe de Rips et le complexe de Čech sont contractiles lorsque  $\gamma = (2 - \vartheta_D)\alpha$  et  $D = 2$ . A droite : si on augmente infinitésimalement la valeur de  $\gamma$ , le complexe de Rips est une sphère tandis que le complexe de Čech est toujours contractile.

## 2.4 Reconstruction de forme

Dans cette section, nous nous intéressons à la reconstruction d'un compact  $X \subset \mathbb{R}^D$  connu seulement à partir d'un ensemble fini de points éventuellement bruités  $P \subset \mathbb{R}^D$ . En utilisant la fonction de défaut de convexité  $h_X$ , nous formulons deux conditions d'échantillonnage garantissant respectivement que le complexe de Čech et le complexe de Rips partagent le même type d'homotopie qu'un offset de  $X$  (Section 2.4.1). Dans la section 2.4.2, nous faisons ensuite un lien entre le fait, pour une forme, d'avoir sa fonction de défaut convexité bornée et le fait d'avoir une borne inférieure sur sa fonction critique. Finalement, nous calculons dans la section 2.4.3 les densités les plus faibles de l'ensemble de points permettant de satisfaire les hypothèses de nos théorèmes pour assurer une reconstruction correcte d'une forme à  $\mu$ -reach positif.

### 2.4.1 Conditions d'échantillonnage

Nous rassemblons ici plusieurs idées introduites dans ce chapitre afin de déduire des conditions sous lesquelles le complexe de Čech et le complexe de Rips d'un ensemble fini de points récupèrent la topologie d'une forme échantillonnée. Durant cette section,  $X$  désigne un compact de  $\mathbb{R}^D$  et  $P$  est un ensemble fini de points dont la distance de Hausdorff à  $X$  est inférieure ou égale à  $\varepsilon$ .

**Reconstruction à partir du complexe de Čech** L'hypothèse  $d_H(X, P) \leq \varepsilon$  implique la séquence suivante d'inclusions :

$$P^\alpha \subset X^{\alpha+\varepsilon} \subset P^{\alpha+2\varepsilon} \subset X^{\alpha+3\varepsilon}.$$

Or, lorsque quatre espaces sont emboîtés  $P_0 \subset X_0 \subset P_1 \subset X_1$ , que  $X_1$  se rétracte par déformation vers  $X_0$  et que  $P_1$  se rétracte par déformation sur  $P_0$ , alors  $X_0$  se rétracte par déformation vers  $P_0$  [AL10].

En appliquant ce résultat dans notre contexte et en utilisant le théorème d'isotopie ainsi que la caractérisation des points critiques donnée dans le lemme 2.2, on déduit immédiatement que  $X^{\alpha+\varepsilon}$  se rétracte par déformation vers  $P^\alpha$  dès lors que les deux conditions suivantes sont remplies :

$$\begin{aligned} h_X(t) &< t, & \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon], \\ h_P(t) &< t, & \forall t \in [\alpha, \alpha + 2\varepsilon]. \end{aligned}$$

Puisque  $d_H(X, P) \leq \varepsilon$ , le lemme 2.4 implique que  $h_P(t) \leq h_X(t + \varepsilon) + 2\varepsilon$ . Les deux conditions ci-dessus sont donc vérifiées dès lors que la condition plus forte est remplie :  $h_X(t) < t - 3\varepsilon, \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$ . Comme  $h_X$  est positive, cette condition implique que  $2\varepsilon < \alpha$ . Puisque  $h_X$  est croissante, elle implique aussi que  $h_X(t) < t$  pour tout  $t \in [\alpha - 2\varepsilon, \alpha + 3\varepsilon]$ , montrant que les  $\eta$ -offsets de  $X$  ont tous le même type d'homotopie quand  $\eta$  est dans l'intervalle  $[\alpha - 2\varepsilon, \alpha + 3\varepsilon]$ . Le théorème suivant résume ces observations :

**Théorème 2.4.** *Soient  $\varepsilon, \alpha > 0$  tels que  $2\varepsilon < \alpha$ . Soit  $P$  un ensemble fini de points dont la distance de Hausdorff à un compact  $X$  est inférieure ou égale à  $\varepsilon$ . Le complexe de Čech  $\text{Cech}(P, \alpha)$  a le même type d'homotopie que  $X^\eta$  pour tout  $\eta \in [\alpha - 2\varepsilon, \alpha + 3\varepsilon]$  dès lors que  $h_X(t) < t - 3\varepsilon$  pour tout  $t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$ .*

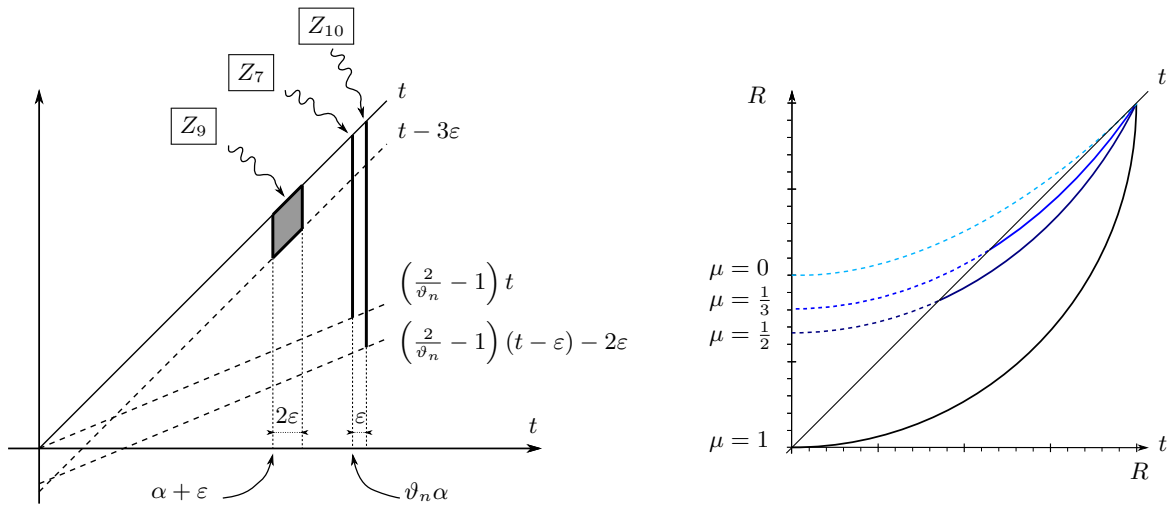


FIGURE 2.8 – A gauche : pour  $i \in \{2.3, 2.4, 2.5\}$ , les hypothèses du théorème  $i$  sont représentées par les régions  $Z_i$  qui ne croisent pas le graphe de la fonction de défaut de convexité. Plus précisément, si  $c_P \cap Z_7 = \emptyset$ , le théorème 2.3 implique que  $\text{Rips}(P, \alpha) \simeq P^\alpha$ . Si  $h_X \cap Z_9 = \emptyset$ , le théorème 2.4 implique que  $\text{Cech}(P, \alpha) \simeq X^{\alpha-2\varepsilon}$ . Si  $h_X \cap Z_{10} = \emptyset$ , le théorème 2.5 implique que  $\text{Rips}(P, \alpha) \simeq X^{2\alpha-\vartheta_D\alpha-2\varepsilon}$ . A droite : borne supérieure sur  $h_X$  pour  $\mu \in \{0, \frac{1}{3}, \frac{1}{2}, 1\}$  fournie par le lemme 2.6.

**Reconstruction pour le complexe de Rips** Si on suppose en plus que la condition  $c_P(\vartheta_D\beta) < 2\alpha - \vartheta_D\beta$  est vérifiée, nous pouvons utiliser le théorème 2.3 pour s'assurer que le  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$  se rétracte par déformation sur le complexe de Čech  $\check{C}ech(P, \alpha)$ . En utilisant le lemme 2.4, nous obtenons que  $c_P(\vartheta_D\beta) \leq h_P(\vartheta_D\beta) \leq h_X(\vartheta_D\beta + \varepsilon) + 2\varepsilon$  et l'hypothèse du théorème 2.3 est remplie dès lors que  $h_X(\vartheta_D\beta + \varepsilon) < 2\alpha - \vartheta_D\beta - 2\varepsilon$ . Comme  $h_X$  est positive, cette condition implique que  $2\varepsilon < 2\alpha - \vartheta_D\beta$ . Puisque  $h_X$  est croissante, elle implique aussi que  $h_X(t) < t - 3\varepsilon, \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$  et l'hypothèse du théorème 2.4 est aussi vérifiée. On en déduit le théorème suivant :

**Théorème 2.5.** *Soient  $\varepsilon, \alpha$  et  $\beta$  trois réels positifs tels que  $\alpha \leq \beta$  et  $2\varepsilon < 2\alpha - \vartheta_D\beta$ . Soit  $P$  un ensemble fini de points dont la distance de Hausdorff au compact  $X$  est inférieure ou égale à  $\varepsilon$ . Sous ces hypothèses, tout  $(\alpha, \beta)$ -quasi complexe de Rips de  $P$  a le même type d'homotopie que  $X^\eta$  pour tout  $\eta \in [2\alpha - \vartheta_D\beta - 2\varepsilon, \vartheta_D\beta + \varepsilon]$  dès lors que  $\alpha$  est une valeur inerte de  $P$  et que  $h_X(\vartheta_D\beta + \varepsilon) < 2\alpha - \vartheta_D\beta - 2\varepsilon$ .*

## 2.4.2 Liens avec la fonction critique

Dans cette section, nous montrons qu'une forme possède une fonction de défaut de convexité bornée si et seulement si elle admet une borne inférieure sur sa fonction critique. Afin de préciser cette affirmation, nous rappelons la définition de la fonction critique qui a été introduite afin de pouvoir formuler des conditions d'échantillonnage pour la reconstruction de formes dont le *reach* peut éventuellement être nul [CCSL09]. Même si la fonction de distance à  $X$  n'est pas différentiable, on peut définir une *fonction de gradient généralisé*  $\nabla_X : \mathbb{R}^D \setminus X \rightarrow \mathbb{R}^D$  qui coïncide avec le gradient usuel sur les points où  $d(\cdot, X)$  est différentiable et qui s'annule précisément sur les points critiques [CCSL09] :

$$\nabla_X(y) = \frac{y - \text{Center}(\Gamma_X(y))}{d(y, X)}.$$

La fonction critique  $\chi_X : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$  est définie par

$$\chi_X(t) = \inf_{d(y, X)=t} \|\nabla_X(y)\|.$$

Clairement, la fonction critique s'annule en  $t$  si et seulement si  $t$  est une valeur critique de la fonction de distance. Ainsi, d'après le lemme 2.2, nous avons les équivalences :  $\chi_X(t) = 0 \iff c_X(t) = t \iff h_X(t) = t$ . Les deux prochains lemmes précisent les liens entre fonction critique et fonction de défaut de convexité. Le premier fournit une borne inférieure sur  $\chi_X$  en  $t$ , en supposant connue une borne supérieure sur  $c_X$  en  $t$ .

**Lemme 2.5.** *Pour tout compact  $X \subset \mathbb{R}^D$ , tout  $0 \leq \mu \leq 1$  et tout  $t \geq 0$ , l'implication suivante est vérifiée :*

$$c_X(t) < (1 - \mu)t \implies \chi_X(t) > \mu.$$

*Démonstration.* Soit  $y \in \mathbb{R}^D$  tel que  $d(y, X) = t$  et montrons que  $\|\nabla_X(y)\| > \mu$ . Soit  $\sigma = \Gamma_X(y)$  l'ensemble des points de  $X$  dont la distance à  $y$  est minimum ; voir la partie gauche de la figure 2.9. Supposons que la plus petite boule englobante de  $\sigma$  ait un centre  $z$  et un rayon  $s$ . Puisque  $s \leq t$ , on obtient que  $c_X(s) \leq c_X(t) < (1 - \mu)t$  et donc  $t - \|y - z\| \leq d(z, X) \leq c_X(s) < (1 - \mu)t$ . On en déduit que  $\|\nabla_X(y)\| = \frac{\|z - y\|}{t} > \mu$ .  $\square$

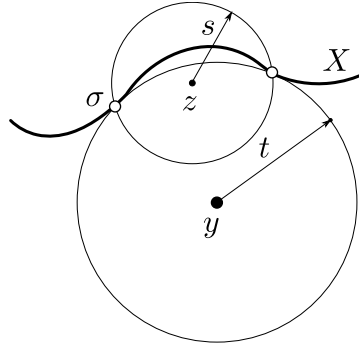


FIGURE 2.9 – Notations pour la preuve du lemme 2.5.

Le lemme suivant peut être vu comme une réciproque du lemme précédent, il fournit une borne supérieure sur  $h_X$  sur l'intervalle  $[0, R]$  étant donnée une borne inférieure sur la fonction critique  $\chi_X$  sur l'intervalle  $[0, R]$ . Ce lemme étend un résultat de [AL10] qui, intuitivement, établit que l'enveloppe convexe d'un ensemble de points  $\sigma \subset X$  ne peut pas être trop loin d'une forme  $X$ , en supposant que  $\sigma$  puisse être inclus dans une boule de rayon  $t$  et que  $X$  ait un *reach* positif. La preuve de ce lemme est donnée dans l'annexe A.3.

**Lemme 2.6.** *Soit  $\mu \in (0, 1]$  et  $R \geq 0$ . Soit  $X \subset \mathbb{R}^D$  un compact tel que  $\chi_X(t) \geq \mu$  pour tout  $t \in (0, R)$ . Alors, pour tout  $0 \leq t \leq R$ , l'inégalité suivante est vérifiée :*

$$h_X(t) \leq \frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left(\frac{t}{R}\right)^2}}{\mu(2 - \mu)} R.$$

La borne supérieure sur  $h_X$  est l'arc d'une ellipse qui tend vers l'arc d'une parabole lorsque  $\mu \rightarrow 0$  ; voir la partie droite de la figure 2.8. Notons que puisque  $h_X(t) \leq t$  pour tout  $t$ , cette borne supérieure est utile seulement sous la diagonale. Pour  $\mu = 1$ , on obtient que  $h_X(t) \leq R - \sqrt{R^2 - t^2}$  comme dans [AL10]. De manière équivalente, le graphe de  $h_X$  est en-dessous du cercle de rayon  $R$  et de centre  $(0, R)$ .

### 2.4.3 Reconstruction de formes à $\mu$ -reach positif

Les formes à  $\mu$ -reach positif regroupent une large classe d'objets qui, contrairement aux formes à reach positif, peuvent posséder des arêtes vives concaves. Précisément, pour  $0 < \mu \leq 1$ , les auteurs de [CCSL09] définissent le  $\mu$ -reach de  $X$  comme  $r_\mu(X) = \inf \{t > 0 \mid \chi_X(t) < \mu\}$ . Cette terminologie est justifiée par le fait que  $r_1(X)$  coïncide avec la notion usuelle du reach de  $X$ .



Etant donnée une forme  $X$  dont le  $\mu$ -reach est supérieur ou égal à  $R > 0$  et un ensemble fini de points  $P$  tel que  $d_H(P, X) \leq \varepsilon$ , on calcule la valeur de plus grand ratio  $\frac{\varepsilon}{R}$  pour lequel le complexe Čech  $\text{Cech}(P, \alpha)$  ou le complexe Rips  $\text{Rips}(P, \alpha)$  fournissent une reconstruction ayant le même type d'homotopie que  $X$  pour une certaine valeur du paramètre  $\alpha$ . Les calculs pour obtenir ces valeurs ont été réalisés grâce à un logiciel de calcul formel. Dans la section C, nous donnons tous les détails pour le calcul du ratio lorsque  $\mu = 1$ ,  $R = 1$  et  $D = +\infty$ .

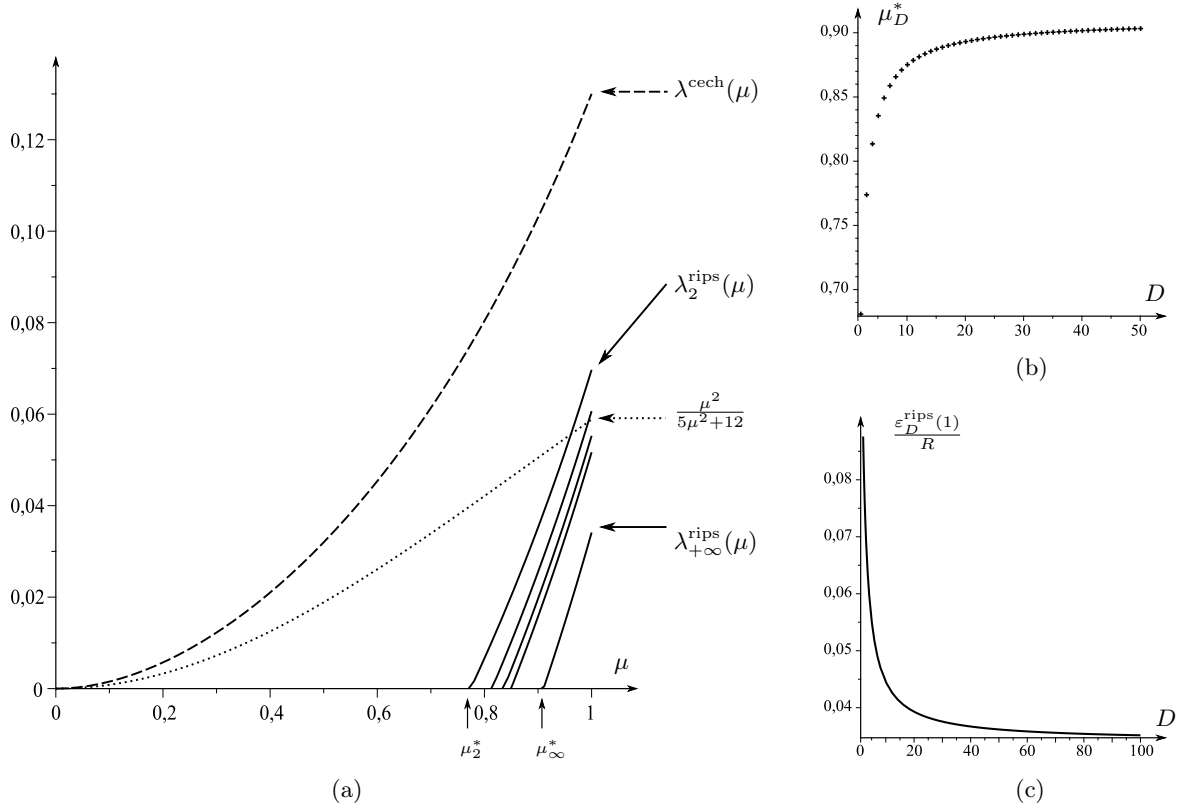


FIGURE 2.10 – (a) Meilleurs ratios  $\frac{\varepsilon}{R}$  que l'on obtient pour une reconstruction correcte d'une forme à  $\mu$ -reach positif avec le complexe de Čech ou le complexe de Rips pour  $D \in \{2, 3, 4, 5, +\infty\}$ ; comparaison avec le ratio  $\frac{\mu^2}{5\mu^2+12}$  obtenu dans [CCSL09]. (b) et (c)  $\mu_D^*$  et  $\lambda_D^{\text{rips}}(1)$  tracées en fonction de  $D$ .

**Reconstruction avec le complexe de Čech** Notons que l'hypothèse  $r_\mu(X) \geq R$  équivaut à l'hypothèse  $\chi_X(t) \geq \mu$  pour tout  $t \in (0, R)$ . Une forme  $X$  avec  $r_\mu(X) \geq R$  vérifie donc les hypothèses du lemme 2.6 et possède alors une fonction de défaut de convexité  $h_X$  bornée par une fonction dépendant uniquement de  $R$  et  $\mu$ . En insérant cette borne dans le théorème 2.4, on obtient que si  $\alpha + 3\varepsilon \leq R$  alors le complexe de Čech  $\text{Cech}(P, \alpha)$  a le même type d'homotopie que  $X^\eta$  pour tout  $0 < \eta < R$  dès lors que l'inégalité suivante tient pour tout  $t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$  :

$$\frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left(\frac{t}{R}\right)^2}}{\mu(2 - \mu)} R < t - 3\varepsilon.$$

En éliminant la racine carrée, on peut remplacer l'inégalité ci-dessus par  $H_{\mu,\varepsilon}(t) < 0$  où  $H_{\mu,\varepsilon}(t)$  est un polynôme de degré 2 en  $t$ . Il s'ensuit que l'inégalité est vérifiée quand la différence absolue entre les deux racines  $t_\mu^1(\varepsilon) \leq t_\mu^2(\varepsilon)$  de  $H_{\mu,\varepsilon}(t)$  est plus grande que  $2\varepsilon$ . Lorsque c'est le cas, les valeurs admissibles pour  $\alpha$  sont situées dans l'intervalle  $I_\mu(\varepsilon) = [t_\mu^1(\varepsilon) - 3\varepsilon, t_\mu^2(\varepsilon) - \varepsilon]$ . La condition  $t_\mu^2(\varepsilon) - t_\mu^1(\varepsilon) > 2\varepsilon$  peut être réécrite comme la positivité d'un polynôme de degré 2 en  $\varepsilon$  avec deux racines, l'une positive, l'autre négative. Ainsi, la condition est vérifiée tant que  $\varepsilon$  est inférieur à la racine positive dont la valeur divisée par  $R$  est :

$$\lambda^{\text{cech}}(\mu) = \frac{-3\mu + 3\mu^2 - 3 + \sqrt{-8\mu^2 + 4\mu^3 + 18\mu + 2\mu^4 + 9 + \mu^6 - 4\mu^5}}{-7\mu^2 + 22\mu + \mu^4 - 4\mu^3 + 1}.$$

On obtient finalement le théorème de reconstruction suivant :

**Théorème 2.6.** *Soit  $P$  un ensemble fini de points et  $X$  un compact dont le  $\mu$ -reach  $R$  est positif. Si*

$$d_H(P, X) \leq \varepsilon < \lambda^{\text{cech}}(\mu) R$$

*alors  $\text{Cech}(P, \alpha)$  et  $X^\eta$  partagent le même type d'homotopie pour tout  $\eta \in (0, R)$  et  $\alpha \in I_\mu(\varepsilon)$ .*

Il est intéressant de noter que  $\lambda^{\text{cech}}(\mu)$  ne dépend pas de la dimension ambiante  $D$ . En traçant  $\lambda^{\text{cech}}(\mu)$  comme une fonction de  $\mu$  (voir la figure 2.10(a)), on observe que cette valeur est positive pour tout  $\mu \in (0, 1]$  et on améliore la borne supérieure  $\frac{\mu^2}{5\mu^2+12}$  établie dans [CCSL09]. Cependant, pour  $\mu = 1$ , on obtient le ratio  $\lambda^{\text{cech}}(1) = \frac{-3+\sqrt{22}}{13} \approx 0.13$  qui est moins bon que le ratio  $3 - \sqrt{8} \approx 0.17$  obtenu dans [NSW08].

**Reconstruction avec le complexe de Rips** En combinant le théorème 2.5 avec  $\beta = \alpha$  et le lemme 2.6, on obtient que si  $\vartheta_D \alpha + \varepsilon \leq R$  alors le complexe  $\text{Rips}(P, \alpha)$  a le même type d'homotopie que  $X^\eta$  pour tout  $0 < \eta < R$  dès lors que

$$\frac{1 + \mu(1 - \mu) - \sqrt{1 - \mu(2 - \mu) \left( \frac{\vartheta_D \alpha + \varepsilon}{R} \right)^2}}{\mu(2 - \mu)} R < 2\alpha - \vartheta_D \alpha - 2\varepsilon.$$

Comme précédemment, nous pouvons éliminer la racine carrée et remplacer l'inégalité ci-dessus par  $H_\mu(\varepsilon, \alpha) < 0$  où  $H_\mu(\varepsilon, \alpha)$  est un polynôme de degré 2 en  $\varepsilon$  et  $\alpha$ . Afin de trouver la plus grande valeur pour  $\varepsilon$  telle que  $H_\mu(\varepsilon, \alpha) < 0$ , on peut supposer que  $\frac{\partial H_\mu(\varepsilon, \alpha)}{\partial \alpha} = 0$ . En insérant la valeur de  $\alpha$  pour laquelle  $\frac{\partial H_\mu(\varepsilon, \alpha)}{\partial \alpha} = 0$  dans  $H_\mu(\varepsilon, \alpha)$ , nous obtenons un polynôme de degré 2 en  $\varepsilon$  dont la plus grande racine  $\varepsilon_D^{\text{rips}}(\mu)$  donne la borne supérieure de  $\varepsilon$  pour laquelle l'inégalité ci-dessus est vérifiée. Posons  $\lambda_D^{\text{rips}}(\mu) = \frac{\varepsilon_D^{\text{rips}}(\mu)}{R}$ . En utilisant un logiciel de calcul formel, on obtient la valeur suivante pour  $\lambda_D^{\text{rips}}(\mu)$  :

$$\lambda_D^{\text{rips}}(\mu) = \frac{(2 - \vartheta_D)(2 + 2\sqrt{2} - \vartheta_D)\sqrt{2 - \mu - \mu^2} - (4 + 2\sqrt{2} - \sqrt{2}\vartheta_D)(1 + \mu - \mu^2)}{\mu(2 - \mu)(12 + 8\sqrt{2} - 4\vartheta_D - 4\vartheta_D\sqrt{2} + (\vartheta_D)^2)}.$$

La figure 2.10(a) montre un tracé de  $\lambda_D^{\text{rips}}(\mu)$  en fonction de  $\mu$ . Nous observons que ce ratio est positif seulement sur un sous-intervalle  $(\mu_D^*, 1]$  de  $(0, 1]$ . Nous pouvons donc seulement garantir que les complexes de Rips fournissent une reconstruction correcte de formes à  $\mu$ -reach positif lorsque  $\mu > \mu_D^*$ . Sur la figure 2.10(b), nous traçons  $\mu_D^*$  comme fonction de  $D$ .  $\mu_D^*$  augmente avec  $D$  et, grâce à un logiciel de calcul formel, nous pouvons montrer que  $\mu_D^*$  tend vers  $\sqrt{2\sqrt{2}-2} \approx 0.91$  quand  $D \rightarrow +\infty$ . Sur la figure 2.10(c), nous traçons  $\lambda_D^{\text{rips}}(1)$  comme fonction de  $D$ . Ce ratio décroît avec  $D$  et nous avons pu prouver similairement que  $\lim_{D \rightarrow +\infty} \lambda_D^{\text{rips}}(1) = \frac{2\sqrt{2-\sqrt{2}-\sqrt{2}}}{2+\sqrt{2}} \approx 0.034$ .

Intéressons-nous maintenant à l'intervalle des valeurs du paramètre  $\alpha$  possible pour la reconstruction. Pour cela, supposons que  $\mu_D^* < \mu$  et  $\varepsilon < \lambda_D^{\text{rips}}(\mu)$ . On montre avec du calcul formel que l'inégalité  $H_\mu(\varepsilon, \alpha) < 0$  est vérifiée pour  $\alpha \in I_D(\mu, \varepsilon, R) = (\alpha_D^{\min}(\mu, \varepsilon, R), \alpha_D^{\max}(\mu, \varepsilon, R))$  où  $I_D(\mu, \varepsilon, R)$  est un intervalle non-vide. L'expression analytique  $I_D(\mu, \varepsilon, R)$  étant volumineuse, nous la donnons seulement lorsque  $D \rightarrow \infty$  et  $\mu = 1$  :

$$\begin{cases} \alpha_\infty^{\min}(1, \varepsilon, R) = \frac{1}{4} \left( R - (\sqrt{2} - 1)\varepsilon - \sqrt{\Delta} \right) \\ \alpha_\infty^{\max}(1, \varepsilon, R) = \frac{1}{4} \left( R - (\sqrt{2} - 1)\varepsilon + \sqrt{\Delta} \right) \end{cases} \text{ avec } \Delta = -(17+12\sqrt{2})\varepsilon^2 - (14+10\sqrt{2})R\varepsilon + R^2$$

Nous obtenons alors le théorème suivant de reconstruction suivant :

**Théorème 2.7.** *Soit  $P$  un ensemble fini de points et  $X$  un compact dont le  $\mu$ -reach  $R$  est positif. Si*

$$\mu_D^* < \mu \quad \text{et} \quad d_H(P, X) \leq \varepsilon < \lambda_D^{\text{rips}}(\mu) R$$

alors  $\text{Rips}(P, \alpha)$  et  $X^\eta$  partagent le même type d'homotopie pour tout  $\eta \in (0, R)$  et  $\alpha \in I_D(\mu, \varepsilon, R) = (\alpha_D^{\min}(\mu, \varepsilon, R), \alpha_D^{\max}(\mu, \varepsilon, R))$ .

Les deux extrémités de l'intervalle  $I_\infty(\mu, \varepsilon, 1)$  sont tracées en fonction de  $\mu$  et  $\varepsilon$  sous forme de deux surfaces sur la figure 2.12. La figure 2.11 représente les coupes de ces deux surfaces pour  $\mu = 0.92$  et  $\mu = 1$ , c'est à dire les deux fonctions  $(\alpha_\infty^{\min}(\mu, \cdot, 1), \alpha_\infty^{\max}(\mu, \cdot, 1))$  pour  $\mu = 0.92$  et  $\mu = 1$ . Conformément à ce que suggèrent ces tracés, on a pu montrer avec un logiciel de calcul formel que, lorsque  $\mu \rightarrow \mu_\infty^*$ ,  $\varepsilon_\infty^{\text{rips}}(\mu)$  tend vers 0 et  $I_\infty(\mu, \varepsilon, \alpha)$  tend vers un singleton  $\{\alpha'R\}$  où  $\alpha' \approx 0,27$ . Finalement, on a pu montrer que, lorsque  $\varepsilon \ll 1$  et  $\mu > \mu_\infty^*$ , on a :

$$\begin{cases} \alpha_\infty^{\min}(\mu, \varepsilon, R) = \varphi(\mu)R + O(\varepsilon) \\ \alpha_\infty^{\max}(\mu, \varepsilon, R) = \psi(\mu)R + O(\varepsilon) \end{cases}$$

Où  $\varphi$  est une fonction continue décroissante vérifiant  $\varphi(0) \approx 0.27$  et  $\varphi(1) = 0$  et  $\psi$  est une fonction continue croissante avec  $\psi(0) \approx 0.27$  et  $\psi(1) = \frac{1}{2}$ . De manière surprenante, on note que  $\alpha_D^{\min}(\mu, \varepsilon, R)$  ne tend pas vers 0 lorsque  $\varepsilon \rightarrow 0$  et  $\mu_D^* < \mu < 1$ .

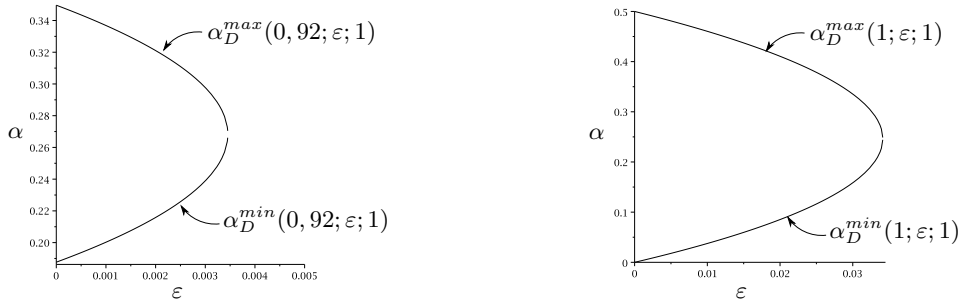


FIGURE 2.11 – Tracés des deux fonctions  $(\alpha_D^{\min}(\mu; \cdot; 1), \alpha_D^{\max}(\mu; \cdot; 1))$  pour  $\mu = 0,92$  (gauche) et  $\mu = 1$  (droite) lorsque  $D \rightarrow \infty$ .

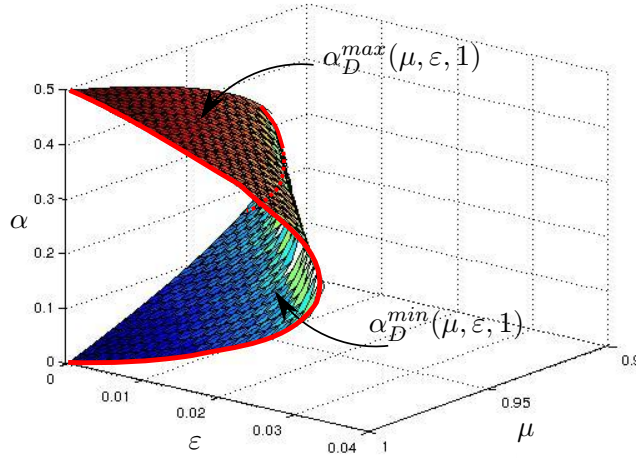


FIGURE 2.12 – Tracé de  $\alpha_D^{\min}(\mu, \varepsilon, 1)$  et  $\alpha_D^{\max}(\mu, \varepsilon, 1)$  en fonction de  $\mu$  et  $\varepsilon$  lorsque  $D \rightarrow \infty$ . Les deux coupes selon  $\mu = 0,92$  et  $\mu = 1$  de la figure 2.11 sont représentées en rouge. Etant donné  $\mu$  et  $\varepsilon$ , il faut que  $\alpha$  soit entre ces deux surfaces pour pouvoir garantir le type d'homotopie du complexe de Rips.

#### 2.4.4 Reconstruction avec le complexe de Rips « intrinsèque »

Les complexes de Rips étudiés dans ce chapitre sont des complexes de Rips « extrinsèques », au sens où on considère la distance euclidienne ambiante pour les calculer. Lorsque la forme échantillonnée est une variété, il est également possible de considérer des complexes de Rips « intrinsèques » en prenant la distance géodésique de la variété pour les calculer. Il a été montré que le Rips intrinsèque a le type d'homotopie d'une variété échantillonnée lorsque l'échantillonnage est inclus dans la variété et qu'il possède une distance de Gromov-Hausdorff à la variété nulle [Hau95] ou suffisamment faible [Lat01]. Latschev indique qu'il est possible de calculer des constantes de reconstruction à partir de sa preuve. Il semble également possible d'extraire des constantes de reconstruction pour le Rips intrinsèque avec les preuves mises en œuvres dans ce chapitre. En effet, même si nous avons considéré des complexes de Rips extrinsèques, le lemme A.2 indique que les hypothèses du théorème 2.3 sont stables sous faible

perturbation métrique. Il semble donc plausible de pouvoir utiliser les quasi-Rips complexes afin de montrer un résultat analogue au théorème 2.3 dans le cadre de Rips intrinsèque. Cela pourrait permettre de montrer que le complexe de Rips intrinsèque a le même type d'homotopie qu'une variété lorsqu'il est construit à partir d'un ensemble de points ayant une faible distance de Gromov-Hausdorff à celle-ci. Il serait intéressant d'obtenir des constantes de reconstruction dans ce cadre, en adaptant notre méthode au cas du Rips intrinsèque, afin de pouvoir les comparer ensuite aux constantes de reconstruction pouvant être extraites de la preuve de Latschev.

# Structure de données complexes simpliciaux

## Sommaire

<b>3.1</b>	<b>Introduction</b>	<b>35</b>
<b>3.2</b>	<b>Structure de données squelette/bloqueurs</b>	<b>37</b>
<b>3.3</b>	<b>Taille du complexe de Rips</b>	<b>39</b>
3.3.1	Bornes théoriques	39
3.3.2	Evaluation expérimentale	41
<b>3.4</b>	<b>Opérations élémentaires sur la structure de données squelette/bloqueurs</b>	<b>41</b>
3.4.1	Algorithmes de complexité polynomiale en la taille du graphe et le nombre de bloqueurs	42
3.4.2	Algorithmes de conversion entre structures de données	48
<b>3.5</b>	<b>Conclusion et perspectives</b>	<b>49</b>

## 3.1 Introduction

Les complexes simpliciaux sont intensément utilisés pour représenter des objets géométriques notamment dans l'étude et la simulation d'éléments finis ou encore l'informatique graphique. Ces applications ont motivé l'introduction de nombreuses structures de données à même de représenter les complexes simpliciaux de dimension 2 et 3 aussi appelés *triangulations* et *tétraédrisation*. Parmi ces structures, on peut citer les structures dites *Winged-Edge* [Bau72], *DCEL* [MP78] ou encore *Quad edge* [GS85] pour représenter les triangulations. Les structures dites *Facet-Edge* [DL87] et *Handle-Face* [LT97] permettent quant-à-elles de représenter les tétraédrisations<sup>1</sup>. Ces structures stockent plus d'informations que les seules arêtes du complexe en incluant des relations d'incidence supplémentaires comme par exemple, pour chaque arête du complexe, la liste des faces la contenant. Ces informations supplémentaires permettent de réaliser efficacement certaines opérations classiques de géométrie algorithmique, comme le calcul d'une enveloppe convexe. Ces structures ont l'inconvénient de mal se généraliser à des complexes simpliciaux de plus grandes dimensions. Or, l'analyse de données de grande dimension nécessite des structures capables de représenter des complexes simpliciaux de dimension

1. La bibliothèque *CGAL* implémente de telles structures pour les triangulations et tétraédrisations [cga].

arbitraire. Plusieurs extensions ont été proposées notamment le  $n$ - $G$ -map de Lienhardt [Lie91] ou la structure de données *cell-tuple* de Brisson [Bri89]. Ces structures sont conçues pour représenter des complexes simpliciaux homéomorphes à une variété. D'autres structures de données permettent de représenter des complexes simpliciaux arbitraires comme le graphe d'incidence [Ede87], le graphe d'incidence simplifié [DFGH04], ou plus récemment l'arbre simplicial de Boissonnat et Maria [BM12b]. Toutes ces méthodes stockent l'ensemble des simplexes d'un complexe simplicial ainsi qu'une représentation des relations d'incidence entre ceux-ci. Par exemple, Boissonnat et Maria proposent de représenter un complexe simplicial par un arbre dont les nœuds sont en bijection avec l'ensemble des simplexes du complexe. La relation d'incidence est alors stockée implicitement dans cet arbre.

Cependant, les complexes de grande dimension présentent souvent un nombre gigantesque de simplexes et de faces maximales rendant prohibitive toute représentation explicite de ces derniers dans beaucoup d'applications. Ce problème motive une autre approche consistant à représenter implicitement les simplexes dans une représentation compacte. Zomorodian a développé le *Tidy set* dans ce but [Zom10b]. Partant d'un *flag-complexe*, il propose une méthode permettant de simplifier ce complexe en un ensemble simplicial dont on peut ensuite calculer les groupes d'homologie. Son approche évite de calculer l'ensemble des simplexes : tous les calculs sont faits uniquement à partir d'un graphe encodant le 1-squelette du complexe. Expérimentalement, l'auteur remarque que la taille de cet ensemble est souvent comparable au nombre de sommets du complexe initial (qui est largement inférieure au nombre de simplexes du complexe). Les groupes d'homologie peuvent alors être calculés efficacement. Néanmoins, un inconvénient de cette méthode est qu'il est difficile d'extraire un complexe simplicial après simplification. Nous proposons une représentation qui a le même but que le Tidy Set : représenter de manière compacte les complexes proches de *flag-complexes*. Nous introduisons pour cela une structure capable de représenter de manière efficace un complexe en stockant le graphe de son 1-squelette ainsi que ses *bloqueurs*. Les bloqueurs<sup>2</sup> sont les non-faces minimales du complexe : ils représentent les différences du complexe avec son expansion. Lorsque le complexe est un *flag-complexe*, il ne possède pas de bloqueurs. Nous verrons qu'il est alors possible de réaliser plusieurs opérations simpliciales en utilisant uniquement le 1-squelette dans notre structure de données et sans énumérer les simplexes de dimension supérieure. Lorsque le complexe n'est pas un *flag-complexe* et qu'il possède un nombre constant de bloqueurs, les complexités de ces opérations simpliciales sont polynomiales en la taille du 1-squelette du complexe. Enfin, contrairement à la construction de Zomorodian, notre structure encode un complexe simplicial : il est donc possible d'extraire l'ensemble des simplexes à tout moment de notre structure. Réciproquement, il est également possible de calculer efficacement la représentation d'un complexe simplicial dans notre structure de données étant donné un ensemble de simplexes.

Nous commençons par définir notre structure de données. Une comparaison entre le nombre de simplexes et la taille du 1-squelette de plusieurs complexes simpliciaux est ensuite menée afin de justifier l'introduction de notre structure de données. Nous présentons ensuite différentes opérations possibles sur celle-ci.

---

2. Les bloqueurs sont également appelés face manquantes (*missing faces*) ou simplexes vides (*empty simplices*) dans la littérature [Nag08].

Ce chapitre reprend une partie des articles [ALS11a, ALS12a] co-écrits avec Dominique Attali et André Lieutier. Outre une modification sur la forme, ont également été ajoutés les algorithmes permettant de collapser un simplexe de toute dimension, d'énumérer l'ensemble des simplexes et enfin de calculer la représentation squelette/bloqueurs d'un complexe simplicial étant donné son ensemble de simplexes.

## 3.2 Structure de données squelette/bloqueurs

Dans cette section, nous introduisons une nouvelle représentation pour les complexes simpliciaux. Pour cela, nous stockons le 1-squelette  $G$  de  $K$  ainsi qu'un ensemble de simplexes appelés *bloqueurs* qui indiquent à quel point  $K$  diffère du *flag*-complexe de  $G$  (voir Figures 3.2 et 3.3).

**Définition 3.1.** Soit  $i \geq 0$ . Un simplexe  $\sigma \subset \text{Vert}(K)$  est un bloqueur d'ordre  $i$  de  $K$  s'il satisfait les trois conditions suivantes (1)  $\dim \sigma > i$ ; (2)  $\sigma$  n'appartient pas à  $K$ ; (3) toutes les faces propres de  $\sigma$  appartiennent à  $K$ . L'ensemble des bloqueurs d'ordre  $i$  de  $K$  est noté  $\text{Blockers}_i(K)$ .

De manière équivalente, les bloqueurs d'ordre  $i$  de  $K$  sont les simplexes minimaux pour l'inclusion de  $\text{Expand}(K^{(i)}) \setminus K$ ; voir Figures 3.1 et 3.2. Une conséquence importante est que la paire  $(K^{(i)}, \text{Blockers}_i(K))$  encode entièrement le complexe simplicial  $K$ . En effet, le complexe simplicial ayant  $S$  pour  $i$ -squelette et  $B$  comme ensemble de bloqueurs d'ordre  $i$  peut être récupéré à partir de la paire  $(S, B)$  grâce à la formule suivante :

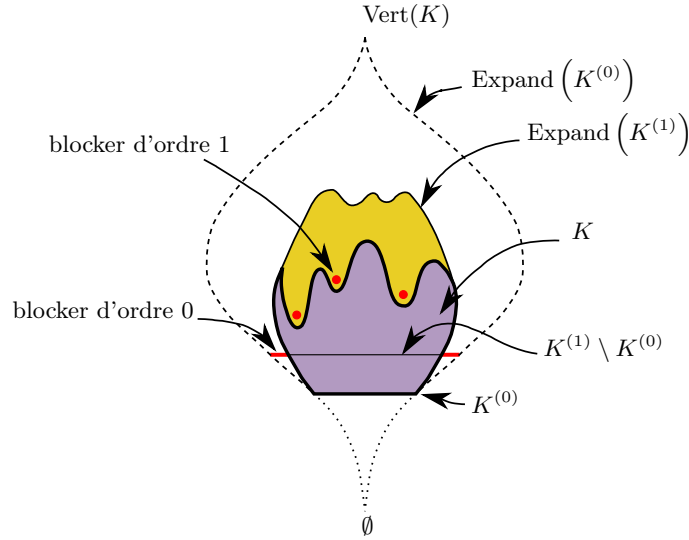
$$K = \{\sigma \in \text{Expand}(S) \mid \sigma \text{ n'a pas de face dans } B\}. \quad (3.1)$$

Par la suite, les bloqueurs d'ordre 1 seront simplement désignés par bloqueurs.

Les complexes simpliciaux qui nous intéresseront seront toujours "proches" de *flag*-complexes. Autrement dit, ils comporteront peu de bloqueurs. Nous proposons de représenter ces complexes avec la paire  $(K^{(1)}, \text{Blockers}_1(K))$ . Nous verrons que cette structure de données est particulièrement adaptée dans ce cas puisque la complexité des opérations élémentaires sera polynomiale en la taille du 1-squelette en supposant que le nombre de bloqueurs est constant.

Nous allons maintenant détailler plus formellement notre structure de données. Soit  $\mathcal{N}(v) = \mathcal{N}_K(v)$  l'ensemble des sommets  $w \neq v$  tels que  $vw \in K$  et soit  $\mathcal{B}(v) = \mathcal{B}_K(v)$  l'ensemble des bloqueurs contenant  $v$ . Encoder la paire  $(K^{(1)}, \text{Blockers}_1(K))$  revient à encoder la paire  $(\mathcal{N}(v), \mathcal{B}(v))$  pour chaque sommet  $v$  de  $K$ . Notre structure de données contient donc un tableau  $V$  pour les sommets et stocke pour chaque sommet  $v$  l'ensemble des voisins  $\mathcal{N}(v)$  ainsi qu'un ensemble de pointeurs vers les bloqueurs de  $\mathcal{B}(v)$  comme illustré par la figure 3.3. La taille de notre structure de données est donc  $\sum_{v \in \text{Vert}(K)} (1 + \#\mathcal{N}(v) + 2\#\mathcal{B}(v))$ . En effet,



FIGURE 3.1 – Diagramme de Hasse de  $K$ .

supposons que chaque sommet d'un bloqueur paye 1 pour son sommet correspondant dans  $V$ . A la fin de cette opération, le coût pour chaque sommet de  $V$  est  $\#\mathcal{B}(v)$ .

Une borne supérieure grossière sur la dimension et le nombre de bloqueurs d'un complexe simplicial comportant  $n$  sommets peut être obtenue facilement<sup>3</sup>. Considérons un bloqueur  $\sigma$  passant par un sommet  $v$ . Puisque  $\sigma \subset \{v\} \cup \mathcal{N}(v)$ , on obtient que  $\#\sigma \leq 1 + \#\mathcal{N}(v)$  et donc  $\dim \sigma \leq \max_{v \in \sigma} \#\mathcal{N}(v)$ . On en déduit que  $N_{\max} = \max_{v \in \text{Vert}(K)} \#\mathcal{N}(v)$  est une borne supérieure sur la dimension des bloqueurs et que  $O(2^{N_{\max}})$  est une borne supérieure sur le nombre de bloqueurs passant par  $v$ . Le nombre total de bloqueurs dans notre structure de données est donc  $O(n2^{N_{\max}})$ .

Il est possible d'obtenir de meilleures bornes dès lors que l'on possède en plus des informations géométriques. Par exemple, supposons que  $K$  soit le complexe de Čech d'un ensemble de points  $P \subset \mathbb{R}^D$  avec un paramètre  $\alpha$ . Dans ce cas, Xavier Goac a suggéré qu'une borne moins grossière peut être obtenue avec le théorème de Helly. En effet, ce théorème implique clairement qu'il n'est pas possible d'avoir de bloqueurs de dimension strictement plus grande que  $D$ . L'existence d'un bloqueur  $\sigma$  de dimension  $\dim \sigma = \#\sigma - 1 > D$  impliquerait l'existence d'une famille de convexes violant le théorème de Helly. Pour s'en convaincre par l'absurde, prenons  $\sigma$  un bloqueur de  $\text{Cech}(P, \alpha)$  de dimension strictement supérieure à  $D$  et notons  $\mathcal{B}$  les boules centrées en les sommets de  $\sigma$  de rayon  $\alpha$ . Notons que chaque face de  $\sigma$  de dimension  $\dim \sigma - 1$  appartient à  $\text{Cech}(P, \alpha)$  par définition d'un bloqueur. De manière équivalente, l'intersection de  $\dim \sigma$  boules parmi l'ensemble des boules de  $\mathcal{B}$  est non vide. Ainsi, le théorème de Helly assure que l'intersection de toutes les boules de  $\mathcal{B}$  est non vide puisque  $\dim \sigma > D$  d'où  $\sigma \in \text{Cech}(P, \alpha)$ ; cela rentre en contradiction avec le fait que  $\sigma$  est un bloqueur. On sait donc que la dimension de tout bloqueur de  $\text{Cech}(P, \alpha)$  est bornée par  $D$ .

3. Des bornes plus fines sur le nombre de bloqueurs ont été proposées dans le cas d'un polytope [Nag08].

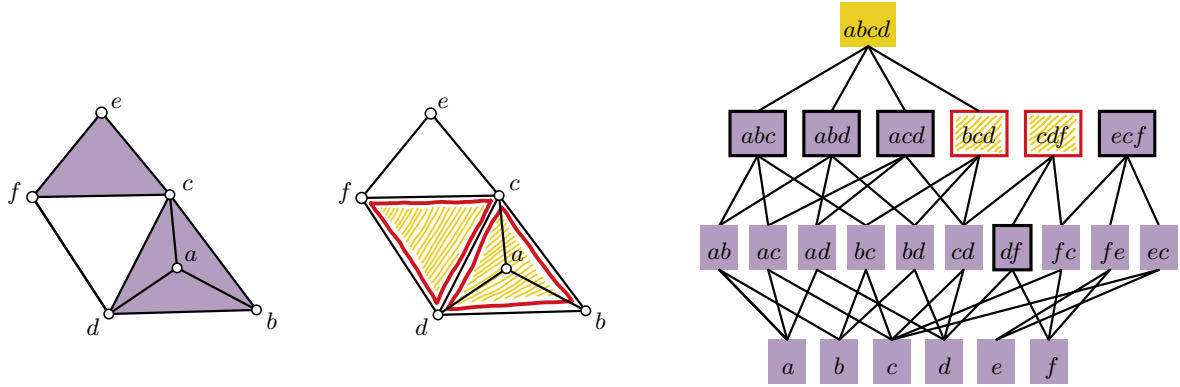


FIGURE 3.2 – Gauche : Un complexe simplicial constitué de six sommets, dix arêtes et quatre triangles. Milieu : le même complexe simplicial représenté cette fois-ci par son 1-squelette et l'ensemble de ses bloqueurs d'ordre 1  $\{cdf, bcd\}$ . Droite : Diagramme de Hasse de l'expansion du 1-squelette. Les nœuds en violet appartiennent au complexe simplicial. Les bloqueurs et les simplexes maximaux pour l'inclusion sont représentés par des nœuds encadrés.

### 3.3 Taille du complexe de Rips

Afin de motiver le besoin de représenter les simplexes dans une forme compacte, nous calculons dans un premier temps une borne inférieure et une borne supérieure sur le nombre de simplexes et la taille du 1-squelette d'un complexe de Rips approximant une forme. Dans un deuxième temps, nous comptons expérimentalement le nombre de simplexes et la taille d'un graphe de plusieurs complexes de Rips construits à partir de nuages de points.

#### 3.3.1 Bornes théoriques

Nous allons établir une borne sur le nombre de simplexes et la taille du graphe d'un complexe de Rips approximant une forme. Pour cela nous supposons que le complexe de Rips est construit avec un paramètre  $\alpha$  à partir d'un ensemble de points  $P$  échantillonnant une variété  $A$  de manière ni trop dense, ni trop clairsemée. Plus précisément, nous supposons que chaque boule centrée en un point de  $A$  de rayon  $\varepsilon$  contient au moins un point de  $P$  et au plus  $\kappa$  points de  $P$  pour  $\varepsilon \leq \alpha$  et  $\kappa > 0$ . Nous exprimons nos bornes inférieure et supérieure en utilisant respectivement les nombres de covering et packing dont nous rappelons la définition (qui peut être trouvée dans [Cla06b]).

Notons  $B(x, r)$  la boule fermée de centre  $x$  et de rayon  $r$  et soit  $U \subset \mathbb{R}^D$ . Un  $\varepsilon$ -cover de  $U$  est un ensemble  $Y \subset U$  tel que l'union des boules centrées en  $Y$  de rayon  $\varepsilon$  contient  $U$ . Un  $\varepsilon$ -packing de  $U$  est un ensemble  $Y \subset U$  avec la propriété que les boules centrées en  $Y$  de rayon  $\varepsilon$  sont deux à deux disjointes. Notons  $\mathcal{C}(x, r, \varepsilon)$  la taille du plus petit  $\varepsilon$ -cover de  $X \cap B(x, r)$  et notons également  $\mathcal{P}(x, r, \varepsilon)$  la taille du plus grand  $\varepsilon$ -packing de  $X \cap B(x, r)$ . Nous définissons  $\mathcal{C}(r, \varepsilon) = \max_{x \in X} \mathcal{C}(x, r, \varepsilon)$  et  $\mathcal{P}(r, \varepsilon) = \min_{x \in X} \mathcal{P}(x, r, \varepsilon)$ . Etant donnée  $X$  une  $d$ -variété à reach positif  $\rho > 0$ , les quantités  $\mathcal{C}(r, \varepsilon)$  et  $\mathcal{P}(r, \varepsilon)$  sont toutes les deux en  $\Theta(r^d/\varepsilon^d)$  quand

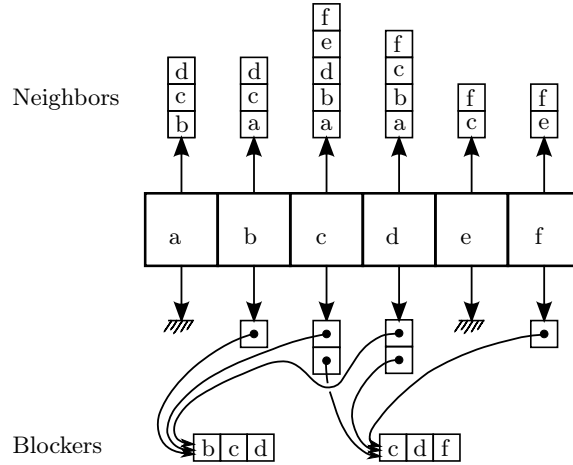


FIGURE 3.3 – Représentation de la structure de données pour le complexe simplicial de la Figure 3.2.

$\varepsilon \rightarrow 0$  et  $r < \rho$ . La constante cachée dans  $\Theta$  dépend seulement de la dimension intrinsèque  $d$ ; elle ne dépend pas de la dimension ambiante  $D$ .

Pour majorer la taille de notre structure de données, notons que le nombre de voisins de  $v$  est borné par

$$N_{\max} = \max_{v \in \text{Vert}(\text{Rips}(P, \alpha))} \#\mathcal{N}(v).$$

Puisque le complexe de Rips est un *flag-complexe*, il ne comporte pas de bloqueurs ainsi la taille initiale de notre structure de données est  $\sum_{v \in \text{Vert}(\text{Rips}(P, \alpha))} (1 + \#\mathcal{N}(v)) \leq n(1 + N_{\max})$  où  $n$  désigne le nombre de points de  $P$ . Notons que les voisins d'un sommet  $v$  du complexe  $\text{Rips}(P, \alpha)$  sont les points de  $P \setminus \{v\}$  dans la boule  $B(v, 2\alpha)$ . Cette boule peut être couverte avec  $\mathcal{C}(v, 2\alpha, \varepsilon)$  boules de rayon  $\varepsilon$  centrées en un point de  $A \cap B(v, 2\alpha)$ , chacune d'entre elles contenant au plus  $\kappa$  points de  $P$ . On en déduit que  $\#\mathcal{N}(v) \leq \kappa \mathcal{C}(v, 2\alpha, \varepsilon)$  et  $N_{\max} \leq \kappa \mathcal{C}(2\alpha, \varepsilon)$ . La taille de notre structure de données est donc bornée par  $n$  fois une quantité dépendant du ratio  $\alpha/\varepsilon$ , de la constante  $\kappa$  et de la dimension intrinsèque  $d$ . En particulier, cette borne ne dépend pas de la dimension  $D$ .

Nous donnons maintenant une borne inférieure et une borne supérieure sur le nombre de simplexes du complexe de Rips. Pour cela, nous allons commencer par borner le nombre de  $k$ -simplexes. Soit un  $k$ -simplexe  $\sigma$  contenant le sommet  $v$ . Si  $\sigma \in \text{Rips}(P, \alpha)$  alors l'ensemble de ses sommets est contenu dans  $B(v, 2\alpha)$ . De plus, si l'ensemble des sommets de  $\sigma$  est contenu dans  $B(v, \alpha)$  alors  $\sigma \in \text{Rips}(P, \alpha)$ . Autrement dit, l'ensemble des  $k$ -simplexes de  $\text{Rips}(P, \alpha)$  ayant  $v$  comme sommet contient l'ensemble des  $k$ -simplexes obtenus en choisissant  $v$  et  $k$  voisins distincts de  $v$  dans  $\text{Rips}(P, \alpha/2)$ .

Soit

$$N'_{\min} = \min_{v \in \text{Vert}(\text{Rips}(P, \alpha/2))} \#\mathcal{N}(v).$$

Si on note  $n$  le nombre de sommets de  $P$ , on peut alors déduire une borne inférieure et une

### 3.4. Opérations élémentaires sur la structure de données squelette/bloqueurs 41

borne supérieure sur le nombre de  $k$ -simplexes :

$$\frac{n}{k+1} \binom{N'_{\min}}{k} \leq \#\{\sigma \in \text{Rips}(P, \alpha) \mid \dim \sigma = k\} \leq \frac{n}{k+1} \binom{N_{\max}}{k}.$$

En sommant sur toutes les dimensions, on obtient que  $\frac{n}{N'_{\min}+1} 2^{N'_{\min}} \leq \#\text{Rips}(P, \alpha) \leq n 2^{N_{\max}}$ . Nous avons vu que  $N_{\max} \leq \kappa \mathcal{C}(2\alpha, \varepsilon)$ . Nous allons maintenant donner une borne inférieure de  $N'_{\min}$ . On rappelle que les voisins d'un sommet  $v$  dans  $\text{Rips}(P, \alpha/2)$  sont les points de  $P \setminus \{v\}$  inclus dans la boule  $B(v, r)$ . Cette boule peut être couverte avec  $\mathcal{P}(v, \alpha - \varepsilon, \varepsilon)$  boules deux à deux disjointes de rayon  $\varepsilon$  centrées en un point de  $A \cap B(v, \alpha - \varepsilon)$ , chacune d'entre elles contenant au moins un point de  $P$ . On en déduit que  $N'_{\min} \geq \mathcal{P}(\alpha - \varepsilon, \varepsilon) - 1$ .

Pour conclure, si  $N$  est le nombre typique de voisins dans le graphe de proximité, une approche consistant à énumérer tous les simplexes devra considérer un nombre d'éléments  $\frac{n}{N+1} 2^N$ , ce nombre est à comparer avec la taille de notre structure de données qui est grossièrement de  $O(nN)$ . Le figure 3.4, illustre la différence d'ordre de grandeur entre le nombre de simplexes et la taille du 1-squelette.

#### 3.3.2 Evaluation expérimentale

Nous évaluons maintenant le nombre de simplexes et la taille du graphe de plusieurs complexes de Rips construits à partir de notre jeu de données (la description de notre jeu de données se trouve dans la section 4.4.1 du chapitre 4). La figure 3.4, représente le nombre de  $k$ -simplexes pour un complexe de Rips construit à partir d'un ensemble de points qui échantillonne la surface d'un 2-cube plongé dans  $\mathbb{R}^3$ . On constate que le nombre de simplexes dans ce cas simple est déjà gigantesque, rendant difficile toute représentation ayant une taille similaire au nombre de simplexes du complexe représenté. Pour conclure notre comparaison entre notre structure de données et les structures stockant tous les simplexes, nous présentons dans la table 4.1 du chapitre 4 plusieurs statistiques sur les complexes de Rips construits à partir de notre jeu de données notamment le nombre de simplexes et la taille du graphe. Dans chaque cas, l'ensemble des points  $P$  échantillonne une variété  $A$  et le paramètre  $\alpha$  du Rips est choisi de telle sorte que  $\text{Rips}(P, \alpha) \simeq A$ . Nous constatons de nouveau que la taille du graphe représentant le complexe est bien inférieure au nombre de simplexes justifiant l'introduction de notre structure de données.

### 3.4 Opérations élémentaires sur la structure de données squelette/bloqueurs

Nous présentons plusieurs algorithmes permettant de réaliser des opérations courantes sur des complexes simpliciaux encodés dans notre structure de données. Le premier type d'algorithme que nous présentons utilise à son avantage la compacité de notre structure de données et possède une complexité polynomiale avec la taille du graphe et le nombre de bloqueurs. Le

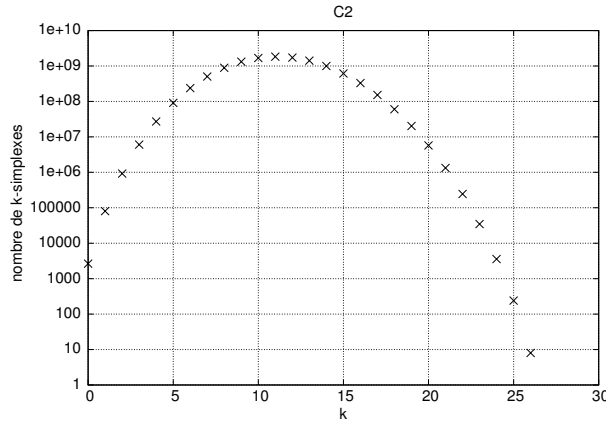


FIGURE 3.4 – Nombre de  $k$ -simplexes dans  $\text{Rips}(P, \alpha)$  en fonction de  $k$  pour un ensemble de points échantillonnant  $C_2$ . Une échelle logarithmique est utilisée pour l'axe  $y$ . Expérimentalement, nous observons que le nombre de  $k$ -simplexes est inférieur mais proche de  $\frac{n}{k+1} \binom{N}{k}$  pour  $N = 25$ . Le nombre de simplexes est donc au moins  $\frac{n}{N+1} 2^N \approx 3.4 \cdot 10^9$ . Ce nombre est à comparer avec notre structure de données qui a juste besoin de stocker un graphe constitué de 2646 sommets et 80304 arêtes pour représenter le même complexe.

deuxième type d'algorithmes sert d'interface avec d'autres structures de données pour complexes simpliciaux. Ils permettent d'importer dans notre structure de données un complexe simplicial donné par son ensemble de simplexes ou encore d'énumérer tous les  $k$ -simplexes. Cette opération est importante puisque de nombreux algorithmes, comme ceux permettant de calculer les groupes d'homologie, nécessitent de connaître l'ensemble des simplexes.

### 3.4.1 Algorithmes de complexité polynomiale en la taille du graphe et le nombre de bloqueurs

Dans cette partie, nous décrivons des algorithmes réalisant les opérations classiques suivantes :

- tester si un simplexe appartient au complexe ;
- calculer le *link* d'un simplexe dans un complexe ;
- tester si un complexe simplicial est un cône ;
- collapser un simplexe ;
- contracter une arête.

Pour chaque algorithme, un pseudo-code et un calcul de complexité sont donnés dans l'annexe de ce chapitre. On note que les opérations de contraction d'arête ou de collapse sont susceptibles de créer des bloqueurs. Or, la taille de notre structure ainsi que la complexité des opérations élémentaires augmentent avec le nombre de bloqueurs. Ainsi on souhaite que ces derniers soient peu nombreux pour que nos opérations restent efficaces. Pour cela, nous proposons dans cette section un test permettant de décider si la suppression d'un bloqueur change le type d'homotopie. Nous verrons dans les chapitres 4 et 5 que l'on peut procéder à des contractions d'arêtes et des collapses sans trop augmenter le nombre de bloqueurs.

### 3.4. Opérations élémentaires sur la structure de données squelette/bloqueurs 43

#### 3.4.1.1 Tester si un simplexe appartient au complexe

Rappelons qu'un simplexe  $\sigma$  appartient à  $K$  si et seulement si  $\sigma$  appartient à  $\text{Expand}(K^{(i)})$  et  $\sigma$  n'a pas de faces dans  $\text{Blockers}_i(K)$  (voir Eq. (3.1) de la section 3.2). En utilisant le cas où  $i = 1$ , on remarque que, pour vérifier si  $\sigma \in K$ , il suffit de tester si  $\sigma \subset \text{Vert}(K)$  en vérifiant que toutes les arêtes de  $\sigma$  appartiennent au 1-squelette et que  $\sigma$  ne contient pas de bloqueurs d'ordre 1; voir l'algorithme 7 dans l'annexe pour le pseudo-code et le calcul de complexité.

#### 3.4.1.2 Calculer le *link*

Puisque le *link* d'un simplexe  $\alpha \in K$  est un complexe simplicial, on peut aussi le représenter par une paire constituée de son 1-squelette et de son ensemble de bloqueurs. La formule ci-dessous exprime chaque élément de cette paire pour cela, on note  $\mathcal{N}(\alpha) = \bigcap_{u \in \alpha} \mathcal{N}(u)$ .

**Lemme 3.1.** *Pour tout simplexe  $\alpha$  du complexe simplicial  $K$ ,*

$$\begin{aligned} \text{Lk}(\alpha)^{(1)} &= \{\sigma \subset \mathcal{N}(\alpha) \mid \dim \sigma \leq 1, \alpha\sigma \in K\}, \\ \text{Blockers}_1(\text{Lk}(\alpha)) &= \{\sigma \subset \mathcal{N}(\alpha) \mid \dim \sigma \geq 2, \alpha\sigma \notin K, \forall \sigma' \subsetneq \sigma, \alpha\sigma' \in K\}. \end{aligned}$$

*Démonstration.* Par définition,  $\sigma$  est un sommet ou une arête du *link* de  $\alpha$  si et seulement si  $\dim \sigma \leq 1$ ,  $\alpha \cap \sigma = \emptyset$  et  $\alpha\sigma \in K$ , ce qui donne la première formule.

Par définition,  $\sigma$  est un bloqueur du *link* de  $\alpha$  si et seulement si  $\dim \sigma \geq 2$ ,  $\sigma \notin \text{Lk}(\alpha)$  et pour toutes ses faces propres  $\emptyset \neq \sigma' \subsetneq \sigma$ , on a que  $\sigma' \in \text{Lk}(\alpha)$ , ce qui induit la deuxième formule.  $\square$

Nous donnons maintenant une caractérisation des bloqueurs du *link* d'un simplexe à partir de laquelle nous dérivons un algorithme pour calculer le *link* d'un simplexe; voir l'algorithme 6 dans l'Annexe pour le pseudo-code ainsi qu'une discussion sur la complexité de celui-ci.

**Lemme 3.2.** *Soit  $\alpha \in K$  et  $\sigma \subset \text{Vert}(K)$ . Le simplexe  $\sigma$  est un bloqueur du *link* de  $\alpha$  si et seulement si  $\sigma \subset \text{Vert}(\text{Lk}(\alpha))$ ,  $\dim \sigma \geq 2$ , il existe  $\beta \in \text{Blockers}_1(K)$  tel que  $\sigma = \beta \setminus \alpha$  et il n'existe pas de bloqueur de  $K$  de la forme  $\alpha'\sigma'$  avec  $\emptyset \neq \alpha' \subset \alpha$  et  $\sigma' \subsetneq \sigma$ .*

*Démonstration.* Supposons que  $\sigma$  soit un bloqueur du *link* de  $\alpha$ . Grâce au lemme 3.1, nous savons que  $\dim \sigma \geq 2$ ,  $\sigma \subset \mathcal{N}(\alpha)$ ,  $\alpha\sigma \notin K$  et  $\forall \sigma' \subsetneq \sigma, \alpha\sigma' \in K$ . En particulier, il est facile de vérifier que tous les sommets et toutes les arêtes de  $\sigma$  appartiennent à  $K$  et que  $\alpha \cap \sigma = \emptyset$ . Ainsi, la condition  $\alpha\sigma \notin K$  implique qu'il existe un bloqueur  $\beta$  de  $K$  tel que  $\beta \subset \alpha\sigma$ . La condition  $\forall \sigma' \subsetneq \sigma, \alpha\sigma' \in K$  implique que le bloqueur  $\beta$  contient  $\sigma$  et donc  $\sigma = \beta \setminus \alpha$ . De plus, il n'existe pas de bloqueur de  $K$  de la forme  $\alpha'\sigma'$  avec  $\alpha' \subset \alpha$  et  $\sigma' \subsetneq \sigma$ .

Réciproquement, considérons  $\sigma = \beta \setminus \alpha \subset \text{Vert}(\text{Lk}(\alpha))$  tel que  $\dim \sigma \geq 2$ ,  $\beta$  soit un bloqueur de  $K$  et tel qu'il n'existe pas de bloqueur  $\alpha'\sigma'$  de  $K$  avec  $\emptyset \neq \alpha' \subset \alpha$  et  $\sigma' \subsetneq \sigma$ . Puisque  $\beta \subset \alpha\sigma$ , cela implique que  $\alpha\sigma$  est une coface du bloqueur  $\beta$  et donc n'appartient pas

à  $K$ . D'autre part, pour tout  $\emptyset \neq \sigma' \subsetneq \sigma$ , le simplexe  $\sigma'$  est une face propre du bloqueur  $\beta$  et appartient donc à  $K$ . On peut vérifier que pour tout  $\sigma' \subsetneq \sigma$ , le simplexe  $\alpha\sigma'$  appartient à  $K$  puisque toutes ses arêtes appartiennent à  $K$  et qu'aucune de ses faces n'est un bloqueur de  $K$ . On obtient grâce au Lemme 3.1 que  $\sigma \in \text{Blockers}_1(\text{Lk}(\alpha))$ .  $\square$

Un corollaire immédiat est que si  $K$  est un *flag-complexe*, le *link* de n'importe lequel de ses simplexes est également un *flag-complexe*. Le pseudo-code pour calculer le *link* d'un simplexe ainsi que le calcul de complexité sont donnés dans l'algorithme 6 de l'annexe.

### 3.4.1.3 Tester si un complexe simplicial est un cône

Le lemme suivant induit un algorithme simple pour tester si un complexe simplicial est un cône. Rappelons qu'un complexe simplicial  $K$  est un *cône* s'il existe un sommet  $o$  tel que l'implication suivante est vérifiée :  $\sigma \in K \implies \sigma \cup \{o\} \in K$ . Pour vérifier si un complexe est un cône, il suffit de chercher un sommet  $o$  qui soit relié à tous les autres sommets du complexe et qui ne soit pas inclus dans un bloqueur ; voir l'algorithme 9 dans l'annexe pour le pseudo-code et sa complexité.

**Lemme 3.3.** *Un complexe simplicial  $L$  est un cône d'apex  $o$  si et seulement si  $\mathcal{N}_L(o) = \text{Vert}(L) \setminus \{o\}$  et  $\mathcal{B}_L(o) = \emptyset$ .*

*Démonstration.* Supposons que  $L$  soit un cône d'apex  $o$ . Clairement, chaque sommet  $v \neq o$  de  $L$  est connecté à  $o$  par une arête ou de manière équivalente  $\mathcal{N}_L(o) = \text{Vert}(L) \setminus \{o\}$ . Par contradiction, supposons que  $o\tau$  soit un bloqueur de  $L$  pour un simplexe  $\tau$  tel que  $o \cap \tau = \emptyset$ . Par définition d'un bloqueur,  $o\tau \notin L$  et toutes les faces propres de  $o\tau$  appartiennent à  $L$ . En particulier  $\tau \in L$  et par définition d'un cône  $o\tau \in L$ , ce qui est impossible. Réciproquement, supposons que  $\mathcal{B}_L(o) = \emptyset$  et  $\mathcal{N}_L(o) = \text{Vert}(L) \setminus \{o\}$ . Soit  $\tau \in L$ . Il est facile de voir que  $o\tau \in L$  puisque toutes les arêtes de  $o\tau$  appartiennent à  $L$  et que  $o\tau$  ne contient pas de bloqueurs.  $\square$

### 3.4.1.4 Collapser un simplexe

Intéressons-nous à l'opération qui, étant donné un complexe simplicial  $K$  et un simplexe  $\sigma \in K$ , retire de  $K$  le *star* de  $\sigma$  dans  $K$ . On note cette opération  $K \mapsto K \setminus \text{St}_K(\sigma)$ . Cette opération de simplification est très populaire en topologie puisqu'elle préserve le type d'homotopie lorsque le *link* de  $\sigma$  dans  $K$  est la fermeture d'un simplexe [EH10]. Dans ce cas, cette opération est appelée un *collapse élémentaire*. D'autres conditions plus générales mais plus difficiles à vérifier sur le *link* existent pour assurer que cette opération préserve le type d'homotopie. Elles seront étudiées dans le chapitre 5. Nous nous penchons ici uniquement sur le moyen de retirer le *star* d'un simplexe d'un complexe encodé dans notre structure de données.<sup>4</sup> Le lemme suivant exprime le complexe  $K \setminus \text{St}_K(\sigma)$  en terme de squelette/bloqueurs

4. Notons simplement que tester si l'opération  $K \mapsto K \setminus \text{St}_K(\sigma)$  est un *collapse élémentaire* est déjà possible dans notre structure. En effet, il suffit de calculer le *link* de  $\sigma$  dans notre structure et de vérifier si le 1-squelette du complexe obtenu est un graphe complet.

### 3.4. Opérations élémentaires sur la structure de données squelette/bloqueurs 45

et permet de réaliser l'opération  $K \mapsto K \setminus \text{St}_K(\sigma)$  dans notre structure.

**Lemme 3.4.** Notons  $\mathcal{B}(\sigma) = \{\beta \in \text{Blockers}(K) \mid \beta \cap \sigma \neq \emptyset\}$  l'ensemble des bloqueurs de  $K$  contenant  $\sigma$  et  $K' = K \setminus \text{St}_K(\sigma)$ .

$$\text{si } \dim \sigma = 0 \text{ alors } \begin{cases} K'^{(1)} = K^{(1)} \setminus (\sigma \cup \{\sigma x, \sigma x \in \text{Edges}(K)\}) \\ \text{Blockers}(K') = \text{Blockers}(K) \setminus \mathcal{B}(\sigma) \end{cases}$$

$$\text{si } \dim \sigma = 1 \text{ alors } \begin{cases} K'^{(1)} = K^{(1)} \setminus \sigma \\ \text{Blockers}(K') = \text{Blockers}(K) \setminus \mathcal{B}(\sigma) \end{cases}$$

$$\text{si } \dim \sigma > 1 \text{ alors } \begin{cases} K'^{(1)} = K^{(1)} \\ \text{Blockers}(K') = (\text{Blockers}(K) \setminus \mathcal{B}(\sigma)) \cup \{\sigma\} \end{cases}$$

*Démonstration.* Nous montrons la correction du lemme seulement dans le cas où  $\sigma$  est un sommet, les autres cas étant similaires. Lorsque  $\dim \sigma = 0$ , le 1-squelette de  $\text{St}_K(\sigma)$  est constitué de  $\sigma$  ainsi que toutes les arêtes  $\sigma x$  lorsque  $\sigma x \in \text{Edges}(K)$  ce qui justifie la formule du 1-squelette de  $K'$ . Concernant l'ensemble des bloqueurs de  $K'$ , notons la suite d'équivalences suivante :

$$\begin{aligned} \beta \in \text{Blockers}(K') &\iff \beta \not\subseteq K' \text{ et } \forall \tau \subsetneq \beta, \tau \in K' \\ &\iff \beta \not\subseteq K' \text{ et } \forall \tau \subsetneq \beta, (\tau \in K \text{ et } \tau \cap \sigma = \emptyset) \\ &\iff \beta \not\subseteq K' \text{ et } (\forall \tau \subsetneq \beta, \tau \in K) \text{ et } \beta \cap \sigma = \emptyset \\ &\iff \beta \not\subseteq K \text{ et } (\forall \tau \subsetneq \beta, \tau \in K) \text{ et } \beta \cap \sigma = \emptyset \\ &\iff \beta \in \text{Blockers}(K) \setminus \mathcal{B}(\sigma) \end{aligned}$$

Ainsi  $\text{Blockers}(K') = \text{Blockers}(K) \setminus \mathcal{B}(\sigma)$ . □

Une remarque importante est que si  $K$  est un *flag-complexe*, collapser un sommet ou une arête n'altère pas cette propriété. Autrement dit, collapser un sommet ou une arête dans un *flag-complexe* ne crée pas de bloqueurs. Pour cette raison, ces deux opérations seront étudiées dans le chapitre 5 où on cherchera à les appliquer pour simplifier un complexe de Rips.

Au lieu de représenter les complexes simpliciaux dans une paire  $(K^{(1)}, \text{Blockers}_1(K))$  nous aurions aussi pu choisir de les encoder avec une paire  $(K^{(2)}, \text{Blockers}_2(K))$ . Dans ce cas, nous aurions stocké  $K^{(2)}$ , le 2-squelette de  $K$ , dans une structure permettant de représenter une triangulation comme une structure dite *DCEL* [MP78] par exemple. Nous aurions pu alors collapser des sommets, des arêtes mais également des triangles sans rajouter de bloqueurs d'ordre 2 dans cette structure. Une telle structure offre de nouvelles possibilités de simplification par collapses par rapport à ce qui est proposé dans le chapitre 5.



Plus généralement, on peut représenter les complexes simpliciaux dans une paire  $(K^{(i)}, \text{Blockers}_i(K))$  si l'on a une structure à même de traiter les complexes simpliciaux de dimension  $i$ . Dans cette structure, collapser un  $i$ -simplexe ne crée pas de bloqueurs d'ordre  $i$ .

### 3.4.1.5 Contracter une arête

#### Définition de la contraction d'arête

On rappelle que  $\text{Vert}(K)$  dénote l'ensemble des sommets de  $K$ . Considérons  $a, b \in \text{Vert}(K)$  et  $c \notin \text{Vert}(K)$ . Pour décrire la contraction d'arête  $ab \mapsto c$ , on définit l'application simpliciale  $f$  qui associe à  $a$  et  $b$  le sommet  $c$  et qui est égale à l'identité pour tous les autres sommets :

$$f(v) = \begin{cases} c & \text{si } v \in \{a, b\}, \\ v & \text{si } v \notin \{a, b\}. \end{cases} \quad (3.2)$$

On étend ensuite  $f$  à tous les simplexes  $\sigma = \{v_0, \dots, v_k\}$  de  $K$ , avec  $f(\sigma) = \{f(v_0), \dots, f(v_k)\}$ . La contraction d'arête  $ab \mapsto c$  est l'opération qui associe à  $K$  le complexe  $K' = \{f(\sigma) \mid \sigma \in K\}$ . Par construction,  $f$  est surjective et  $K'$  est un complexe simplicial. Notons que la contraction d'arête  $ab \mapsto c$  est définie même quand  $ab$  n'appartient pas à  $K$ .

#### Mise à jour de la structure de données après une contraction d'arête

Nous décrivons maintenant comment mettre à jour notre structure de données après une contraction d'arête. Plus précisément, étant donné  $K$  un complexe simplicial et  $K'$  le complexe simplicial obtenu après la contraction d'arête  $ab \mapsto c$ , notre but est de calculer la paire  $(\mathcal{N}_{K'}(c), \mathcal{B}_{K'}(c))$ . Clairement,  $\mathcal{N}_{K'}(c) = (\mathcal{N}(a) \setminus \{b\}) \cup (\mathcal{N}(b) \setminus \{a\})$ . Il suffit donc d'exprimer l'ensemble des bloqueurs  $\mathcal{B}_{K'}(c)$ . Nous commençons par un lemme technique permettant de caractériser les simplexes du *link* de  $c$  après la contraction d'arête  $ab \mapsto c$ .

**Lemme 3.5.** *Soit  $\sigma \subset \text{Vert}(K) \setminus \{a, b\}$ . Le simplexe  $c\sigma$  appartient à  $K'$  si et seulement si  $a\sigma$  appartient à  $K$  ou  $b\sigma$  appartient à  $K$ . De manière équivalente,  $\sigma \in \text{Lk}(c)$  si et seulement si  $\sigma \in \text{Lk}(a) \cup \text{Lk}(b)$ .*

*Démonstration.* En utilisant le fait que  $f^{-1}[\{c\sigma\}] = \{a\sigma, b\sigma, ab\sigma\} \cap K$  et la surjectivité de  $f$ , on obtient que  $c\sigma \in K' \iff f^{-1}[\{c\sigma\}] \neq \emptyset \iff \{a\sigma, b\sigma, ab\sigma\} \cap K \neq \emptyset \iff \{a\sigma, b\sigma\} \cap K \neq \emptyset \iff a\sigma \in K \text{ ou } b\sigma \in K$ .  $\square$

Le lemme suivant permet ensuite de calculer  $\mathcal{B}_{K'}(c)$  puisqu'il caractérise les bloqueurs passant par  $c$ . Pour cette preuve, il sera utile de noter l'équivalence suivante qui découle de la définition de  $\text{Blockers}_0(K)$ .

$$\forall \sigma \subset \text{Vert}(K), \sigma \notin K \iff \sigma \text{ a une face dans } \text{Blockers}_0(K)$$

**Lemme 3.6.** *Soit  $K'$  le complexe simplicial obtenu après la contraction d'arête  $ab \mapsto c$ . Soit  $\sigma \subset \text{Vert}(K) \setminus \{a, b\}$  un simplexe tel que  $\dim \sigma \geq 1$ . Le simplexe  $c\sigma$  appartient à  $\text{Blockers}_1(K')$  si et seulement si les deux conditions suivantes sont remplies :*

### 3.4. Opérations élémentaires sur la structure de données squelette/bloqueurs 47

- (i)  $\sigma \in K$  ; toute face propre de  $\sigma$  appartient à  $\text{Lk}(a) \cup \text{Lk}(b)$  ;
- (ii)  $\sigma = \alpha\beta$  où  $a\beta \in \text{Blockers}_0(K)$  et  $b\alpha \in \text{Blockers}_0(K)$ .

*Démonstration.* En premier lieu, notons que les faces propres de  $c\sigma$  appartiennent à  $K'$  si et seulement si (i) est vérifiée. En effet, en utilisant le lemme 3.5, (i) est équivalent à  $\sigma \in K'$  et  $c\tau \in K'$  pour toutes les faces propres  $\tau \subsetneq \sigma$ . Ainsi,  $c\sigma \in \text{Blockers}_1(K') \implies$  (i). Nous allons maintenant montrer que  $c\sigma \in \text{Blockers}_1(K') \implies$  (ii). Puisque  $c\sigma \notin K'$ , ni  $a\sigma$  ni  $b\sigma$  n'appartiennent à  $K$ . On en déduit que  $a\sigma$  a une face dans  $\text{Blockers}_0(K)$  et puisque  $\sigma \in K$ , cette face contient  $a$ . Notons cette face  $a\beta$  où  $\beta \subset \sigma$ . Symétriquement, puisque  $b\sigma \notin K$ , il existe une face  $\alpha \subset \sigma$  telle que  $b\alpha \in \text{Blockers}_0(K)$ . Montrons maintenant que  $\alpha\beta = \sigma$ . Par contradiction, supposons que  $\alpha\beta$  est une face propre de  $\sigma$ . (i) implique que  $\alpha\beta \in \text{Lk}(a) \cup \text{Lk}(b)$ . Si  $\alpha\beta$  appartient à  $\text{Lk}(a)$ , alors  $\beta$  étant une face de  $\alpha\beta$ , il appartient aussi à  $\text{Lk}(a)$  ce qui contredit  $a\beta \in \text{Blockers}_0(K)$ . Symétriquement, si  $\alpha\beta$  appartient à  $\text{Lk}(b)$ , on obtient aussi une contradiction.

Réciproquement, montrons que (i) et (ii)  $\implies c\sigma \in \text{Blockers}(K')$ . Nous avons vu que (i) implique que toutes les faces propres de  $c\sigma$  appartiennent à  $K'$ . Pour prouver que  $c\sigma \notin K'$ , observons que ni  $a\sigma$  ni  $b\sigma$  n'appartiennent à  $K$ . En effet,  $a\sigma = a\alpha\beta \notin K$  puisque sa face  $a\beta \in \text{Blockers}_0(K)$  et  $b\sigma = b\alpha\beta \notin K$  puisque sa face  $b\alpha \in \text{Blockers}_0(K)$ .  $\square$

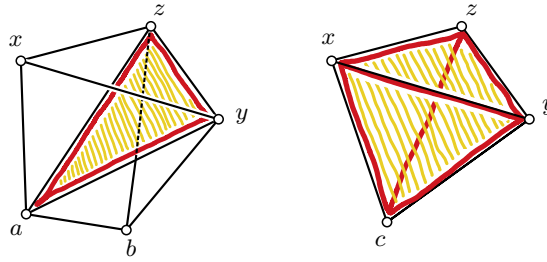


FIGURE 3.5 – Le triangle  $ayz$  est un 2-bloqueur. Les *links* de  $a$ , de  $b$  et de  $ab$  valent respectivement  $\text{Lk}(a) = \{x, y, z, xy, xz, b, by, bz\}$ ,  $\text{Lk}(b) = \{y, z, yz, a, ay, az\}$  et  $\text{Lk}(ab) = \{y, z\}$ . Observons que  $\sigma = \alpha\beta$  avec  $\alpha = x$  et  $\beta = yz$  vérifie les conditions (i) et (ii) du lemme 3.6. Ainsi, la contraction d'arête  $ab \mapsto c$  engendre la création du 3-bloqueur  $cxyz$  et la destruction du 2-bloqueur  $ayz$ .

Les lemmes 3.5 et 3.6 nous permettent de dériver un algorithme pour calculer  $\mathcal{B}_{K'}(c)$ . Le pseudo-code de cet algorithme et sa complexité sont donnés dans l'annexe D.3.

#### 3.4.1.6 Garder uniquement les bloqueurs utiles pour le type d'homotopie

Plus le nombre de bloqueurs présent dans notre structure est faible, plus notre structure a une taille réduite. Or, nous venons de voir que les opérations de contraction d'arête ou de collapse pouvaient créer des bloqueurs. Nous cherchons ici une condition nous permettant de "retirer" un bloqueur d'un complexe simplicial sans en changer le type d'homotopie. Plus

précisément, étant donné un complexe simplicial  $K$  représenté par une paire  $(S, B)$ , étudions l'opération qui retire un bloqueur  $\sigma$  de l'ensemble de bloqueur  $B$ , tout en laissant intact le 1-squelette  $S$ . Le résultat est un complexe  $K'$  représenté par une paire  $(S, B \setminus \{\sigma\})$ . Par définition d'un bloqueur, toutes les faces de  $\sigma$  sauf  $\sigma$  appartiennent à  $K$ . Ainsi, enlever  $\sigma$  de l'ensemble des bloqueurs de  $K$  a pour effet d'ajouter le simplexe correspondant au complexe (plus éventuellement des cofaces). On peut retrouver  $K$  à partir de  $K'$  en enlevant précisément le *star* de  $\sigma$  dans  $K'$ , i.e.  $K = K' \setminus \text{St}_{K'}(\sigma)$ . La proposition 1.2 du chapitre 2 établit que, lorsque  $\text{Lk}_{K'}(\sigma)$  est contractile,  $K$  et  $K'$  ont le même type d'homotopie. Tester si un complexe est contractile est difficile (ce problème sera discuté dans le chapitre 5) mais il est possible de vérifier une condition plus simple comme le fait que le *link* soit un cône. Un cône est toujours contractile et nous avons vu une manière efficace de tester si un complexe est un cône.

Nous introduisons la définition suivante afin de caractériser un bloqueur que l'on peut enlever sans changer le type d'homotopie.

**Définition 3.2.** *Soit  $K$  un complexe simplicial et  $K'$  le complexe simplicial encodé par la paire  $(K^{(1)}, \text{Blockers}_1(K) \setminus \{\sigma\})$ . Le bloqueur  $\sigma$  de  $K$  est simplement redondant si le link de  $\sigma$  dans  $K'$  est un cône. Les complexes  $K$  et  $K'$  ont le même type d'homotopie. Lorsque le link de  $\sigma$  dans le complexe simplicial  $K'$  est contractile, on dira que le bloqueur  $\sigma$  est redondant, dans ce cas,  $K$  et  $K'$  partagent encore le même type d'homotopie.*

Une manière simple de limiter le nombre de bloqueurs est donc de parcourir tous les bloqueurs et de supprimer tous ceux qui sont simplement redondants. Cette technique sera utilisée dans le chapitre 4.

### 3.4.2 Algorithmes de conversion entre structures de données

#### 3.4.2.1 Calculer l'expansion du graphe

Plusieurs algorithmes ont déjà été proposés pour calculer le  $k$ -squelette d'un *flag*-complexe [Zom10b, BM12b]. Zomorodian introduit et compare notamment plusieurs approches dans [Zom10b]. L'algorithme qui est le plus rapide d'après ses expérimentations est celui qu'il appelle algorithme incrémental. Pour calculer l'expansion d'un 1-squelette, cet algorithme ajoute les sommets incrémentalement. Lorsqu'un sommet est ajouté, il ajoute ensuite toutes les cofaces contenant ce sommet récursivement. Il n'est pas difficile d'adapter cet algorithme dans notre structure de données. L'algorithme 11 donné dans l'annexe D.6 permet d'énumérer les simplexes du  $k$ -squelette d'un complexe représenté par une paire squelette/bloqueurs. Une preuve de sa correction peut être obtenue en modifiant légèrement la preuve originale de Zomorodian ; elle est également donnée en annexe.

### 3.4.2.2 Représenter dans notre structure un complexe donné par une liste de simplexes

Nous cherchons maintenant à importer dans notre structure de données un complexe simplicial  $K$  donné par une liste de simplexes. Nous supposons que cette liste nous est donnée sous la forme  $L = (L_0, \dots, L_k)$  où  $L_i$  est la liste des  $i$ -simplexes de  $K$ . La construction du squelette  $K^{(1)}$  est facile puisqu'il est constitué uniquement des simplexes de  $K$  de dimension inférieure ou égale à 1 et donc  $K^{(1)} = L_0 \cup L_1$ . Il nous reste à voir comment reconstituer l'ensemble de bloqueurs. Rappelons que les bloqueurs de  $K$  sont les éléments minimaux pour l'inclusion de  $\text{Expand}(K^{(1)}) \setminus K$ . Un moyen direct pour énumérer les bloqueurs de  $K$  est de parcourir les  $i$ -simplexes  $\sigma$  de  $\text{Expand}(K^{(1)})$  par dimension croissante. Pour chaque simplexe  $\sigma$ , on ajoute  $\sigma$  à l'ensemble des bloqueurs de  $K$  si  $\sigma \notin K$  et si aucune des faces de  $\sigma$  n'est un bloqueur. Cette solution impose de calculer tous les simplexes de l'expansion de  $K$ . Il est également possible d'obtenir un algorithme plus efficace en évitant de parcourir les co-faces de bloqueurs dans l'expansion du complexe. Afin de l'introduire, notons  $E_i$  l'ensemble des  $i$ -simplexes de l'expansion de  $K$ . Observons que  $E_i = \{\sigma \mid \dim \sigma = i \text{ et } \forall \tau \subsetneq \sigma, \tau \in K\}$ . Ainsi, l'ensemble des bloqueurs de dimension  $i$  est alors égal à  $E_i \setminus L_i$ . Notons que l'on peut calculer  $E_{i+1}$  à partir de  $L_i$ . En effet, pour  $\sigma \in L_i$  et  $v \in \text{Vert}(K) \setminus \sigma$ , remarquons que  $\sigma v \in E_{i+1}$  si et seulement si, pour tout sommet  $w$  de  $\sigma$ , le simplexe  $\sigma v \setminus \{w\}$  appartient à  $L_i$ . L'algorithme 1 calcule  $E_{i+1}$  à partir de  $L_i$  en utilisant cette remarque. Grâce à celui-ci, on peut alors introduire l'algorithme 2 qui permet d'énumérer les bloqueurs du complexe, étant donnée une liste de simplexes  $L = (L_0, \dots, L_k)$ .

---

#### Algorithme 1 NEXT-EXPAND( $L_i$ )

---

```

1:  $E_{i+1} \leftarrow \emptyset$ 
2: for  $\sigma \in E_i$  do
3:   for all  $v \in L_0 \setminus \sigma$  do
4:     test  $\leftarrow$  true
5:     for all  $w \in \sigma$  do
6:       if  $\sigma v \setminus \{w\} \notin L_i$  then
7:         test  $\leftarrow$  false
8:         break
9:     if test then
10:       $E_{i+1} \leftarrow E_{i+1} \cup \sigma v$ 
11: return  $E_{i+1}$ 

```

---

## 3.5 Conclusion et perspectives

La structure de données squelette/bloqueurs que nous avons introduite dans ce chapitre a pour avantage sa compacité ainsi que la faible complexité nécessaire pour réaliser les opérations simpliciales classiques. Des expériences seront menées dans les prochains chapitres sur cette structure. Il s'agira d'appliquer des contractions d'arêtes et des collapsés sur un complexe

---

**Algorithme 2** ENUMERE-BLOQUEURS( $L = (L_0, \dots, L_k)$ )
 

---

```

1:  $B \leftarrow \emptyset$ 
2: for  $i = 2$  to  $\dim K$  do
3:    $E_i \leftarrow \text{NEXT-EXPAND}(L_{i-1})$ 
4:   {Enumère les  $i$ -simplexes de  $E_i$  de trouver les bloqueurs d'ordre  $i$ }
5:   for all  $\sigma \in E_i$  do
6:     if  $\sigma \notin K$  then
7:        $B = B \cup \{\sigma\}$ 
8: return  $B$ 

```

---

encodé dans cette structure de données. A chaque fois le complexe de départ sera un complexe de Rips qui est un *flag*-complexe. Cependant dans le chapitre 4, le complexe s'éloigne d'un *flag*-complexe puisqu'on applique des contractions d'arêtes susceptibles de créer des bloqueurs. Néanmoins, on s'apercevra qu'il est possible de limiter le nombre de bloqueurs. Notre structure de données est donc compacte et efficace dans ce cas.

Une perspective de travail serait l'analyse et l'expérimentation du comportement de cette structure de données pour des complexes qui ne sont pas initialement des *flag*-complexes. Une possibilité serait d'étudier le comportement de cette structure pour stocker des complexes de Čech, par exemple. Se pose alors la question de savoir combien de bloqueurs sont nécessaires pour représenter de tels complexes. Egalement, une question d'intérêt est de connaître la dimension de tels bloqueurs puisque ceux-ci alourdissent notre structure. Dans le cas où l'on construit un complexe  $\text{Cech}(P, \alpha)$  et que  $P \subset \mathbb{R}^D$ , nous avons pu voir que la dimension des bloqueurs du complexe de Čech était toujours inférieure à  $D$  en raison du théorème de Helly. Lorsque  $P$  échantillonne une  $d$ -variété  $A$  peu courbée, et si on choisit un paramètre  $\alpha$  tel que  $\text{Cech}(P, \alpha) \simeq A$ , on peut s'attendre intuitivement à ce que la dimension des bloqueurs ne soit pas beaucoup plus grande que  $d$ . En effet, sous de bonnes conditions d'échantillonnage, les sommets d'un bloqueur du Čech sont proches d'un hyperplan de dimension  $d$ . Serait-il alors possible de montrer que la dimension de ce bloqueur est proche de  $d$  en utilisant encore le théorème de Helly ?

# Simplification d'un complexe simplicial par contraction d'arête

## Sommaire

<b>4.1</b>	<b>Introduction</b>	<b>51</b>
<b>4.2</b>	<b>Une condition sur le <i>link</i> pour qu'une contraction d'arête préserve le type d'homotopie</b>	<b>52</b>
4.2.1	Contraction d'arête	52
4.2.2	Condition du <i>link</i>	53
4.2.3	Tester la condition du <i>link</i> efficacement	56
4.2.4	Réciproque de la condition du <i>link</i>	57
<b>4.3</b>	<b>Simplification par contraction d'arêtes</b>	<b>58</b>
4.3.1	Un algorithme	58
4.3.2	Priorité sur les arêtes	59
<b>4.4</b>	<b>Résultats expérimentaux</b>	<b>62</b>
4.4.1	Jeu de données	62
4.4.2	Taille de la structure de données au cours des contractions	63
4.4.3	Taille des complexes après contraction sur le jeu de données	64

## 4.1 Introduction

Les données réelles sont souvent représentées par des ensembles de points de taille gigantesque et de grande dimension. Les complexes simpliciaux bâtis à partir de telles données ont alors des tailles rendant leur stockage et leur traitement délicat. Ainsi voit-on apparaître la nécessité de pouvoir simplifier de tels complexes en préservant leur topologie si possible, puisque celle-ci nous permet de savoir si notre approximation ne s'éloigne pas trop du complexe initial, du moins qualitativement. Dans ce but, Kaczyński et al. ont proposé de simplifier par collapses un complexe simplicial donné par son ensemble de simplexes [KMS98]. Cependant, le nombre de simplexes dans les complexes de Rips devient rapidement colossal comme l'a montré l'étude du chapitre 3. Afin d'éviter de traiter l'ensemble des simplexes d'un complexe simplicial, Zomorodian s'est intéressé aux *flag-complexes*, un ensemble de complexes dont les simplexes sont définis implicitement par un graphe, et a proposé de construire le *simplicial set* de tels complexes [Zom10b]. Celui-ci peut être construit efficacement à partir du seul graphe, il

possède une taille modeste (souvent largement inférieur au nombre de simplexes) et on peut en extraire les groupes d'homologie. Une autre possibilité pour réduire la taille d'un complexe est d'utiliser des contractions d'arêtes. Cette opération, issue de l'informatique graphique, permet de simplifier une triangulation en réduisant son nombre de sommets tout en maintenant une erreur géométrique minimale [HDD<sup>+</sup>93, GH97]. Une condition locale garantissant qu'une contraction d'arête préserve la topologie a été donnée par Dey et al. [DEGN98]. Les auteurs ont montré que cette condition, appelée *condition du link*, permet de caractériser les contractions d'arêtes qui induisent des transformations homéomorphes lorsque le complexe simplicial est homéomorphe à une variété de dimension 2 ou 3 [DEGN98]. Ce résultat a par la suite été étendu dans le cas où le complexe simplicial est homéomorphe à une variété de dimension arbitraire [Nev06]. Dans ce chapitre, nous présentons une variante de ce résultat. Précisément, pour tout complexe simplicial, nous montrons que la contraction d'une arête préserve le type d'homotopie lorsque la condition du *link* est vérifiée. Une bonne nouvelle est que cette condition peut être vérifiée efficacement avec la structure de données squelette/bloqueurs introduite dans le chapitre 3. On peut alors simplifier des complexes simpliciaux encodés dans cette structure tout en préservant le type d'homotopie. Il suffit pour cela de contracter des arêtes vérifiant la condition du *link*. Cependant de telles contractions peuvent engendrer des bloqueurs et alourdir notre structure de données. Pour pallier ce problème, nous proposons une heuristique induisant un ordre sur les contractions et nous montrons comment enlever certains bloqueurs. Deux résultats expérimentaux notables sont :

- le faible nombre de bloqueurs que nous rencontrons, ce qui permet une exécution très efficace ;
- la possibilité d'arriver à des complexes ayant un nombre de sommets très faible, proche du minimum de sommets nécessaires pour décrire de tels complexes.

*Ce chapitre reprend une partie des articles [ALS11a, ALS12a] co-écrits avec Dominique Attali et André Lieutier.*

## 4.2 Une condition sur le *link* pour qu'une contraction d'arête préserve le type d'homotopie

Dans cette section, nous donnons une condition locale sur le *link* d'une arête  $ab$  d'un complexe simplicial  $K$  garantissant que la contraction de l'arête  $ab$  ne change pas le type d'homotopie de  $K$ . Nous commençons par rappeler la définition de l'opération de contraction d'arête puis nous donnons l'énoncé du théorème du nerf que nous utilisons pour prouver la condition du *link*.

### 4.2.1 Contraction d'arête

Nous rappelons ici brièvement la définition de l'opération de contraction d'arête qui a été définie dans le chapitre 3. Soient  $a, b \in \text{Vert}(K)$  et  $c \notin \text{Vert}(K)$ . Pour définir la contraction

d'arête  $ab \mapsto c$ , on définit l'application simpliciale  $f$  qui à  $a$  et  $b$  associe  $c$  et qui est constante pour tous les autres sommets :

$$f(v) = \begin{cases} c & \text{si } v \in \{a, b\}, \\ v & \text{si } v \notin \{a, b\}. \end{cases} \quad (4.1)$$

Nous étendons  $f$  sur les autres simplexes  $\sigma = \{v_0, \dots, v_k\}$  de  $K$ , en posant  $f(\sigma) = \{f(v_0), \dots, f(v_k)\}$ . La contraction d'arête  $ab \mapsto c$  est l'opération qui change  $K$  en  $K' = \{f(\sigma) \mid \sigma \in K\}$ . Par construction,  $f$  est surjective et  $K'$  est un complexe simplicial.

### 4.2.2 Condition du *link*

Soit  $\sigma$  un simplexe du complexe simplicial  $K$ . Nous rappelons que le *link* de  $\sigma$  dans  $K$  est le complexe simplicial :

$$\text{Lk}_K(\sigma) = \{\tau \in K \mid \tau \cup \sigma \in K, \tau \cap \sigma = \emptyset\}.$$

Lorsque le contexte est sans ambiguïté, le *link* de  $\sigma$  dans  $K$  est noté simplement  $\text{Lk}(\sigma)$ . En particulier, lorsque l'on contracte une arête  $ab \in K$  sur un nouveau sommet  $c \in K'$ , on peut écrire sans ambiguïté  $\text{Lk}(a)$ ,  $\text{Lk}(b)$ ,  $\text{Lk}(ab)$  pour les *links* de  $a$ ,  $b$  et  $ab$  dans  $K$  et  $\text{Lk}(c)$  pour le *link* de  $c$  dans  $K'$ . Nous pouvons maintenant formuler le résultat suivant (voir figure 4.1).

**Théorème 4.1.** *Soit  $K$  un complexe simplicial. La contraction d'arête  $ab \in K$  préserve le type d'homotopie lorsque  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$ .*

Sur la figure 4.1<sup>1</sup>, l'arête  $ab$  vérifie la condition  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  et sa contraction préserve en effet le type d'homotopie (les complexes sont représentés en suivant la convention choisie à la fin de l'introduction). On observe sur la figure 4.2 que la réciproque du théorème 4.1 n'est pas vraie en général.

Une preuve du théorème 4.1 est proposée dans la suite de cette section. Nous commençons par rappeler le théorème du nerf sur lequel repose notre preuve. Nous prouvons ensuite notre théorème avant de donner deux lemmes techniques sur le *link*.

Rappelons que le *nerf* d'une collection finie d'ensembles  $\mathcal{F}$ , est le complexe simplicial constitué de toutes les sous-collections non-vides de  $\mathcal{F}$  dont les ensembles ont une intersection non-vide,

$$\text{Nrv } \mathcal{F} = \{\mathcal{G} \subset \mathcal{F} \mid \bigcap \mathcal{G} \neq \emptyset\}$$

Le théorème du nerf a plusieurs versions [Bjo95], une des plus anciennes étant due à Leray [Ler45]. Pour la preuve du théorème 4.1, nous utiliserons la version suivante :

---

1. Lorsque l'on représente un complexe simplicial dans une figure, on adopte la convention qu'en plus de représenter le 1-squelette, soit nous grisons les simplexes maximaux pour l'inclusion, soit nous encadrons en rouge les bloqueurs. Lorsqu'aucun triangle n'est grisé ou encadré en rouge, l'ensemble des bloqueurs est vide ou, de manière équivalente, le complexe est un *flag-complexe*.



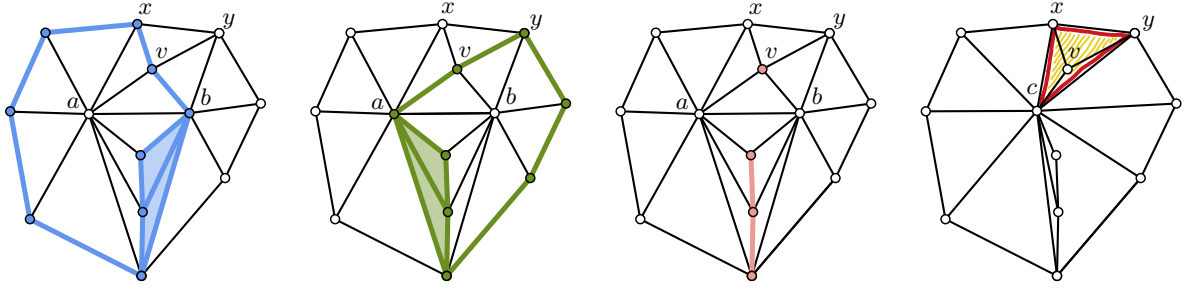


FIGURE 4.1 – De gauche à droite : le *link* de  $a$ , le *link* de  $b$ , le *link* de  $ab$  et le complexe simplicial obtenu après la contraction d'arête  $ab$ . On remarque que  $ab$  vérifie la condition du *link*. De manière équivalente, aucun bloqueur ne contient  $ab$  (voir le lemme 4.2). La contraction d'arête  $ab$  préserve donc le type d'homotopie. Notons que cette contraction d'arête crée un bloqueur  $cxy$  et que  $\alpha = x$  and  $\beta = y$  vérifient les hypothèses du lemme 3.6 du chapitre 3 caractérisant les bloqueurs d'un complexe après une contraction d'arête. La contraction de n'importe quelle arête parmi  $cv$ ,  $xv$  ou  $yv$  enlève le bloqueur  $cxy$ .

**Théorème 4.2** (Théorème du nerf). *Soit  $X$  un espace triangulable qui soit l'union d'une collection finie d'ensembles fermés  $\mathcal{F}$ , i.e.  $X = \bigcup \mathcal{F}$ . Si pour toute sous-collection  $\mathcal{G} \in \text{Nrv } \mathcal{F}$ , l'intersection  $\bigcap \mathcal{G}$  est contractile, alors l'espace sous-jacent à  $\text{Nrv } \mathcal{F}$  a la même type d'homotopie que  $X$ .*

*Preuve du théorème 4.1.* La preuve de ce théorème repose sur le fait que l'application simpliciale  $f$  préserve le type d'homotopie lorsque les préimages  $|f^{-1}[\text{Cl}(\tau')]|$  sont contractibles pour tout  $\tau' \in K'$  [Bar10]. Nous donnons une preuve de ce résultat avant de montrer que préimages  $|f^{-1}[\text{Cl}(\tau')]|$  sont contractibles.

Supposons que  $ab \in K$  vérifie la condition du *link*  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  et soit  $K'$  le complexe obtenu après la contraction d'arête  $ab \mapsto c$ . Nous allons considérer deux couvertures, une pour  $|K|$  ainsi qu'une pour  $|K'|$ , dont les nerfs  $N$  et  $N'$  seront isomorphiques et pour lesquels on établira que  $|K| \simeq |N|$  et  $|N'| \simeq |K'|$ . Par abus de notation, l'espace sous-jacent de la fermeture de  $\sigma$  sera noté  $|\sigma|$ .

Considérons dans un premier temps la collection d'ensembles  $\{|\sigma'|, \sigma' \in K'\}$  qui couvre  $|K'|$  et notons  $N'$  son nerf. Clairement, pour chaque sous-collection non vide  $\Sigma' \subset K'$ , l'intersection  $\bigcap_{\sigma' \in \Sigma'} \sigma'$  est soit vide soit un simplexe de  $K'$  ainsi l'intersection  $\bigcap_{\sigma' \in \Sigma'} |\sigma'|$  des ensembles de la sous-collection  $\{|\sigma'|, \sigma' \in \Sigma'\}$  est soit vide soit contractile. Le théorème du nerf implique que  $|K'| \simeq |N'|$ . Soit  $\bar{f} : |K| \rightarrow |K'|$  l'application simpliciale induite par l'application sur les sommets  $f : \text{Vert}(K) \rightarrow \text{Vert}(K')$  définie dans l'équation eq. (4.1). Considérons la collection d'ensembles  $\{|\bar{f}^{-1}[\sigma']|, \sigma' \in K'\}$  obtenue en prenant les préimages des ensembles de la première collection. Cette collection couvre  $|K|$  et on note son nerf  $N$ . Les deux nerfs  $N$  et  $N'$  sont isomorphes puisque la surjectivité de  $\bar{f}$  implique que pour tout  $\Sigma' \subset K'$ , nous avons l'équivalence :

$$\bigcap_{\sigma' \in \Sigma'} |\bar{f}^{-1}[\sigma']| = \bar{f}^{-1} \left[ \bigcap_{\sigma' \in \Sigma'} |\sigma'| \right] \neq \emptyset \iff \bigcap_{\sigma' \in \Sigma'} |\sigma'| \neq \emptyset.$$

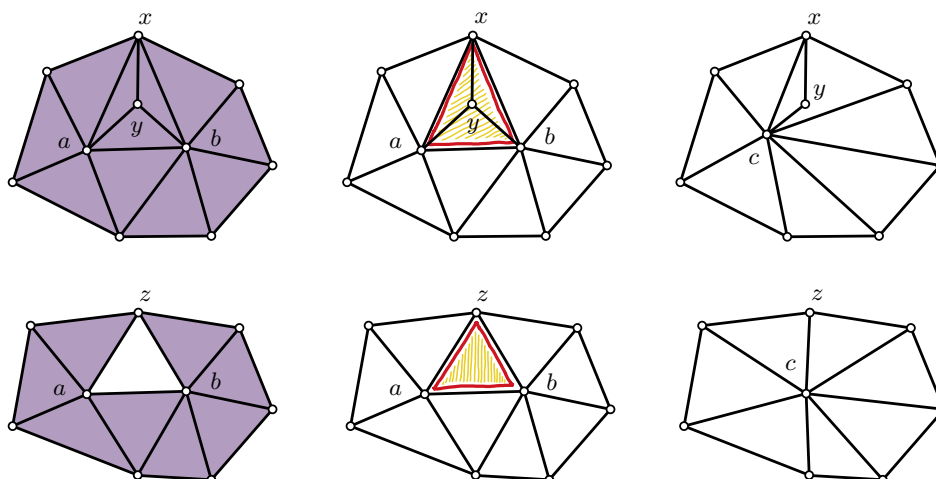


FIGURE 4.2 – En haut et en bas : deux complexes simpliciaux de dimension 2 avec un bloqueur passant par  $ab$ . A droite, la contraction d'arête  $ab \mapsto c$  peut changer le type d'homotopie (en bas) ou le préserver (en haut).

De plus, si l'intersection  $\bigcap_{\sigma' \in \Sigma'} |\sigma'|$  est non vide, alors il existe un simplexe  $\tau' \in K'$  tel que  $\tau' = \bigcap_{\sigma' \in \Sigma'} \sigma'$  et en utilisant le lemme 4.1 qui est prouvé juste après, l'intersection  $\bigcap_{\sigma' \in \Sigma'} \bar{f}^{-1}[|\sigma'|] = \bar{f}^{-1}[|\tau'|] = |f^{-1}[\text{Cl}(\tau')]|$  est contractile.

Pour résumer, nous avons établi que  $|K| \simeq |N| \approx |N'| \simeq |K'|$ , montrant que  $|K|$  et  $|K'|$  ont le même type d'homotopie.  $\square$

La preuve du théorème 4.1 nécessite d'établir que  $|f^{-1}[\text{Cl}(\tau)]|$  est contractile pour tout simplexe  $\tau$  de  $K' = f(K)$  lorsque  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$ . Un résultat plus fort existe, lorsque la condition du *link* est vérifiée,  $f^{-1}[\text{Cl}(\tau)]$  est collapsable [Mel11].

**Lemme 4.1.** *Supposons que  $ab \in K$  vérifie  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  et soit  $K'$  le complexe simplicial obtenu après la contraction d'arête  $ab \mapsto c$ . La préimage  $f^{-1}[\text{Cl}(\sigma)]$  de la fermeture de tout simplexe  $\sigma \in K'$  est non vide et collapsable.*

*Démonstration.* Pour tout  $\sigma \in K'$ , nous donnons l'expression de la préimage  $f^{-1}[\text{Cl}(\sigma)]$  dont on peut déduire facilement la collapsibilité. Rappelons que la fermeture d'un simplexe est  $\text{Cl}(\sigma) = \bigcup_{\emptyset \neq \tau \subset \sigma} \{\tau\}$  et notons que la préimage d'une union est l'union des préimages. Nous considérons trois cas :

Cas 1 : Si  $\sigma = c$  alors  $f^{-1}[\text{Cl}(c)] = f^{-1}[\{c\}] = \{a, b, ab\}$  est collapsable.

Cas 2 : Si  $\sigma \cap c = \emptyset$ , alors pour toutes les faces  $\tau$  of  $\sigma$ , nous avons aussi  $\tau \cap c = \emptyset$  et donc  $f^{-1}[\{\tau\}] = \{\tau\}$ . On en déduit :

$$f^{-1}[\text{Cl}(\sigma)] = \bigcup_{\emptyset \neq \tau \subset \sigma} f^{-1}[\{\tau\}] = \bigcup_{\emptyset \neq \tau \subset \sigma} \{\tau\} = \text{Cl}(\sigma).$$

Cas 3 : Si  $\sigma \neq c$  et  $\sigma \cap c \neq \emptyset$  alors  $\sigma = c\sigma'$  où  $\sigma' \in K'$  et  $\sigma' \cap c \neq \emptyset$ . Comme  $\sigma'$  appartient au *link* de  $c$  dans  $K'$ ,  $\sigma' \in \text{Lk}(a) \cup \text{Lk}(b)$  grâce au lemme 3.5. De plus,  $ab$  vérifie la condition du *link*  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  ce qui implique que  $\sigma'$  appartient soit à  $\text{Lk}(ab)$ , soit à  $\text{Lk}(a) \setminus \text{Lk}(ab)$ , soit à  $\text{Lk}(b) \setminus \text{Lk}(ab)$ . Observons que c'est aussi vrai pour toutes les faces  $\tau$  de  $\sigma'$ , on en déduit que pour tout  $\emptyset \neq \tau \subset \sigma'$ ,

$$f^{-1}[\{c\tau\}] = \begin{cases} \{a\tau, b\tau, ab\tau\} & \text{si } \tau \in \text{Lk}(ab), \\ \{a\tau\} & \text{si } \tau \in \text{Lk}(a) \setminus \text{Lk}(ab), \\ \{b\tau\} & \text{si } \tau \in \text{Lk}(b) \setminus \text{Lk}(ab). \end{cases}$$

Puisque  $\text{Cl}(\sigma) = \{c\} \cup \text{Cl}(\sigma') \cup \bigcup_{\emptyset \neq \tau \subset \sigma'} \{c\tau\}$ , on obtient :

$$f^{-1}[\text{Cl}(\sigma)] = \{a, b, ab\} \cup \text{Cl}(\sigma') \cup \bigcup_{\emptyset \neq \tau \subset \sigma'} f^{-1}[\{c\tau\}].$$

En notant  $\Sigma_1 \cdot \Sigma_2 = \{\sigma_1\sigma_2 \mid \sigma_1 \in \Sigma_1, \sigma_2 \in \Sigma_2\}$  et en posant  $L = \text{Cl}(\sigma') \cap \text{Lk}(ab)$  on obtient que

$$f^{-1}[\text{Cl}(c\sigma')] = \begin{cases} \text{Cl}(ab\sigma') & \text{si } \sigma' \in \text{Lk}(ab), \\ \text{Cl}(a\sigma') \cup \{b, ab\} \cup \{b, ab\} \cdot L & \text{si } \sigma' \in \text{Lk}(a) \setminus \text{Lk}(ab), \\ \text{Cl}(b\sigma') \cup \{a, ab\} \cup \{a, ab\} \cdot L & \text{si } \sigma' \in \text{Lk}(b) \setminus \text{Lk}(ab). \end{cases}$$

Ainsi, si  $\sigma' \in \text{Lk}(ab)$ , la préimage  $f^{-1}[\text{Cl}(c\sigma')]$  est clairement collapsable. Si  $\sigma' \in \text{Lk}(a) \setminus \text{Lk}(ab)$ , on peut toujours trouver un ensemble de simplexes  $\lambda_1, \dots, \lambda_k$  dont la fermeture est égale au complexe simplicial  $L = \text{Cl}(\sigma') \cap \text{Lk}(ab)$  et tel que, pour tout  $1 \leq i, j \leq k$ , le simplexe  $\lambda_i$  n'est ni une face ni une coface du simplexe  $\lambda_j$ . Autrement dit, l'ensemble des simplexes  $\lambda_1, \dots, \lambda_k$  est l'ensemble des simplexes maximaux pour l'inclusion de  $L$ . Par construction,  $ab\lambda_i$  est la seule coface propre de  $b\lambda_i$  dans  $f^{-1}[\text{Cl}(c\sigma')]$ . Après une séquence de  $k$  collapsés élémentaires consistant à enlever les paires de simplexes  $(b\lambda_i, ab\lambda_i)$ , il reste un complexe simplicial  $\text{Cl}(a\sigma') \cup \{b, ab\}$  qui est collapsable. Le cas  $\sigma' \in \text{Lk}(b) \setminus \text{Lk}(ab)$  est similaire.  $\square$

### 4.2.3 Tester la condition du *link* efficacement

Nous avons vu une condition locale sur le *link* d'une arête  $ab$  qui garantit que la contraction d'arête  $ab$  préserve le type d'homotopie. Une formulation équivalente faisant intervenir les bloqueurs est possible : une arête  $ab$  vérifie la condition du *link* si et seulement si aucun bloqueur ne contient  $ab$  [Mel11].

**Lemme 4.2.** *Soit  $K$  un complexe simplicial. L'arête  $ab \in K$  vérifie la condition du link  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  si et seulement si aucun bloqueur de  $K$  ne contient  $ab$ .*

*Démonstration.* Il est facile de voir que pour tout complexe simplicial  $K$  et pour toute arête  $ab \in K$ , on a l'inclusion suivante  $\text{Lk}(ab) \subset \text{Lk}(a) \cap \text{Lk}(b)$ . Nous allons montrer que  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$  implique qu'aucun bloqueur de  $K$  ne contient  $ab$ . Par l'absurde, supposons qu'un simplexe  $ab\tau$  soit un bloqueur de  $K$  pour un simplexe  $\tau$  tel que  $ab \cap \tau = \emptyset$ . Par définition d'un

bloqueur, toutes les faces propres de  $ab\tau$  appartiennent à  $K$ , en particulier  $a\tau \in K$  et  $b\tau \in K$ . D'un autre côté,  $ab\tau$  n'appartient pas à  $K$ . Ainsi  $\tau \in \text{Lk}(a)$ ,  $\tau \in \text{Lk}(b)$  et  $\tau \notin \text{Lk}(ab)$ , d'où  $\text{Lk}(ab) \neq \text{Lk}(a) \cap \text{Lk}(b)$ . Réciproquement, supposons qu'aucun bloqueur de  $K$  ne contienne  $ab$  et montrons qu'alors  $\text{Lk}(a) \cap \text{Lk}(b) \subset \text{Lk}(ab)$ . Soit un simplexe  $\sigma \in \text{Lk}(a) \cap \text{Lk}(b)$ . Par définition,  $a\sigma \in K$ ,  $b\sigma \in K$  et  $ab \cap \sigma = \emptyset$ . Montrons que le simplexe  $ab\sigma$  appartient à  $K$  et donc  $\sigma \in \text{Lk}(ab)$ . Raisonnons par l'absurde. Supposons qu'il existe des faces de  $ab\sigma$  qui n'appartiennent pas à  $K$  et soit  $\tau_{\min}$  la face minimale pour l'inclusion parmi ces faces. Autrement dit,  $\tau_{\min}$  est un bloqueur d'ordre 0. Puisque  $\tau_{\min} \subset ab\sigma$ ,  $\tau_{\min} \notin K$ ,  $a\sigma \in K$  et  $b\sigma \in K$ , on a nécessairement  $ab \subset \tau_{\min}$  ce qui contredit l'hypothèse selon laquelle aucun bloqueur d'ordre 1 de  $K$  ne contient  $ab$ .  $\square$

Ainsi tester si une arête vérifie la condition du *link* peut être réalisé simplement en utilisant la structure de données squelette/bloqueurs introduite dans le chapitre 3. En effet, le lemme précédent montre qu'il suffit de parcourir l'ensemble des bloqueurs passant par  $a$  et de vérifier pour chacun d'entre eux qu'il ne contient pas  $b$ , voir l'algorithme 10 dans l'annexe pour le pseudo-code de ce parcours ainsi qu'une discussion sur sa complexité.

#### 4.2.4 Réciproque de la condition du *link*

Nous avons vu qu'une contraction d'arête  $ab \mapsto c$  préserve le type d'homotopie dès que la condition du *link* est vérifiée ou, de manière équivalente, dès qu'aucun bloqueur de  $K$  ne contient  $ab$ . Intéressons-nous maintenant à la réciproque. Lorsque  $K$  est homéomorphe à une variété, une contraction d'arête  $ab \mapsto c$  préserve l'homéomorphisme si et seulement si la condition du *link* est vérifiée [DEG99, Nev06]. Cependant, cette réciproque est fautive pour le type d'homotopie. En effet, la figure 4.2 montre un complexe simplicial  $K$  et une arête  $ab$  telle que la contraction de cette arête dans  $K$  préserve le type d'homotopie alors que la condition du *link* n'est pas vérifiée (puisque le bloqueur  $abx$  passe par  $ab$ ). Notons néanmoins dans ce cas que le bloqueur  $abx$  est redondant (voir la définition 3.2) et qu'il peut donc être enlevé pour donner un nouveau complexe simplicial ayant le même type d'homotopie et dans lequel la contraction d'arête préserve le type d'homotopie et la condition du *link* est vérifiée (aucun bloqueur ne contient alors l'arête  $ab$ ). Cela nous amène à introduire la conjecture suivante.

**Conjecture 4.1.** *Soit  $K$  un complexe simplicial sans bloqueur redondant et  $ab \in K$ . Si  $ab \mapsto c$  préserve le type d'homotopie alors aucun bloqueur ne contient  $ab$ .*

L'énoncé suivant découle de la conjecture précédente et caractérise totalement les contractions d'arêtes qui préservent le type d'homotopie.

**Conjecture 4.2.** *Soit  $K$  un complexe simplicial sans bloqueur redondant. La contraction d'arête  $ab \mapsto c$  préserve le type d'homotopie si et seulement si aucun bloqueur de  $K$  ne contient l'arête  $ab$ .*

Dans le chapitre 3, nous avons vu une structure de données compacte à même de représenter les complexes simpliciaux proches de *flag*-complexes et capable de contracter des arêtes

efficacement. Nous venons de voir une condition, facilement calculable dans notre structure, qui garantit qu'une contraction d'arête préserve le type d'homotopie. La prochaine section utilise tous ces éléments et propose un algorithme de simplification par contraction d'arêtes.

### 4.3 Simplification par contraction d'arêtes

#### 4.3.1 Un algorithme

Pour décrire notre algorithme de simplification d'un complexe simplicial  $K$ , nous introduisons une priorité sur les arêtes que l'on note  $\varphi$ . Dans cet algorithme, nous contractons itérativement les arêtes de  $K$  par ordre de priorité décroissante. Plus précisément, nous commençons par mettre toutes les arêtes de  $K$  dans une pile de priorité  $\mathcal{Q}$ . A chaque étape, on extrait  $ab$  l'arête de priorité maximale de  $\mathcal{Q}$ . Si  $ab$  vérifie la condition du *link*, on contracte  $ab$  sur un nouveau sommet<sup>2</sup>  $c = \frac{a+b}{2}$  et on met à jour la structure de données (ce qui inclut la suppression et l'ajout d'arêtes dans  $\mathcal{Q}$ ). Durant cette étape, nous enlevons également tous les bloqueurs simplement redondants dont la suppression ne change pas le type d'homotopie (voir la définition 3.2). L'algorithme s'arrête lorsqu'il n'y a plus d'arêtes dans  $\mathcal{Q}$  (il est fini puisque chaque contraction ôte un sommet à  $K$ ). Voir l'algorithme 3 pour le pseudo-code du procédé de simplification.

---

**Algorithme 3** Simplifie un complexe simplicial  $K$  par contraction d'arêtes

---

```

1:  $\mathcal{Q} = \text{Edges}(K)$ 
2: while  $\mathcal{Q} \neq \emptyset$  do
3:   Extraire l'arête  $ab$  de priorité maximale dans  $\mathcal{Q}$ 
4:   if  $(\text{Lk}_K(ab)) = \text{Lk}_K(a) \cap \text{Lk}_K(b)$  then
5:     Contracter l'arête  $ab$  sur  $c = \frac{a+b}{2}$  dans  $K$ 
     {On met à jour les arêtes de  $\mathcal{Q}$ }
6:     for all  $ax \in \text{Edges}(K)$  do
7:        $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{ax\}$ 
8:        $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{cx\}$ 
9:     for all  $bx \in \text{Edges}(K)$  do
10:       $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{bx\}$ 
11:       $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{cx\}$ 
12:   Enlever les bloqueurs de  $K$  qui sont simplement redondants

```

---

Cet algorithme est implémentable dans plusieurs structures de données. Nous l'avons implémenté dans la structure squelette/bloqueurs décrite au chapitre 3 puisque les complexes utilisés par la suite seront initialement des *flag-complexes* et que cette structure est compacte dans ce cas. En particulier, nous avons vu dans ce chapitre comment contracter une arête dans cette structure ou encore comment enlever les bloqueurs simplement redondants.

---

2. Lorsque le complexe simplicial est abstrait on peut prendre  $c = a$  à la place de  $c = \frac{a+b}{2}$

### 4.3.2 Priorité sur les arêtes

L'algorithme de simplification introduit une fonction de priorité sur les arêtes. Nous discutons ici d'un choix possible pour cette priorité. Deux cas sont à distinguer : le cas où le complexe est abstrait et le cas où ses sommets sont dans un espace vectoriel normé. Nous étudions ici le deuxième cas qui est le plus facile. En effet, lorsque les points sont dans un espace vectoriel normé, la distance permet de réaliser des opérations utiles pour la priorité ce qui n'est plus possible lorsque le complexe est abstrait.

Une priorité sur les arêtes pour un algorithme de simplification par contraction d'arête similaire à celui que nous avons introduit a été donné par Garland et al [GH97]. Cette priorité permet de simplifier des surfaces triangulées plongées dans  $\mathbb{R}^3$  en des approximations possédant un nombre de sommets très inférieur au nombre de sommets initial tout en ayant une faible erreur géométrique. Malheureusement, il n'est pas clair de savoir comment étendre cette priorité dans le cas de complexes simpliciaux qui ne sont pas plongés comme le complexe de Rips, par exemple.

Par la suite, pour les complexes géométriques, nous utiliserons une priorité simple. Cette priorité associe à une arête  $ab$  l'inverse de sa distance i.e.  $-d(a, b)$ . Les arêtes les plus petites sont alors contractées en premier ce qui correspond à l'intuition que les détails d'une forme sont les éléments de petite taille que l'on veut simplifier en premier. Les figures 4.3 et 4.4 illustrent le procédé de simplification. On constate en particulier que les contractions d'arêtes semblent lisser la densité des points comme le montre la figure 4.4. Pour être plus précis, on introduit la définition suivante, qui sera réutilisée dans le chapitre suivant :

**Définition 4.1.** Soient  $A, L \subset \mathbb{R}^D$  où  $L$  est fini. On dit que  $L$  est un  $(\varepsilon, \delta)$ -échantillon de  $A$  si (1)  $d_H(A, L) \leq \varepsilon$  et (2)  $L$  est un  $\delta$ -packing de  $A$ . Dans le cas où  $L$  est un  $(\varepsilon, \varepsilon/2)$ -échantillon de  $A$ , on dit que  $L$  est un  $\varepsilon$ -net de  $A$ .

Il semble que l'ensemble de points  $P_N$  obtenu après  $N$  contractions soit de plus en plus clairsemé tout en restant proche de l'ensemble des points initiaux  $P$ . Plus précisément, il semble qu'il existe une fonction  $\varphi$  croissante telle que  $P_N$  soit un  $\varphi(N)$ -net de  $P$ . Il serait intéressant de pouvoir prouver un tel résultat qui apporterait un éclairage théorique nouveau sur les contractions d'arêtes. Une piste est d'utiliser les *net-trees* utilisés récemment par Sheehy [She12] pour extraire une filtration compacte de complexes de Rips. Il s'agit d'une structure hiérarchique qui encode des  $\beta$ -net de  $P$  pour différentes échelles  $\beta$  dans un arbre. Puisque les contractions d'arêtes produisent une structure similaire en pratique, un objectif serait de prouver que celles-ci calculent bien un *net-tree*.

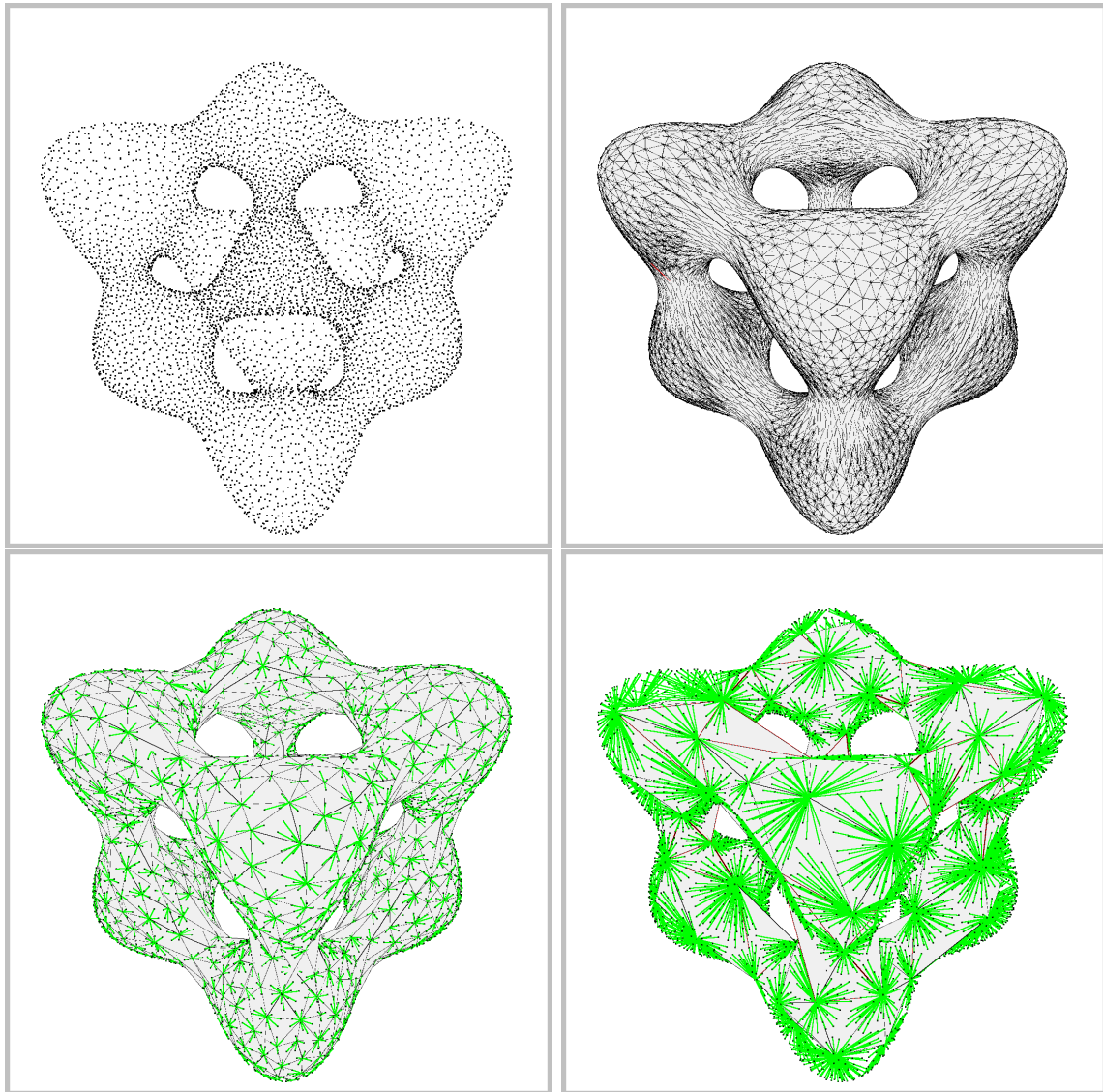


FIGURE 4.3 – Illustration de la simplification par contractions d'arêtes. En haut à gauche, l'ensemble de points `Pillow`. En haut à droite, Le complexe de Rips construit avec ces points. En bas à gauche, le complexe obtenu après 5807 contractions d'arêtes (il reste alors 1000 sommets sur les 6807 sommets initiaux) ; chaque sommet du complexe est lié à tous les points de ses préimages par des segments verts clairs. En bas à droite, le complexe obtenu après 6707 contractions, il ne reste alors que 100 sommets.

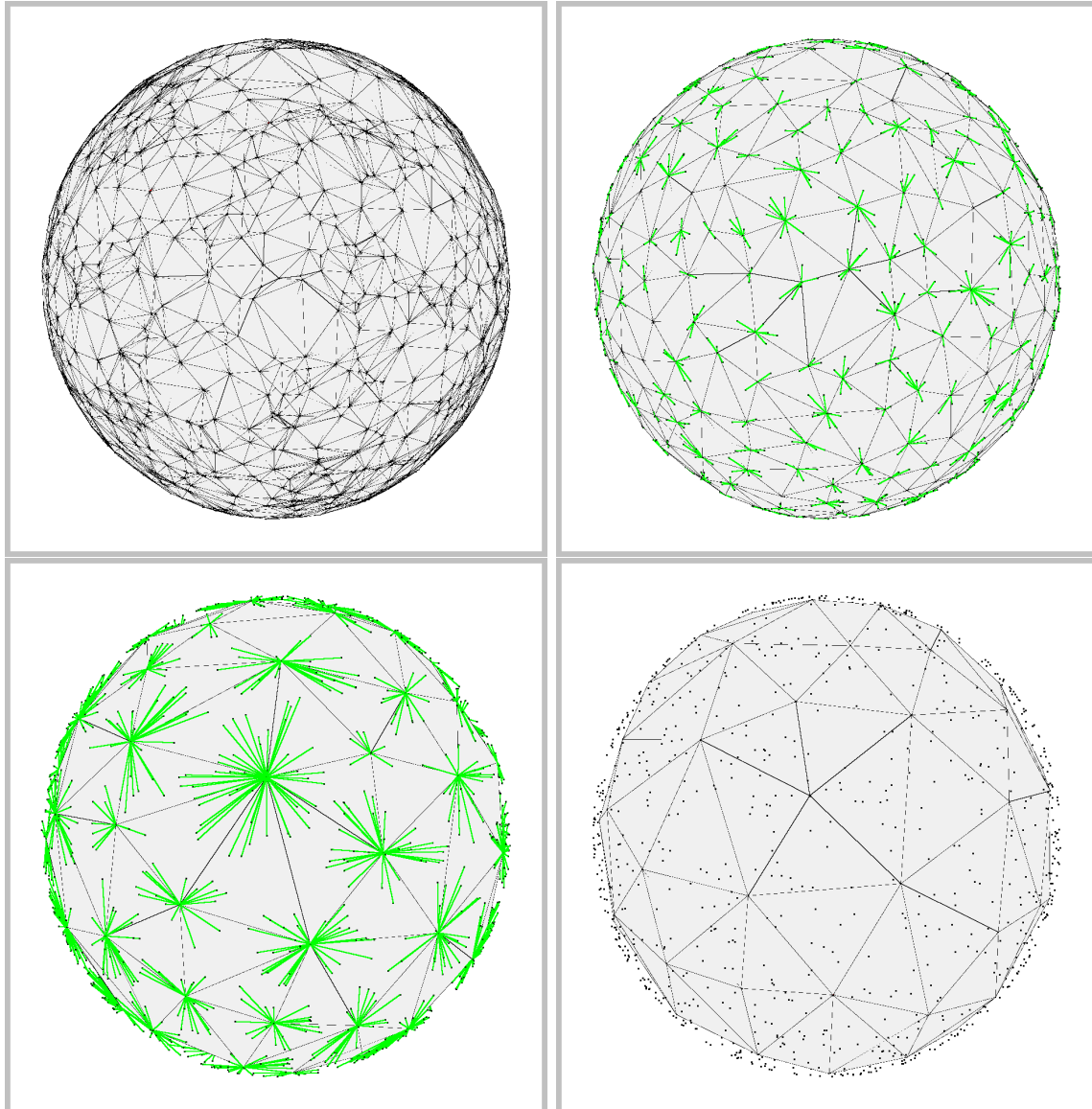


FIGURE 4.4 – Illustration de la simplification par contractions d'arêtes. En haut à gauche, le complexe de Rips d'un ensemble de 2000 points obtenus en échantillonnant uniformément  $S_2$ . En haut à droite, le complexe obtenu après 1500 contractions d'arêtes (il reste alors 500 sommets sur les 2000 sommets initiaux) ; chaque sommet du complexe est lié à tous les points de ses préimages par des segments verts clairs. En bas à gauche, le complexe obtenu après 1900 contractions, il ne reste alors que 100 sommets. En bas à droite, le complexe obtenu après 1900 contractions ainsi que les points initiaux.



## 4.4 Résultats expérimentaux

### 4.4.1 Jeu de données

Chaque élément de notre jeu de données correspond à un ensemble de points  $P \subset \mathbb{R}^D$  échantillonnant une  $d$ -variété  $A$ . Dans tous les cas, nous construisons un complexe de Rips  $K = \text{Rips}(P, \alpha)$ , en choisissant  $\alpha$  de telle sorte que  $K$  possède les mêmes groupes d'homologie que  $A$ . Lorsque la dimension ambiante  $d$  est strictement plus petite que 3, nous avons en plus  $K \simeq A$  pour chaque complexe de Rips. Des statistiques concernant les différents nuages de points  $P$  ainsi que leur complexe de Rips peuvent être trouvées sur la table 4.1.

**Cat.** Le premier jeu de données est constitué de 72 images d'une statue en forme de chat placée sur une table tournante et observée par une caméra fixe [NNM96]. Les images de cette statue sont prises par intervalle de 5 degrés. Chaque image a une taille de  $128^2 = 16384$  pixels et peut être vue comme un point de  $\mathbb{R}^{16384}$ . Puisque le degré du système d'acquisition est 1, l'ensemble des images peut être vu comme un échantillonnage d'une courbe plongée dans  $\mathbb{R}^{16384}$ . Quelques images de cet ensemble sont présentées sur la figure 4.5. Un complexe de Rips bâti à partir de cet ensemble et ayant le même type d'homotopie qu'une courbe est représenté sur la figure 4.6.

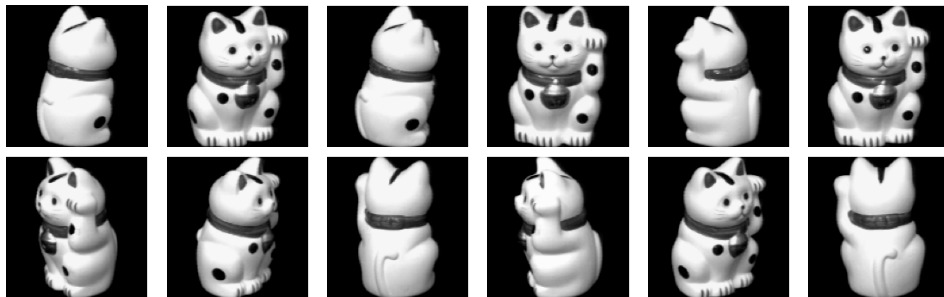


FIGURE 4.5 – Douze points parmi les 72 points de l'ensemble **Cat**. Chaque point représente une image de  $128 \times 128$  pixels d'une statue en forme de chat prise sous différentes poses durant sa rotation autour d'un axe fixe.

**C2, C3, C4.** Pour construire  $P = \text{Cd}$ , nous échantillonnons régulièrement la frontière d'un  $(d + 1)$ -cube  $C_d = \partial[-1, 1]^{d+1}$  plongé dans  $\mathbb{R}^{d+1}$ . Pour se faire, nous échantillonnons chaque  $2d + 2$  face de  $C_d$  en utilisant une grille régulière de  $k^d$  points. En utilisant cette méthode pour  $k = 21$  et  $d = 2$ , on obtient le premier ensemble de points **C2** de taille  $6 \times 21^2 = 2646$ . En appliquant cette méthode avec  $k = 18$  et  $d = 3$ , on obtient le second ensemble **C3** de taille  $8 \times 18^3 = 46656$ . Enfin, avec  $k = 4$  et  $d = 4$ , on obtient **C4** de taille 2560.

**S2, S3** Chaque ensemble de points **Sd** est obtenu en normalisant les points de **Cd**.

$P$	$\#P$	$\#Edges(Rips(P, \alpha))$	$\#Rips(P, \alpha)$	d	D
Cat	72	645	$1, 2 \cdot 10^6$	1	16.384
C2	2.646	80.304	$12 \cdot 10^9$	2	3
S2	2.646	149.484	$> 12 \cdot 10^9$	2	3
C3	46.656	3.029.824	$> 3 \cdot 10^9$	3	4
S3	46.656	7.420.640	$> 3 \cdot 10^9$	3	4
C4	2.560	122.000	$1, 3 \cdot 10^8$	4	5
Tore	1.307	20846	$6, 0 \cdot 10^6$	2	3
Pillow	6.806	112.582	$73 \cdot 10^6$	2	3
Ramses	193.252	1.402.921	$6, 6 \cdot 10^6$	2	3
S03	50.000	1.025.990	$1, 1 \cdot 10^8$	3	9

TABLE 4.1 – Statistiques, pour chaque ensemble  $P$ , du complexe de Rips construit avec  $P$ . Nous indiquons en particulier le nombre de points de  $P$ , le nombre d’arêtes et le nombre de simplexes du complexe de Rips, la dimension intrinsèque de  $A$  et la dimension ambiante des points de  $P$ .

**Ramses.** Cet ensemble de points est obtenu avec un scanner 3D à partir d’une statue représentant Ramses II. La surface de la statue est homéomorphe à  $S_2$  et l’ensemble de ces points est disponible dans la base de données *Aim@Shape*.

**Pillow.** Ce nuage de points de  $\mathbb{R}^3$  échantillonne une surface homéomorphe à une somme connectée de 5 tores. Voir figure 4.3.

**Tore.** Cet ensemble de points  $P \subset \mathbb{R}^3$  est constitué de 1307 points distribués sur un tore.

**S03.** Le dernier jeu de données est obtenu en échantillonnant le groupe orthogonal  $SO(3)$  avec la méthode décrite dans [YJLM10]. Nous obtenons un ensemble de 10000 points  $S03 \subset \mathbb{R}^9$ . Rappelons que le groupe orthogonal  $SO(3)$  est diffeomorphe à  $\mathbb{RP}^3$  qui est une 3-variété qui peut être plongée dans  $\mathbb{R}^9$  en représentant chaque rotation 3D par une matrice  $3 \times 3$ .

#### 4.4.2 Taille de la structure de données au cours des contractions

Nous appliquons maintenant l’algorithme de simplification par contraction d’arête présenté à la section 4.3 sur notre jeu de données. Rappelons qu’une contraction d’arête diminue le nombre de sommets de un. Nous notons  $K_i$  le complexe simplicial obtenu après  $i$  contractions d’arêtes et nous posons  $n_i = \#Vert(K_i) = \#Vert(K_0) - i$ . Pour un sous-ensemble représentatif  $P \in \{C2, S2, C3, S3, S03, Ramses\}$  de notre jeu de données, nous traçons le nombre de bloqueurs, le nombre cumulé de bloqueurs simplement redondant qui ont été retirés à l’étape  $i$ , ainsi que le degré moyen de chaque sommet ; voir la figure 4.7. Le complexe simplicial obtenu par

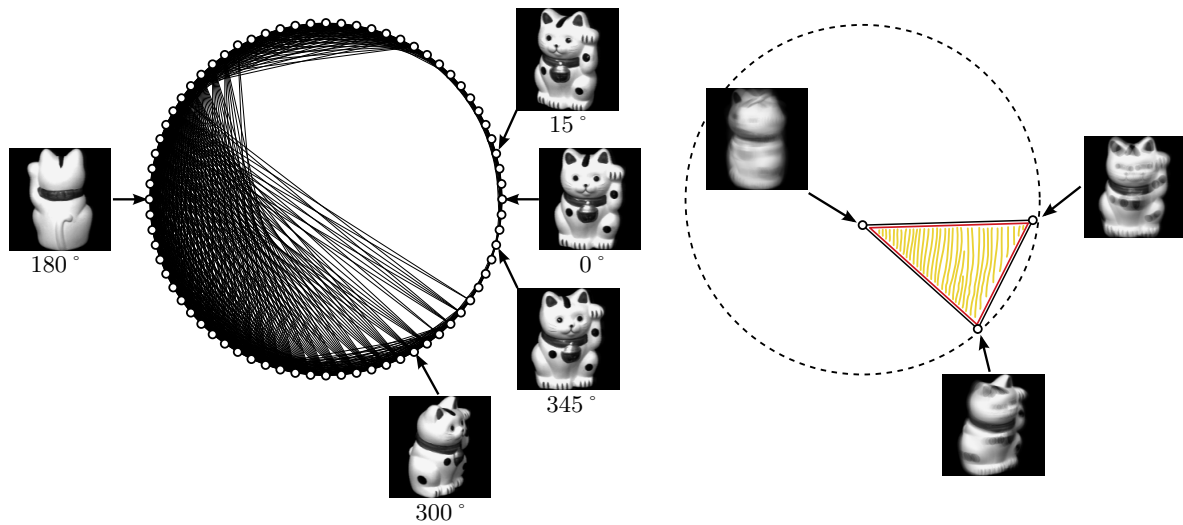


FIGURE 4.6 – Gauche : Complexe de Rips construit avec les points de *Cat*. Droite : Après simplification, on obtient un complexe simplicial constitué de trois sommets, trois arêtes et un unique bloqueur.

contraction d'arêtes à partir de *Cat* est illustré sur la figure 4.6. Il est à noter le faible nombre de bloqueurs observés lors de la simplification, le complexe est donc presque toujours un *flag-complexe* ce qui permet d'effectuer efficacement les contractions dans la structure de données squelette/bloqueurs. Enfin, puisque le nombre de bloqueurs reste faible et que le nombre de sommets et d'arêtes diminue au cours des contractions, la taille de notre structure diminue également durant le procédé de simplification.

Notons que la métrique est un élément clé pour le bon fonctionnement du procédé de simplification. En effet, celle-ci nous permet de donner une priorité aux arêtes qui induit, en pratique, un faible nombre de bloqueurs. Avec une priorité aléatoire sur les arêtes, les résultats obtenus sont beaucoup moins bons puisqu'un nombre important de bloqueurs apparaît au cours des contractions d'arêtes. La simplification est alors plus lente et aboutit sur des complexes avec généralement plus de sommets qu'en utilisant la priorité correspondant à l'inverse de la longueur des arêtes. Par exemple, la contraction du complexe construit avec *Tore* prend 6 minutes avec la priorité aléatoire et seulement 2 secondes avec la priorité correspondant à l'inverse de la longueur des arêtes.

#### 4.4.3 Taille des complexes après contraction sur le jeu de données

Pour plusieurs applications (comme le calcul de groupes d'homologie par exemple), il est préférable que le complexe obtenu après simplification  $K_{out}$  soit de petite taille et comporte peu de sommets. Afin de comparer la taille des complexes que l'on obtient, nous avons besoin de plusieurs définitions. Un complexe homéomorphe à une variété  $A$  est dit *minimal* si tout complexe homéomorphe à  $A$  possède autant ou plus de sommets. Le *nombre de sommets minimal* d'une variété  $A$  est le nombre de sommets d'un de ses complexes minimaux. Enfin,

un complexe simplicial est dit *irréductible* si on ne peut contracter aucune de ses arêtes sans changer le type d'homotopie. Pour chaque variété  $A$  de notre jeu de données, nous indiquons dans la table 4.2 le nombre de sommets minimal de  $A$ . Nous indiquons également dans cette table le nombre de sommets du complexe  $K_{out}$  obtenu après contraction d'arêtes et le temps de calcul nécessaire pour effectuer les contractions d'arêtes<sup>3</sup>. On remarque en particulier que le nombre de sommets de  $K_{out}$  est toujours proche du nombre de sommets minimal de  $A$ . Il serait intéressant de pouvoir caractériser les complexes irréductibles que l'obtient par contractions d'arêtes en partant d'un complexe de Rips reconstruisant une variété. On remarque qu'en pratique, on arrive souvent au bout d'un certain nombre de contractions à un complexe possédant la même dimension que la variété voir même à un complexe homéomorphe à la variété. Or, plusieurs résultats sont connus sur les complexes irréductibles que l'on peut obtenir avec des contractions d'arêtes à partir de complexes homéomorphes à une variété  $A$ . En particulier, si la variété  $A$  échantillonnée est une surface et que le complexe est une triangulation, l'ensemble des complexes irréductibles est fini [BE89] et il est connu dans plusieurs cas. En effet, lorsque  $A = S_2$  il est connu [SR34] que la seule triangulation irréductible est  $K_4$ . Cela explique intuitivement pourquoi notre algorithme de simplification retourne un complexe ayant pour 1-squelette  $K_4$  pour les jeux de données **C2**, **S2**, **Ramses**. Lorsque  $A$  est un tore, il existe 21 triangulations irréductibles possibles comportant 7,8 ou 9 sommets [Law87]. Là encore, on peut remarquer que notre algorithme renvoie un complexe à 9 sommets pour **Tore**. Dans le cas général, si  $A$  est une surface, le nombre de sommets de toute triangulation irréductible est borné par  $\max\{13g - 4, 4\}$  où  $g$  est le genre de la surface [JW10]. Lorsque  $A$  est une variété de dimension 3, caractériser les complexes irréductibles est bien plus difficile. En effet, dans le cas le plus simple de la 3-sphère, on sait reconnaître la variété [Rub95, Tho94] mais même pour ce cas "simple" il existe une infinité de complexes irréductibles [DEGN98]. Lorsque  $A$  est une variété de dimension 4, le problème est difficile puisque décider si deux 4-variétés sont homéomorphes est indécidable. Cependant, sous certaines conditions géométriques, on peut peut-être espérer reconnaître des variétés de dimension supérieure à 3.

Afin de pouvoir appliquer ces résultats sur les complexes irréductibles, on voudrait pouvoir extraire un complexe de  $K$  qui soit homéomorphe à  $A$ . En pratique, on observe qu'on obtient assez souvent un tel complexe après plusieurs contractions d'arêtes en partant d'un complexe  $K = \text{Rips}(P, \alpha)$ . Le chapitre suivant cherche à extraire un complexe homéomorphe à  $A$  de  $K$  par application itérative de collapses sur  $K = \text{Rips}(P, \alpha)$ .

Terminons sur une remarque quant à la complexité du calcul des groupes d'homologie d'un *flag-complexe*  $K$ . Etant donné le graphe encodant le 1-squelette de  $K$ , le problème de calculer les groupes d'homologies d'un *flag-complexe*  $K$  est NP-dur [AS12]. Cependant, ce problème est polynomial en la taille du graphe lorsque sont connues des hypothèses géométriques sur  $K$  comme, par exemple, lorsque les sommets de  $K$  échantillonnent une variété [CO08]. Supposons que  $K$  soit un complexe de Rips approximant une surface et qu'on sache en extraire une triangulation et la contracter en temps polynomial en la taille du graphe. On obtient alors un complexe simplicial  $K_{out}$  dont le nombre de sommets est borné par  $\max\{13g - 4, 4\}$  [JW10]. Le nombre de simplexes de ce complexe est alors borné par  $2^{\max\{13g-4, 4\}}$ . Ainsi le

3. Dans toute cette thèse, les calculs sont effectués sur un processeur AMD cadencé à 3.4 GHz disposant de 8 Go de mémoire vive.

calcul des groupes d'homologie peut se faire à partir de  $K_{out}$  en  $O(2^{3 \max\{13g-4,4\}})$  soit en temps polynomial en la taille du graphe de  $K$ . La complexité totale pour calculer les groupes d'homologie de  $K$  est alors polynomiale en la taille du graphe du 1-squelette de  $K$ . On pourrait obtenir un algorithme similaire à [CO08] permettant de calculer efficacement les groupes d'homologie de surfaces en grande dimension. Contrairement à cette approche, on n'aurait pas besoin d'énumérer les simplexes du complexe de Rips initial mais seulement ceux du complexe obtenu après simplification.

$P$	$\sharp P$	Nombre de sommets de $K_{out}$	Nombre de sommets minimal	Temps de calcul
Cat	72	3	3	2 s
C2	2646	4	4	20 s
S2	2646	4	4	1,3 min
C3	46656	5	5	60 min
S3	46656	5	5	1080 min
C4	2560	6	6	5 min
Tore	1307	8	7	2 s
Pillow	6806	19	$\geq 12$	10 s
Ramses	193252	4	4	1 min
S03	50000	13	11	8 min

TABLE 4.2 – Statistiques, pour chaque ensemble  $P$ , du complexe de Rips construit avec  $P$  ainsi que du complexe obtenu après contraction  $K_{out}$  en utilisant la priorité de la plus petite arête.

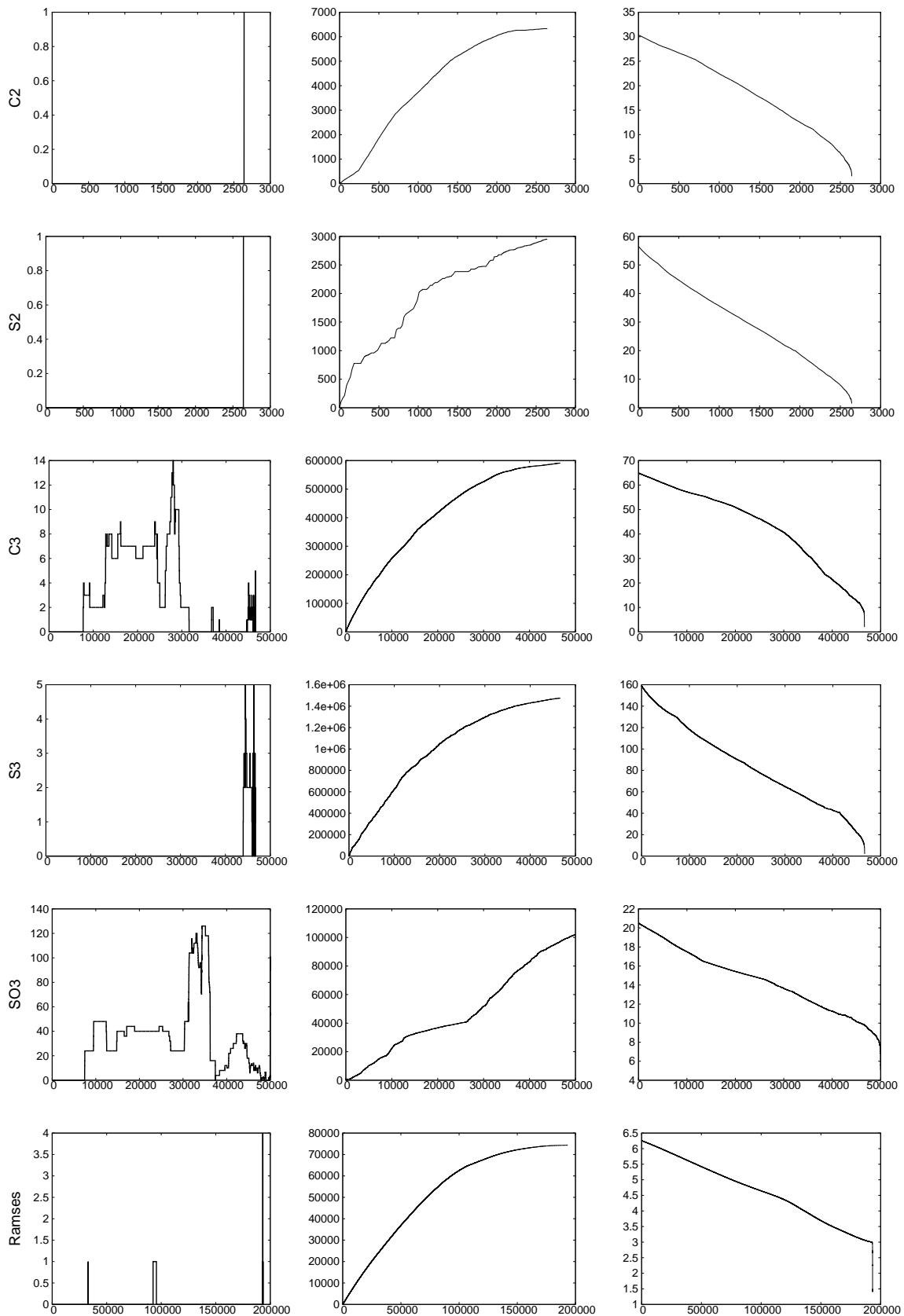


FIGURE 4.7 – Nous traçons en fonction de  $i$ , le nombre de bloqueurs dans  $K_i$  à gauche, le nombre cumulé de bloqueurs redondants enlevés de la structure de données aux étapes antérieures à  $i$  au milieu, et le degré moyen de chaque sommet dans  $K_i$  à droite. De haut en bas, le jeu de données est respectivement C2, S2, C3, S2, S03, et Ramses.



# Simplification de *flag*-complexes par collapses

## Sommaire

<b>5.1</b>	<b>Introduction</b>	<b>69</b>
<b>5.2</b>	<b>Collapser un complexe de Rips dont les sommets approximent un convexe</b>	<b>71</b>
5.2.1	En passant par le complexe de Čech	72
5.2.2	Préserver un <i>flag</i> -complexe durant la simplification	74
5.2.3	Augmenter le paramètre d'échelle	76
5.2.4	Algorithmes pour trouver une séquence de réduction.	77
<b>5.3</b>	<b>Simplifier un complexe de Rips dont les sommets échantillonnent une variété</b>	<b>78</b>
5.3.1	Extraire un sous-ensemble de points sans changer le type d'homotopie	78
5.3.2	Algorithme de simplification.	81
5.3.3	Preuve de correction de l'algorithme de reconstruction dans un cas simple	82
<b>5.4</b>	<b>Résultats expérimentaux</b>	<b>84</b>
5.4.1	Comparaison des différentes stratégies pour REDUCTIBLE	84
5.4.2	Résultats des simplifications sur le jeu de données	85

## 5.1 Introduction

Etant donnée une forme connue seulement à partir d'un nuage de points, le problème de reconstruction topologique vise à construire une approximation de la forme ayant la même topologie que celle-ci. Par même topologie, on entend que la forme et son approximation aient le même type d'homotopie ou encore qu'elles soient homéomorphes. Nous avons pu voir dans le chapitre 2 des conditions assurant que le complexe de Rips possède le même type d'homotopie qu'une forme lorsque les sommets du complexe échantillonnent la forme avec une densité suffisante. Le complexe de Rips possède l'avantage conséquent d'être un *flag*-complexe : son 1-squelette suffit pour le décrire implicitement. En tirant parti de cette propriété, nous avons introduit une structure de données à même de représenter ce complexe de manière compacte dans le chapitre 3. Cependant, la dimension de la forme n'est malheureusement pas directement accessible dans cette représentation puisque la dimension du complexe de Rips



est généralement élevée. Dans ce chapitre, nous cherchons à simplifier le complexe de Rips de manière à obtenir un complexe ayant la même dimension que la forme voire même de manière à obtenir un complexe homéomorphe à la forme.

Extraire la dimension d'une forme à partir d'un ensemble de points l'échantillonnant est plus simple que de construire une approximation homéomorphe à la forme. Des auteurs ont proposé des conditions d'échantillonnage sous lesquelles il est possible de calculer la dimension de la forme en temps polynomial en la dimension des points [GW03]. Le problème de reconstruction d'une approximation homéomorphe à la forme en grande dimensions a été traité par [DGGZ02, BGO07] néanmoins ces algorithmes de reconstruction possèdent une complexité exponentielle en la dimension ambiante. Boissonnat et Ghosh ont proposé une méthode qui calcule localement des triangulations de Delaunay sur des plans tangents et recolle ensuite ces triangulations locales [BG10]. Sous certaines conditions géométriques, leur approche produit un complexe homéomorphe à la forme échantillonnée en temps polynomial en la dimension ambiante.

Nous cherchons à obtenir des résultats similaires en utilisant le complexe de Rips. Dans ce but, nous étudions dans ce chapitre la simplification de *flag*-complexes par une opération préservant le type d'homotopie appelée *collapse*. Cette opération consiste à enlever un simplexe et toutes ses cofaces lorsque le *link* de ce dernier est contractile. Décider si un complexe est contractile est un problème difficile car indécidable [Tan12]. Même si on se réduit au problème de reconnaître un complexe collapsable, le problème est toujours NP-complet [Tan12]. Cependant, plusieurs travaux ont proposé des conditions géométriques assurant qu'un complexe est collapsable [AB11, Chi67]. Chillingworth a notamment démontré qu'un 3-complexe convexe est collapsable. Dans le cas d'un complexe de Rips dont les points se situent dans le plan, on peut reconnaître polynomialement la contractibilité [CdSEG10]. Nous prouvons dans ce chapitre plusieurs conditions portant sur la convexité d'un ensemble de points assurant qu'un complexe est collapsable et donc contractile. En particulier, lorsque les points approximent un convexe, on peut toujours trouver une séquence de collapses pour réduire le complexe de Rips en un point. Nos preuves étant constructives, elles donnent des oracles pour la réduction d'un complexe en un point. En utilisant ces derniers, nous introduisons un algorithme de simplification consistant à collapser itérativement les sommets et les arêtes d'un *flag*-complexe.

Tout le long de ce chapitre, nous utiliserons intensément les complexes de Čech et de Rips dont la définition a été donnée dans le premier chapitre. Lorsque l'on simplifie le complexe de Rips, on utilisera souvent le fait que le *link* de n'importe lequel de ses simplexes est encore un complexe de Rips. Précisément, en notant  $\mathcal{B}(\sigma, \alpha) = \bigcap_{z \in \sigma} B(z, \alpha)$  pour l'intersection commune des boules de rayon  $\alpha$  centrées en  $\sigma$ , le *link* de  $\sigma$  dans  $\text{Rips}(P, \alpha)$  est  $\text{Rips}(P \cap \mathcal{B}(\sigma, \alpha) \setminus \sigma, \alpha)$ .

*Ce chapitre reprend une partie des résultats co-écrits avec Dominique Attali et André Lieutier et publiés dans l'article [ALS13].*

## 5.2 Collapser un complexe de Rips dont les sommets approximent un convexe

Nous cherchons dans cette section à démontrer qu'un complexe de Rips est collapsable lorsque ses sommets approximent un convexe. Vu l'importance de l'opération de collapse dans ce chapitre, nous rappelons la définition des différents types de collapsages (qui ont été définis dans le chapitre 1).

Un *collapse élémentaire* est l'opération qui enlève une paire de simplexes  $(\sigma_{\min}, \sigma_{\max})$  lorsque  $\sigma_{\max}$  est l'unique coface propre de  $\sigma_{\min}$ . Le résultat de cette opération est un complexe simplicial  $K \setminus \{\sigma_{\min}, \sigma_{\max}\}$  qui se déforme en se rétractant sur  $K$ . L'opération inverse qui ajoute les deux simplexes  $\sigma_{\min}$  et  $\sigma_{\max}$  à  $K$  est un *anti-collapse élémentaire* et il s'agit également d'une opération préservant le type d'homotopie.

Nous listons plusieurs généralisations possibles de l'opération de collapse élémentaire. L'opération qui ôte  $\text{St}_K(\sigma_{\min})$  à  $K$  est appelée :

- un *collapse (classique)* si le *star* de  $\sigma_{\min}$  a un unique élément maximal pour l'inclusion  $\sigma_{\max} \neq \sigma_{\min}$  [DEG99]; l'opération inverse est appelée un *anti-collapse (classique)*.
- un *collapse (étendu)* si le *link* de  $\sigma_{\min}$  est un cône; l'opération inverse est appelée un *anti-collapse (étendu)* [BM12a];
- un *collapse (généralisé)* si le *link* de  $\sigma_{\min}$  est contractile; l'opération inverse est appelée *anti-collapse (généralisé)*.

Les trois collapsages (classiques, étendus, généralisés) préservent le type d'homotopie. Les deux premiers peuvent être exprimés comme une composition de collapsages élémentaires (voir le lemme 1.1). Le dernier type de collapse peut lui être exprimé comme une séquence de collapsages et d'anti-collapsages classiques comme le montre la proposition 1.2.

Tester si l'opération qui enlève  $\text{St}_K(\sigma_{\min})$  à  $K$  est un collapse classique ou étendu peut être fait en temps polynomial comme l'a montré le chapitre 3. Cependant, tester si l'opération enlève  $\text{St}_K(\sigma_{\min})$  à  $K$  est un collapse généralisé est plus délicat. Nous nous intéresserons ici à des complexes de Rips dont les sommets approchent un convexe. La convexité sera un élément clé pour nous permettre de trouver des séquences de réduction lorsque l'ensemble des sommets du Rips échantillonne un convexe.

**Convexité et collapsibilité.** Soit  $P \subset \mathbb{R}^D$  un ensemble fini et non vide de points. On note  $\lambda_P = d_H(\text{Hull}(P), P)$  la constante exprimant le défaut de convexité "global" de  $P$ <sup>1</sup>. Le problème que l'on se pose est : existe-t-il une constante  $\beta > 0$  telle que l'énoncé suivant soit vérifié ?

$$\forall \alpha \geq \beta \lambda_P, \text{Rips}(P, \alpha) \text{ est collapsable} \quad (5.1)$$

En premier lieu, notons que si  $\beta = \frac{\text{Diam} P}{2\lambda_P}$ , le complexe  $\text{Rips}(P, \alpha)$  est égal au complexe  $\text{Rips}(P, \frac{\text{Diam} P}{2})$  qui est la fermeture d'un simplexe. Dans ce cas, l'énoncé 5.1 est vérifié puisque

---

1. En particulier, remarquons que  $\lambda_P = h_P(\text{Rad}(P)) = h_P(+\infty)$  où  $h_P$  est la fonction de défaut de convexité introduite dans le chapitre 2.

la fermeture d'un simplexe est collapsable mais  $\beta$  dépend ici de  $P$ .

Dans cette section, on montre que l'énoncé 5.1 tient pour  $\beta = \frac{1}{2-\vartheta_D}$ .<sup>2</sup> Se pose alors la question de trouver la constante minimale  $\beta$  telle que l'énoncé soit encore vérifié. On montre également que, pour  $D < 3$ , cette constante est optimale : l'énoncé 5.1 n'est plus vérifié pour toute constante inférieure à  $\frac{1}{2-\vartheta_D}$ . Se pose alors le problème de calculer efficacement une séquence de collapses du complexe  $\text{Rips}(P, \alpha)$  vers un point lorsque  $\alpha \geq \beta \lambda_P$ .

Le problème de trouver une séquence de collapses d'un complexe vers un point est malheureusement NP-complet [Tan12]. Ainsi, pour trouver une séquence de collapses efficacement, il est nécessaire d'utiliser la "pseudo-convexité" de  $P$  pour guider l'ordre des collapses. La preuve montrant que le complexe  $\text{Rips}(P, \alpha)$  est collapsable lorsque  $\alpha \geq \frac{\lambda_P}{2-\vartheta_D}$  est constructive mais elle impose, entre autres, de construire un complexe de Čech dont la complexité peut être importante. En augmentant la constante  $\beta$ , nous montrons qu'on peut trouver efficacement une séquence de collapses d'un complexe de Rips vers un point. Plus précisément, nous présentons deux preuves constructives montrant que  $\text{Rips}(P, \alpha)$  est collapsable lorsque  $\alpha \geq \frac{\lambda_P}{2-\sqrt{3}}$ . Ces preuves induisent deux algorithmes efficaces, transformant le complexe en un point tout en maintenant un *flag-complexe* durant la simplification. Nous nous en servirons dans la section suivante pour trouver des séquences de réduction lorsque l'on cherchera à simplifier un complexe de Rips approximant une variété.

Par simplicité, on supposera par la suite que les points  $P$  sont en position générale. Cette hypothèse n'est pas trop restrictive puisqu'on peut s'en passer en utilisant des arguments similaires à ceux employés dans le chapitre 2.

### 5.2.1 En passant par le complexe de Čech

La première stratégie repose sur le résultat du chapitre 2 qui donne des conditions sous lesquelles il existe une séquence de collapses élémentaires de  $\text{Rips}(P, \alpha)$  vers  $\text{Cech}(P, \alpha)$ . La séquence est obtenue en considérant la famille de complexes  $\{\text{Cech}(P, t) \cap \text{Rips}(P, \alpha)\}_{t \geq 0}$  et en surveillant les changements qui apparaissent lorsque  $t$  décroît continûment de  $\vartheta_D \alpha$  vers  $\alpha$ . Nous montrons ci-dessous (théorème 5.1) qu'une fois  $\text{Cech}(P, \alpha)$  obtenu, il peut être ensuite réduit à un point par une séquence de collapses élémentaires. En notant  $B^\circ(p, r)$  la boule ouverte centrée en  $p$  de rayon  $r$  et  $P^{\circ r} = \bigcup_{p \in P} B^\circ(p, r)$ , nous avons le théorème suivant :

**Théorème 5.1.** *Soit  $P \neq \emptyset$  un ensemble fini de points dans  $\mathbb{R}^D$  et  $\alpha > \lambda_P$ . Le complexe  $\text{Cech}(P, \alpha)$  est collapsable.*

*Démonstration.* Il suffit d'établir le théorème pour  $\alpha = 1$ . Soit  $o$  un point de l'enveloppe convexe de  $P$  i.e.  $o \in \text{Hull}(P)$ . Notons  $B_x$  pour la boule unité fermée centrée en  $x$  et soit  $t \geq 0$ . On définit le complexe simplicial  $K_t = \text{Nrv}\{B(o, t) \cap B_p \mid p \in P\}$ ; voir la figure 5.1. Notons que  $K_{+\infty} = \text{Cech}(P, 1)$ .  $K_0$  est collapsable. En effet, l'ensemble de sommets de

2. On rappelle que  $\vartheta_D = \sqrt{\frac{2D}{D+1}}$  est la constante introduite dans le chapitre 1.2.

## 5.2. Collapsing a Rips complex whose vertices approximate a convex set 73

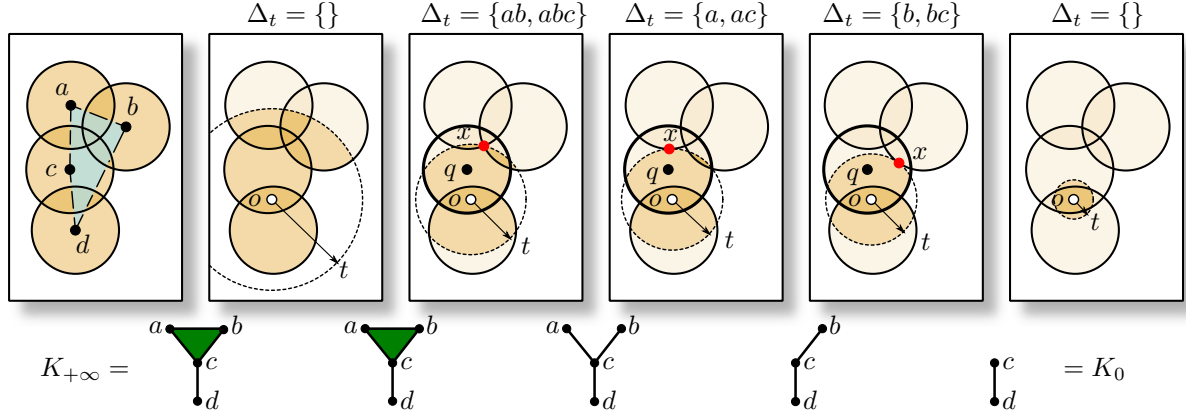


FIGURE 5.1 – Notation pour la preuve du théorème 5.1. A gauche : les quatre boules unité ouvertes contiennent l’enveloppe convexe de  $a$ ,  $b$ ,  $c$  et  $d$ . De gauche à droite : la preuve consiste à balayer l’espace avec une sphère centrée en  $o \in \text{Hull}(\{a, b, c, d\})$  dont le rayon  $t$  décroît continûment de  $+\infty$  vers 0. On déduit du balayage une séquence de collapses réduisant  $\text{Cech}(P, \alpha)$  en un sommet.

$K_0$  est l’ensemble de points  $\tau_o = \{p \in P \mid o \in B_p\} = P \cap B_o$  qui est non vide puisque  $o \in \text{Hull}(P) \subset P^{\circ 1}$ . Ainsi  $K_0 = \text{Cl}(\tau_0)$  est donc collapsable.

Montrons maintenant que lorsque  $t$  décroît continûment de  $+\infty$  vers 0, les changements dans  $K_t$  sont des collapsages élémentaires. Précisément, soit  $\Delta_t$  l’ensemble de simplexes disparaissant au temps  $t$ ,

$$\Delta_t = \left\{ \sigma \subset P \mid \sigma \neq \emptyset \text{ et } d\left(o, \bigcap_{p \in \sigma} B_p\right) = t \right\}.$$

Supposons que  $\Delta_t \neq \emptyset$  et montrons que supprimer les simplexes  $\Delta_t$  de  $K_t$  est un collapse pour tout  $t \in (0, +\infty)$ . Génériquement, on peut supposer que l’ensemble des simplexes  $\Delta_t$  a un unique élément minimal pour l’inclusion  $\sigma_{\min}$ . Par construction, l’intersection  $B(o, t) \cap \bigcap_{p \in \sigma_{\min}} B_p$  est réduite à un seul point  $x$ . Il est facile de vérifier que  $\Delta_t$  a un unique élément maximal pour l’inclusion  $\sigma_{\max} = \{p \in P \mid x \in B_p\}$ . Ainsi,  $\Delta_t$  est constitué de toutes les cofaces de  $\sigma_{\min}$  et ces cofaces sont des faces de  $\sigma_{\max}$ . Pour montrer que la suppression de  $\Delta_t$  de  $K_t$  est un collapse, il suffit d’établir que  $\sigma_{\min} \neq \sigma_{\max}$ . Avec le lemme E.1,  $x$  appartient au bord de  $B_p$  pour tout  $p \in \sigma_{\min}$ . Notons que  $x$  appartient à l’enveloppe convexe des points  $\{o\} \cup \sigma_{\min}$  puisque sinon, on pourrait projeter  $x$  sur  $\text{Hull}(\{o\} \cup \sigma_{\min})$  et obtenir un point  $y$  dans l’intersection commune  $B(o, t) \cap \bigcap_{p \in \sigma_{\min}} B_p$  tel que  $\|y - o\| < \|x - o\|$ , ce qui est impossible. Comme  $\text{Hull}(P) \subset P^{\circ 1}$ , il existe un point  $q \in P$  tel que  $\|q - x\| < 1$ . De manière équivalente,  $x$  appartient à l’intérieur de  $B_q$ . Ainsi,  $q$  appartient à  $\sigma_{\max}$  mais pas à  $\sigma_{\min}$ . On obtient alors que  $\sigma_{\min} \neq \sigma_{\max}$  comme désiré.  $\square$

**Théorème 5.2.** Soit  $P \neq \emptyset$  un ensemble fini de points dans  $\mathbb{R}^D$  et soit  $\alpha > \frac{\lambda_P}{2-\vartheta_D}$ . Le complexe Rips( $P, \alpha$ ) est collapsable.

*Preuve du théorème 5.2.* Nous procédons en deux étapes. Dans un premier temps, nous prouvons qu’il existe une séquence de collapsages élémentaires transformant Rips( $P, \alpha$ ) en Cech( $P, \alpha$ ).

En appliquant le théorème 5.1, on déduit ensuite que  $\text{Cech}(P, \alpha)$  est collapsable et que  $\text{Rips}(P, \alpha)$  l'est donc également.

Pour la première étape, appliquons le théorème 2.3 qui donne des conditions sur  $P$  sous lesquelles  $\text{Rips}(P, \alpha)$  peut être réduit à  $\text{Cech}(P, \alpha)$  grâce à une séquence de collapses élémentaires. On a juste besoin de vérifier que  $P$  satisfait les hypothèses de ce théorème. Définissons l'ensemble des centres de  $P$  à l'échelle  $t$  comme le sous-ensemble

$$\text{Centers}(P, t) = \bigcup_{\substack{\emptyset \neq \sigma \subset P \\ \text{Rad}(\sigma) \leq t}} \{\text{Center}(\sigma)\}$$

où  $\text{Center}(X)$  et  $\text{Rad}(X)$  désignent respectivement le centre et le rayon de la plus petite boule englobante de  $X \subset \mathbb{R}^D$ . L'hypothèse du théorème 2.3 est que  $d_H(\text{Centers}(P, \vartheta_D \alpha) | P) < (2 - \vartheta_D)\alpha$  et elle est satisfaite pour  $P$  puisque pour tout  $\emptyset \neq \sigma \subset P$ , nous avons la séquence d'inclusions  $\text{Center}(\sigma) \subset \text{Hull}(\sigma) \subset \text{Hull}(P) \subset P^{\odot(2-\vartheta_D)\alpha}$ .  $\square$

**Optimalité de la constante  $\frac{1}{2-\vartheta_D}$ .** Afin de montrer que la constante  $\frac{1}{2-\vartheta_D}$  est optimale, nous utilisons l'exemple 2.3.3 du chapitre 2. En particulier, lorsque la dimension ambiante vérifie  $D = 2$ , nous avons construit un ensemble de points  $P$  tel que  $\text{Rips}(P, \alpha)$  n'est pas contractile (donc pas collapsable) et tel que  $h_P(\vartheta_D \alpha)$  est supérieure mais infinitésimalement proche de  $(2 - \vartheta_D)\alpha$ . Or, dans cet exemple,  $\lambda_P = h_P(\vartheta_D \alpha)$  ainsi, pour  $D = 2$ , une constante  $\beta$  permettant à l'énoncé 5.1 d'être vérifié doit être au moins égale à  $\frac{1}{2-\vartheta_D}$ .

### 5.2.2 Préserver un *flag-complexe* durant la simplification

La séquence de réduction que l'on a obtenue dans la section précédente utilise le complexe de Čech et est assez délicate à calculer. En renforçant, la condition sur  $P$ , nous allons voir dans cette section qu'il est possible de simplifier le complexe de Rips tout en préservant sa propriété d'être un *flag-complexe* à tout moment.

**Théorème 5.3.** *Soit  $P \neq \emptyset$  un ensemble fini de points dans  $\mathbb{R}^D$  et  $\alpha > \frac{\lambda_P}{2-\sqrt{3}}$ . Il existe une séquence de collapses étendus réduisant  $\text{Rips}(P, \alpha)$  vers un sommet de telle façon qu'après chaque collapse étendu, le complexe soit toujours un *flag-complexe*.*

*Démonstration.* Comme pour la preuve du théorème 5.1, posons  $\alpha = 1$  et notons  $B_x$  la boule unité fermée centrée en  $x$ . Fixons un point  $o$  dans l'enveloppe convexe de  $P$ . Notre séquence de collapses est obtenue en balayant l'espace avec une sphère centrée en  $o$  dont le rayon  $t \geq 0$  décroît continûment de  $+\infty$  vers 0. Précisément, soit  $G_t$  le graphe dont les sommets sont les points  $a \in P$  tels que  $B(o, t) \cap B_a \neq \emptyset$  et dont les arêtes connectent toutes les paires de points  $a, b \in P$  telles que  $B(o, t) \cap B_a \cap B_b \neq \emptyset$ . Soit  $K_t = \text{Flag}(G_t)$ . Clairement,  $K_\infty = \text{Rips}(P, 1)$  et  $K_0$  est collapsable en utilisant exactement le même argument que pour la preuve du théorème 5.1. Lorsque l'on fait décroître continûment le paramètre  $t$  de  $+\infty$  vers 0, les changements dans le complexe simplicial  $K_t$  se produisent dès lors qu'un sommet ou une arête disparaît

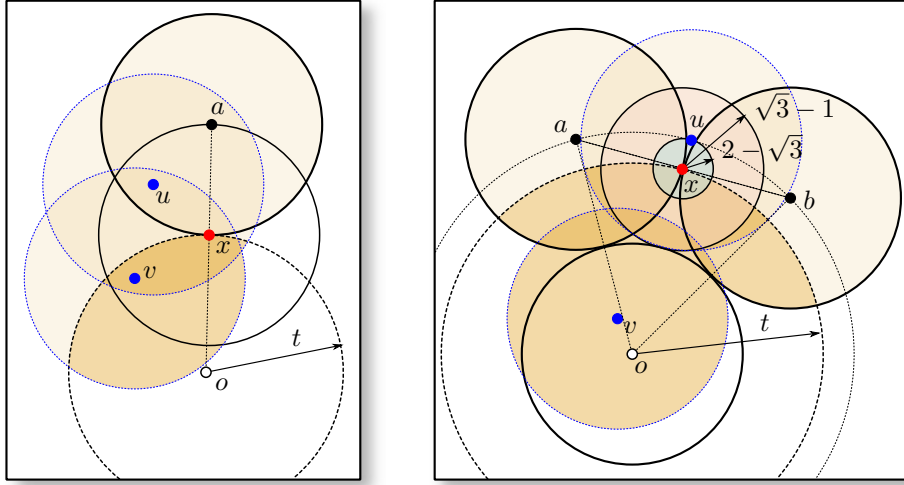


FIGURE 5.2 – Notation pour la preuve du théorème 5.3. Deux types d'événements peuvent se produire : un collapse de sommet (sur la gauche) ou un collapse d'arête (sur la droite). Le cas d'un collapse d'arête est illustré lorsque le triangle  $oab$  est équilatéral.

du graphe  $G_t$ . Génériquement, on peut supposer que ces événements ne se produisent pas simultanément.

Lorsqu'un sommet  $a$  disparaît de  $K_t$  au temps  $t$ , l'intersection  $B(o, t) \cap B_a$  se réduit à un simple sommet  $x$ ; voir la partie gauche de la figure 5.2. Dans cette situation, montrons que le *link* de  $a$  dans  $K_t$  est la fermeture du simplexe  $\tau_x = \{p \in P \setminus \{a\} \mid x \in B_p\} = P \cap B_x \setminus \{a\}$ . Tout d'abord, observons que  $\tau_x$  est non vide puisque  $x$  est situé sur le segment reliant  $o$  à  $a$  et appartient donc à l'enveloppe convexe de  $P$  qui est contenue dans  $P^{2-\sqrt{3}} \subset P^{\odot 1}$ . De plus,  $\tau_x$  est précisément l'ensemble des sommets du *link* puisqu'une arête  $au$  appartient à  $K_t$  si et seulement si  $B(o, t) \cap B_a \cap B_u \neq \emptyset$  où  $u \in P \setminus \{a\}$  ce qui peut être reformulé comme  $u \in P \cap B_x \setminus \{a\}$ . Finalement, toute paire de sommets  $u, v$  dans le *link* est connectée par une arête puisque  $u, v \in \tau_x$  implique que  $B(o, t) \cap B_u \cap B_v \supset \{x\}$ .

Lorsqu'une arête  $ab$  disparaît de  $K_t$  au temps  $t$ , il existe un point  $x$  tel que  $\{x\} = B(o, t) \cap B_a \cap B_b$ ; voir la partie droite de la figure 5.2. Notons que  $x$  appartient à l'enveloppe convexe de  $\{a, b, o\}$  et appartient donc à l'enveloppe convexe de  $P$ . Puisque  $\text{Hull}(P) \subset P^{2-\sqrt{3}}$ , il existe  $u \in P$  tel que  $\|u - x\| \leq 2 - \sqrt{3}$ . En particulier,  $x \in B_u$  et  $u$  appartient donc au *link* de  $ab$  dans  $K_t$ . Le *link* de  $ab$  est alors un cône d'apex  $u$ . En effet, considérons un point  $v \in P$  appartenant au *link* de  $ab$  dans  $K_t$ . De manière équivalente,  $B(o, t) \cap B_a \cap B_v \neq \emptyset$  et  $B(o, t) \cap B_b \cap B_v \neq \emptyset$ , le lemme 5.1 implique alors que  $B(o, t) \cap B(x, \sqrt{3} - 1) \cap B_v \neq \emptyset$ . Puisque  $B(x, \sqrt{3} - 1) \subset B_u$ , on a aussi  $B(o, t) \cap B_u \cap B_v \neq \emptyset$ , ce qui montre que  $uv$  appartient également au *link* de  $ab$  dans  $K_t$ . On a montré que le *link* de  $ab$  dans  $K_t$  est un cône. Ainsi, enlever le *star* de  $ab$  de  $K_t$  correspond à un collapse étendu.  $\square$

La preuve du théorème 5.3 repose sur le lemme géométrique suivant dont la preuve est

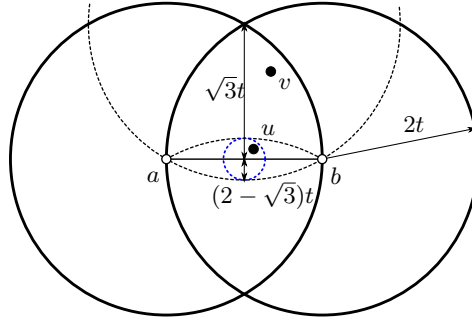


FIGURE 5.3 – Notation pour la preuve du théorème 5.4.

donnée dans l'annexe E.

**Lemme 5.1.** *Soient  $a$  et  $b$  deux points tels que  $B_a$  et  $B_b$  aient une intersection non vide. Soit  $o$  un point tel que  $d(o, B_a \cap B_b) = t > 0$ . Soit  $x$  l'(unique) point de  $B_a \cap B_b$  le plus proche de  $o$ . Toute boule unité qui possède une intersection non vide avec  $B_a \cap B(o, t)$  et  $B_b \cap B(o, t)$  a une intersection non vide avec  $B(x, \sqrt{3} - 1) \cap B(o, t)$ .*

### 5.2.3 Augmenter le paramètre d'échelle

Notre troisième stratégie pour réduire  $\text{Rips}(P, \alpha)$  consiste à augmenter le paramètre d'échelle  $\alpha$ , en ajoutant des arêtes dans le 1-squelette jusqu'à ce qu'on obtienne le graphe complet de  $P$ . Lorsque c'est le cas, le résultat est  $\text{Cl}(P) = \text{Rips}(P, +\infty)$  qui est collapsable. Le théorème suivant donne des conditions sous lesquelles cette stratégie est certaine de réussir.

**Théorème 5.4.** *Soit  $P \neq \emptyset$  un ensemble fini de points de  $\mathbb{R}^D$  et  $\alpha > \frac{\lambda_P}{2 - \sqrt{3}}$ . Lorsque  $t$  décroît continûment de  $+\infty$  vers  $\alpha$ , les seuls changements qui se produisent dans  $\text{Rips}(P, t)$  sont des collapses étendus d'arêtes.*

*Démonstration.* Rappelons que le complexe de Rips à l'échelle  $t$  peut être défini comme l'ensemble des simplexes de diamètre au plus  $2t$ ,  $\text{Rips}(P, t) = \{\sigma \subset P \mid \sigma \neq \emptyset \text{ et } \text{Diam}(\sigma) \leq 2t\}$ . Avec cette définition, il est facile de voir que l'ensemble des simplexes qui disparaissent au temps  $t$  est

$$\Delta_t = \{\sigma \subset P \mid \sigma \neq \emptyset \text{ et } \text{Diam}(\sigma) = 2t\}.$$

Supposons  $\Delta_t \neq \emptyset$  et montrons que la suppression des simplexes  $\Delta_t$  de  $\text{Rips}(P, t)$  est un collapse étendu pour tout  $t \in (\alpha, +\infty)$ . Génériquement, tous les simplexes de  $\Delta_t$  sont des cofaces d'une arête  $ab$  telle que  $\|a - b\| = 2t$ . Pour conclure, il suffit de montrer que le *link* de  $ab$  dans  $\text{Rips}(P, t)$  est un cône. En posant  $L = (B(a, 2t) \cap B(b, 2t)) \setminus \{a, b\}$ , on note que le *link* de  $ab$  dans  $\text{Rips}(P, t)$  est  $\text{Rips}(P \cap L, t)$ . Soit  $c = \frac{a+b}{2}$ . Puisque  $c \in \text{Hull}(P) \subset P^{(2-\sqrt{3})\alpha}$ , on peut trouver un point  $u \in P$  tel que  $\|c - u\| \leq (2 - \sqrt{3})\alpha$ ; voir la figure 5.3. Clairement,  $u \in L$  et donc  $u$  est un sommet du *link* de  $ab$ . De plus, pour tout point  $v \in P \cap L$ ,  $\|v - u\| \leq \|v - c\| + \|c - u\| \leq \sqrt{3}t + (2 - \sqrt{3})\alpha \leq 2t$  et ainsi  $uv$  est une arête du *link* de  $ab$ . En particulier, cela implique que le *link* de  $ab$  est un cône d'apex  $u$ .  $\square$

### 5.2.4 Algorithmes pour trouver une séquence de réduction.

En nous aidant des résultats de cette section, nous présentons quatre alternatives pour essayer de trouver une séquence de réduction d'un complexe de Rips vers un point. Le pseudo-code de chaque procédure est donné dans l'annexe. Chaque procédure teste si le complexe possède une forme dont découle la collapsibilité (par exemple s'il s'agit d'un cône) ou cherche une séquence de réduction vers un point. Si le complexe possède une telle forme ou s'il peut être réduit à un point, la procédure renvoie vrai et faux sinon. Pour les références futures, nous numérotions les stratégies de (S1) à (S4) :

- (S1) ISCONE : prend un complexe simplicial  $L$  en entrée et retourne vrai si et seulement si  $L$  est un cône.
- (S2) REDUCTIBLE\_BY\_SWEEP : prend un complexe simplicial  $L$  en entrée et essaie d'appliquer une séquence de collapsés de sommets et d'arêtes dans l'ordre induit par le balayage d'une sphère centrée en un des sommets de  $L$  comme décrit dans la preuve du théorème 5.3. Retourne vrai si et seulement si la stratégie réussit à réduire  $L$  en un point avec cette séquence.
- (S3) REDUCTIBLE\_BY\_COMPLETION : prend en entrée un *flag*-complexe  $L$  et essaie d'appliquer une séquence d'anti-collapses d'arêtes de longueur croissante jusqu'à ce que le complexe soit un cône ou que son 1-squelette soit le graphe complet comme décrit dans la preuve du théorème 5.4. Retourne vrai si et seulement si la stratégie réussit à compléter  $L$  vers un cône ou un graphe complet.
- (S4) REDUCTIBLE\_BY\_EDGE\_CONTRACTIONS : étant donné un complexe simplicial  $L$ , cette stratégie applique une séquence de contractions d'arêtes  $ab \mapsto \frac{a+b}{2}$  où  $\text{Lk}_L(ab) = \text{Lk}_L(a) \cap \text{Lk}_L(b)$  dans l'ordre de longueur croissante comme expliqué dans le chapitre 4. Retourne vrai si et seulement si le complexe  $L$  est réduit à un sommet après simplification.

Notons REDUCIBLE la fonction prenant en entrée un *flag*-complexe  $L$  et essayant de réduire celui-ci avec une stratégie (Sx) pour  $x \in [1, 4]$ . Comme chaque stratégie (Sx) renvoie vrai si elle réussit à réduire le complexe en un point sans changer le type d'homotopie, l'implication suivante est évidente :

$$\text{REDUCIBLE}(L) \implies L \text{ est contractile}$$

Il est clair qu'on ne peut pas espérer avoir la réciproque de cette implication puisque le problème de reconnaître un complexe contractile est indécidable. Néanmoins, on peut essayer de trouver une classe de complexes la plus grande possible telle que la réciproque soit vraie. En utilisant les théorèmes 5.3 et 5.4, on sait que  $\text{REDUCTIBLE\_BY\_SWEEP}(\text{Rips}(P, \alpha))$  et  $\text{REDUCTIBLE\_BY\_COMPLETION}(\text{Rips}(P, \alpha))$  sont vrais dès lors que  $P$  est non vide et que  $\alpha > \frac{h(P)}{2-\sqrt{3}}$ . Ainsi, on sait qu'on peut réduire un complexe de Rips dont les points sont suffisamment proches d'un convexe.



### 5.3 Simplifier un complexe de Rips dont les sommets échantillonnent une variété

Dans la section précédente, nous avons donné des conditions théoriques garantissant que l'on peut réduire le complexe de Rips en un point par une séquence de collapses lorsque ses sommets sont proches d'un convexe. Ainsi, si l'on échantillonne bien une 0-variété, on peut trouver une séquence de collapse du complexe de Rips vers un point. Le but de cette section est de présenter plusieurs stratégies de simplification pour un complexe de Rips dont les sommets échantillonnent une  $d$ -variété avec  $d > 0$ . On cherche à réduire la dimension du complexe de Rips afin d'obtenir un complexe dont la dimension soit proche de  $d$ . Dans le cas idéal, on voudrait pouvoir obtenir un complexe homéomorphe à la variété. Jusqu'à la fin du chapitre, on suppose que  $P$  est un nuage de points échantillonnant une  $d$ -variété  $A$  plongée dans  $\mathbb{R}^D$ . On suppose également que l'on a trouvé une valeur de  $\alpha$  telle que  $\text{Rips}(P, \alpha) \simeq A$  en s'aidant des conditions données dans le chapitre 2.

Il est fréquent de supposer une hypothèse de dispersion sur  $P$  lorsqu'on cherche à construire une approximation homéomorphe à  $A$ . Bien souvent, lorsque  $P$  est extrait de données réelles, la densité de points est élevée à certains endroits et plus faible ailleurs. Dans un premier temps, en utilisant les résultats de la section précédente, nous présentons une condition théorique permettant d'extraire un sous-ensemble  $L \subset P$  tel que  $L$  ait une densité plus uniforme tout en préservant le type d'homotopie i.e. tel que  $\text{Rips}(L, \alpha) \simeq \text{Rips}(P, \alpha)$ . Sous cette condition, on peut alors réduire  $\text{Rips}(L, \alpha)$  vers  $\text{Rips}(P, \alpha)$  en collapsant tous les sommets de  $P \setminus L$ . Dans un second temps, nous introduisons un algorithme de simplification dans le but de réduire  $\text{Rips}(P, \alpha)$  en une triangulation homéomorphe à  $A$ . Nous montrons que, sous des conditions d'échantillonnage optimales, cet algorithme reconstruit une approximation homéomorphe à  $A$  en supposant que  $A$  soit une courbe et que  $P \subset A$ . Finalement, nous montrons des résultats expérimentaux obtenus sur le jeu de données du chapitre 4.

#### 5.3.1 Extraire un sous-ensemble de points sans changer le type d'homotopie

Les méthodes construisant une approximation homéomorphe à une variété  $A$  à partir d'un nuage de points  $P$  supposent souvent que l'ensemble des points est suffisamment éparpillé [CDR05, BGO07, BG10]. Précisément, ces méthodes supposent que  $P$  est un  $(\varepsilon, \delta)$ -échantillon de  $A$  (voir la section 4.3.2 pour la définition). Il est clair que cette hypothèse n'est que rarement satisfaite pour des nuages de points issus de données réelles. Se pose alors la question de savoir quand est-ce qu'on peut extraire un sous-ensemble  $L \subset P$  clairsemé sans changer la topologie du complexe de Rips, c'est-à-dire en ayant  $\text{Rips}(P, \alpha) \simeq \text{Rips}(L, \alpha)$ . Notons que supprimer la *star* d'un sommet dans un complexe de Rips revient à calculer ce complexe sans ce sommet. Formellement, si  $p$  est un sommet de  $\text{Rips}(P, \alpha)$ ,

$$\text{Rips}(P, \alpha) \setminus \text{St}_{\text{Rips}(P, \alpha)}(p) = \text{Rips}(P \setminus \{p\}, \alpha)$$

Ainsi pour montrer que  $\text{Rips}(P, \alpha) \simeq \text{Rips}(L, \alpha)$ , il suffit de montrer que la suppression du *star* des sommets de  $P \setminus L$  est à chaque fois un *collapse généralisé*. De manière équivalente, il suffit de montrer que les *links* des sommets de  $P \setminus L$  sont à chaque fois *contractiles*. Les résultats de la section précédente nous sont alors utiles. En effet, si  $L$  est suffisamment proche de  $A$  et que  $A$  n'est pas trop courbée, les sommets des *links* des sommets que l'on cherche à enlever seront proches de convexes. Dès lors, les résultats de la section précédente nous permettent de garantir qu'on peut réduire leurs *links* en un point et que leurs suppressions sont des *collapses généralisés*. Le théorème suivant précise formellement cette intuition :

**Théorème 5.5.** *Soit  $A$  une variété telle que  $\text{reach}(A) \geq R > 0$ . Soient  $P$  et  $L$  deux ensembles finis de points non vides tels que  $\max\{d_H(P, A), d_H(L, A)\} \leq \varepsilon$  et  $L \subset P$ . Si  $\varepsilon \leq \alpha \leq \frac{R-\varepsilon}{2}$  et si  $\left(1 - \sqrt{1 - \left(\frac{2\alpha+\varepsilon}{R}\right)^2}\right) R + 2\varepsilon \leq (1 - \frac{\vartheta_D}{2})\alpha$  alors il existe une séquence de *collapses* de sommets permettant de passer de  $\text{Rips}(P, \alpha)$  à  $\text{Rips}(L, \alpha)$ .*

*Démonstration.* Nous allons montrer que  $\text{Rips}(L, \alpha)$  s'obtient à partir de  $\text{Rips}(P, \alpha)$  en *collapsant* tous les sommets de  $P \setminus L$  dans  $\text{Rips}(P, \alpha)$ . Pour cela, il suffit de montrer que pour tout sous-ensemble  $Q$  vérifiant  $L \subsetneq Q \subset P$  et pour tout  $q \in Q \setminus L$ , le *link* de  $q$  dans  $\text{Rips}(Q, \alpha)$  est *collapsable*. Prenons  $q \in Q \setminus L$  et notons  $Q_{2\alpha} = B(q, 2\alpha) \setminus \{q\} \cap Q$  les sommets du *link* de  $q$  dans  $\text{Rips}(Q, \alpha)$ . Nous avons montré dans le théorème 5.2 qu'une condition suffisante pour que le *link* de  $q$  soit *collapsable* est  $\emptyset \neq \text{Hull}(Q_{2\alpha}) \subset Q_{2\alpha}^{(2-\vartheta_D)\alpha}$ . Commençons par remarquer que  $Q_{2\alpha}$  est non vide. En effet, puisque les distances  $d_H(A, P)$  et  $d_H(L, P)$  sont toutes deux majorées par  $\varepsilon$ , la distance  $d_H(P, L)$  est majorée par  $2\varepsilon$ . Ainsi, comme  $q \in P$ , il existe un point  $l \in L$  dans la boule  $B(q, 2\varepsilon)$ . Finalement comme  $\varepsilon \leq \alpha$ ,  $l \in Q_{2\alpha}$ .

Il nous reste à montrer que  $\text{Hull}(Q_{2\alpha}) \subset Q_{2\alpha}^{(2-\vartheta_D)\alpha}$ . En appliquant le lemme 5.2, on obtient que  $\text{Hull}(Q_{2\alpha}) \subset Q_{2\alpha}^{2h_Q(2\alpha)}$ . Ainsi, il nous suffit de vérifier que  $2h_Q(2\alpha)$  est bornée par  $(2-\vartheta_D)\alpha$  pour conclure. En remarquant que  $d_H(Q, A) \leq \varepsilon$  et en utilisant le lemme de stabilité 2.4, on obtient que

$$h_Q(2\alpha) \leq h_A(2\alpha + \varepsilon) + 2\varepsilon$$

Or, en utilisant le lemme 2.4, on obtient la borne suivante sur le défaut de convexité de  $A$

$$h_A(2\alpha + \varepsilon) \leq \left(1 - \sqrt{1 - \left(\frac{2\alpha + \varepsilon}{R}\right)^2}\right) R$$

En combinant ces deux inégalités, on obtient finalement que

$$2h_Q(2\alpha) \leq (2 - \vartheta_D)\alpha$$

Ainsi, le complexe  $\text{Rips}(Q_{2\alpha}, \alpha)$  qui est le *link* de  $p$  dans  $\text{Rips}(Q, \alpha)$  est *collapsable* d'après le théorème 5.2.  $\square$

**Lemme 5.2.** *Soit  $\emptyset \neq Q \subset \mathbb{R}^D$  et  $q \in Q$  tel que  $Q_{2\alpha} = (Q \cap B(q, 2\alpha)) \setminus \{q\}$  soit non vide. Si  $h_Q(2\alpha) < \alpha$  alors  $\text{Hull}(Q_{2\alpha}) \subset Q_{2\alpha}^{2h_Q(2\alpha)}$*

*Démonstration.* Par définition de  $h_Q$ , pour tout  $\sigma \subset Q$  tel que  $\text{Rad}(\sigma) \leq 2\alpha$ , l'inclusion  $\text{Hull}(\sigma) \subset Q^{h_Q(2\alpha)}$  est vérifiée. Ainsi,  $\text{Hull}(Q_{2\alpha}) \subset Q^{h_Q(2\alpha)}$ . Prenons un point  $x \in \text{Hull}(Q_{2\alpha})$  et montrons qu'on peut trouver un point  $y \in Q_{2\alpha}$  tel que  $\|x - y\| \leq 2h_Q(2\alpha)$ . On distingue trois cas :

Cas 1 : Si  $\|x - q\| \geq h_Q(2\alpha)$  alors les sphères  $S(x, h_Q(2\alpha))$  et  $S(q, h_Q(2\alpha))$  intersectent toutes les deux le segment  $xq$ . On note  $z$  le point dans l'intersection le plus proche de  $x$ . Par construction,  $\|x - z\| \leq h_Q(2\alpha)$  et  $\|z - q\| \geq h_Q(2\alpha)$ . Comme  $h_Q(2\alpha) \leq \alpha$ , la boule  $B(z, h_Q(2\alpha))$  est contenue dans la boule  $B(q, \alpha)$  et comme  $z \in \text{Hull}(Q_{2\alpha}) \subset Q^{h_Q(2\alpha)}$ , la boule  $B(z, h_Q(2\alpha))$  contient un point  $y$  de  $Q$  dans son intérieur. Par construction  $y \neq q$  et donc  $y \in Q_{2\alpha}$  avec  $\|x - y\| \leq \|x - z\| + \|z - y\| \leq 2h_Q(2\alpha)$ .

Cas 2 : Si  $\|x - q\| < h_Q(2\alpha)$  et si  $q$  est le seul point de  $Q$  contenu dans  $B(q, h_Q(2\alpha))$  alors on considère l'hyperplan passant par  $x$  et orthogonal à la droite connectant  $x$  au centre  $q$ . Cet hyperplan délimite deux demi-espaces et on note  $H$  celui dont l'intérieur ne contient pas  $q$ . Si  $x = q$ , alors on prend pour  $H$  un des demi-espaces fermés dont le bord passe par  $x$ . Par construction,  $H$  contient  $x \in \text{Hull}(Q_{2\alpha})$  et intersecte donc  $Q_{2\alpha}$ . Soit  $p \in H \cap Q_{2\alpha}$  et soit  $z$  l'intersection du segment  $xp$  avec la sphère  $S(x, h_Q(2\alpha))$ . Comme  $h_Q(2\alpha) \leq \alpha$ , la boule  $B(z, h_Q(2\alpha))$  est contenue dans  $B(q, \alpha)$  et on conclut comme dans le cas précédent.

Cas 3 : Si  $\|x - q\| < h_Q(2\alpha)$  et s'il existe un point  $y \in Q$  distinct de  $q$  dans l'intérieur de la boule  $B(q, h_Q(2\alpha))$  alors on conclut immédiatement.

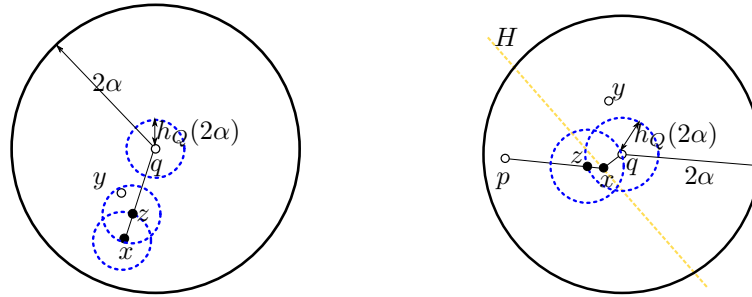


FIGURE 5.4 – Notation pour la preuve du lemme 5.2 dans le cas 1 (gauche) et dans le cas 2 (droite).

□

Supposons que  $L$  et  $P$  soient deux nuages de points échantillonnant  $A$  où  $L$  est un ensemble de points clairsemés (par exemple un  $(\varepsilon, \delta)$ -échantillon). Supposons également que  $L$  et  $P$  vérifient les hypothèses du théorème 5.5. Remarquons que, sous ces hypothèses, on sait que  $\text{Rips}(P, \alpha) \simeq \text{Rips}(L, \alpha) \simeq A$ . Les hypothèses du théorème 5.5 permettent en plus d'assurer qu'on peut obtenir  $\text{Rips}(L, \alpha)$  à partir de  $\text{Rips}(P, \alpha)$  en collapsant itérativement tous les sommets de  $P \setminus L$  dans  $\text{Rips}(P, \alpha)$ .

### 5.3.2 Algorithme de simplification.

Le complexe de Rips est simplifié en deux étapes : la première collapse itérativement les sommets et la seconde collapse itérativement les arêtes ; voir l'algorithme 4 pour le pseudo-code. Durant la simplification, le complexe reste un *flag*-complexe puisque cette propriété n'est pas altérée en collapsant un sommet ou une arête. Pour  $k \in \{0, 1\}$ , l'étape  $k$  procède comme suit. Initialement, un ensemble de  $k$ -simplexes est stocké dans une file de priorité  $Q$ . Chaque  $k$ -simplexe reçoit comme priorité son diamètre. Ainsi, tous les sommets reçoivent la même priorité alors que la plus grande arête reçoit la plus grande priorité. Durant l'étape  $k$ , on extrait itérativement le  $k$ -simplexe  $\sigma$  de priorité maximale et on l'enlève du complexe courant  $K$  avec toutes ses cofaces dès lors que  $\text{REDUCIBLE}(\text{Lk}_K(\sigma))$  renvoie vrai ; voir algorithme 5 pour le pseudo-code. Idéalement, on voudrait que la fonction  $\text{REDUCIBLE}(\text{Lk}_K(\sigma))$  retourne vrai si et seulement si l'opération qui enlève  $\sigma$  et ses cofaces est un collapse généralisé. Autrement dit, on voudrait qu'idéalement la fonction  $\text{REDUCIBLE}$  prenne comme entrée un complexe simplicial  $L$  et retourne vrai dès qu'il existe une séquence d'opérations élémentaires préservant le type d'homotopie (collapses, anti-collapses et contractions d'arêtes) allant de  $L$  vers un point et qu'elle retourne faux sinon. Malheureusement, comme mentionné précédemment, reconnaître un complexe contractile est indécidable. A la place, nous utilisons une des quatre stratégies de la section 5.2.4 pour trouver une telle séquence et implémenter la fonction  $\text{REDUCIBLE}$ .

---

**Algorithme 4**  $\text{SIMPLIFIE}(\text{Complexe simplicial } K)$

---

$L =$  une approximation clairsemée de  $P = \text{Vert}(K)$   
 $\text{SIMPLIFIE}(K, P \setminus L)$  {Collapse les sommets de  $P \setminus L$  dans  $K$ }  
 $\text{SIMPLIFIE}(K, \text{Edges}(K))$  {Collapse les arêtes de  $K$ }

---



---

**Algorithme 5**  $\text{SIMPLIFIE}(\text{Complexe simplicial } K, \text{File de priorité } Q)$

---

**while**  $Q \neq \emptyset$  **do**  
    Extraire le simplexe  $\sigma$  de  $Q$  de priorité maximale  
    **if**  $\text{REDUCIBLE}(\text{Lk}_K(\sigma))$  **then**  
         $K = K \setminus \text{St}_K(\sigma)$

---

**Approximation clairsemée.** Plusieurs solutions existent pour calculer une approximation clairsemée  $L$  d'un nuage de points  $P$ . Une possibilité est d'utiliser un algorithme max-min comme pour le choix des témoins dans les complexes de témoins [DSC04]. Cet algorithme consiste à mettre initialement un sommet au hasard dans  $L$ . A chaque itération, on rajoute ensuite à  $L$  le sommet de  $P$  le plus loin de  $L$  et on s'arrête lorsque  $L$  est suffisamment proche de  $P$ . Cet algorithme renvoie bien un ensemble de points clairsemés [Cla06a]. Une autre possibilité est d'utiliser les contractions d'arêtes. On contracte des arêtes dans le complexe  $\text{Rips}(P, \alpha)$  jusqu'à obtenir un complexe dont les sommets sont suffisamment clairsemés (voir la section pour une intuition sur le fait que les contractions d'arêtes donnent des ensembles de points de plus en plus clairsemés 4.3.2). Dans la suite, nous utilisons la stratégie consistant à effectuer des contractions d'arêtes pour obtenir une approximation clairsemée d'un ensemble de points

$P^3$ .

### 5.3.3 Preuve de correction de l'algorithme de reconstruction dans un cas simple

Nous montrons ici que l'algorithme de simplification renvoie un complexe homéomorphe à  $A$  lorsque  $A$  est une courbe à *reach* positif et que  $P \subset A$  est un échantillon non bruité. Pour cela, nous avons besoin de quelques définitions. Etant donné  $\alpha$ , on appelle *graphe de Gabriel* le graphe  $\text{Gabriel}(P, \alpha)$ . Les sommets de ce graphe sont les points de  $P$  et ses arêtes connectent les points  $a$  et  $b$  de longueur  $\|a - b\| \leq 2\alpha$  et tels que la plus petite boule englobante de  $a$  et  $b$  ne contienne aucun autre point de  $P$  dans son intérieur. Etant donné  $\varepsilon > 0$ , on dit qu'un ensemble de points  $P \subset A$  est un  $\varepsilon$ -échantillon de  $A$  si pour tout  $x \in A$ , il existe au moins deux points de  $P$  dans la boule ouverte centrée en  $x$  de rayon  $\varepsilon^4$ . Deux points  $a$  et  $b$  sont dits  *$P$ -contigus sur  $A$*  s'il existe un intervalle  $\Gamma \subset A$  d'extrémités  $a$  et  $b$  et tel que  $\Gamma \cap P = \{a, b\}$ . Enfin, étant donnés deux points  $a$  et  $b$ , on leur associe l'ensemble des boules

$$\mathcal{B}_{ab} = \{B(z, r) \mid \|z - a\| = \|z - b\| = r \leq \|a - b\|\}$$

On note également  $E_{ab} = \bigcap \mathcal{B}_{ab}$  l'intersection de ces boules (voir la figure 5.5). Rappelons le lemme suivant dont une preuve peut être trouvée dans [AB99].

**Lemme 5.3.** *Soit  $B$  une boule fermée de rayon strictement inférieur à  $\text{reach}(A)$ . L'intersection  $A \cap B$  est soit vide soit un intervalle (i.e. une 1-boule topologique fermée).*

Ainsi lorsque  $\|a - b\| < \text{reach}(A)$ , toutes les boules de  $\mathcal{B}_{ab}$  intersectent  $A$  selon un intervalle qui est contenu dans  $E_{ab}$ . On en déduit les deux lemmes suivants :

**Lemme 5.4.** *Soient  $a$  et  $b$  deux points vérifiant  $\|a - b\| < \text{reach}(A)$ . Alors, pour tout  $B \in \mathcal{B}_{ab}$ ,  $A \cap B = A \cap E_{ab}$ .*

**Lemme 5.5.** *Soit  $P \subset A$ . Pour tout  $a, b \in P$  vérifiant  $\|a - b\| \leq 2\alpha < \text{reach}(A)$ , les deux implications suivantes sont vérifiées :*

$$\begin{aligned} ab \notin \text{Gabriel}(P, \alpha) &\implies E_{ab} \cap P \neq \{a, b\} \\ ab \in \text{Gabriel}(P, \alpha) &\implies \bigcup \mathcal{B}_{ab} \cap P = \{a, b\} \end{aligned}$$

On peut alors montrer la proposition suivante qui donne des conditions sous lesquelles le graphe  $\text{Gabriel}(P, \alpha)$  est homéomorphe à  $A$ .

**Proposition 5.1.** *Supposons que  $\text{reach}(A) > 0$  et soit  $P$  un  $\varepsilon$ -échantillon de  $A$ . Pour  $\varepsilon < \alpha < \text{reach}(A)$ , les deux graphes de Gabriel  $\text{Gabriel}(P, \varepsilon)$  et  $\text{Gabriel}(P, \alpha)$  sont égaux. De plus ces deux graphes coïncident avec le graphe constitué des sommets de  $P$  et dont les arêtes relient les points qui sont  $P$ -contigus sur  $A$ . Enfin,  $\text{Gabriel}(P, \alpha) \approx A$ .*

3. On pourra remarquer l'analogie entre le choix des points d'une approximation clairsemée et le choix des *landmarks* dans les complexes des témoins [CS03].

4. Notons que la définition d' $\varepsilon$ -échantillon diffère légèrement de la condition  $(\varepsilon, \delta)$ -échantillon.

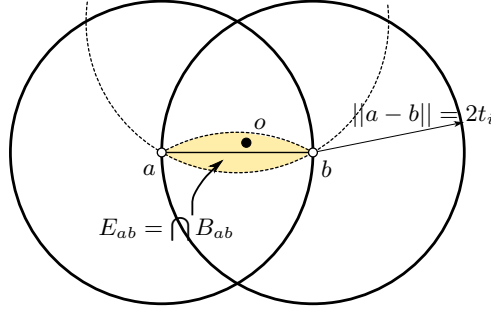


FIGURE 5.5 – Notation pour la preuve de la proposition 5.2

*Preuve de la proposition 5.1.* Soient  $a$  et  $b$  deux points  $P$ -contigus sur  $A$ . Par définition, il existe un intervalle  $\Gamma \subset A$  connectant  $a$  à  $b$  et tel que  $\Gamma \cap P = \{a, b\}$ . Soit  $x$  un point sur  $\Gamma$  équidistant de  $a$  et  $b$  et montrons que  $\|x - a\| = \|x - b\| < \varepsilon$ . Dans le cas contraire,  $B(x, \varepsilon)$  intersecterait  $A$  sur un sous-intervalle de  $\Gamma$  et on aurait  $B(x, \varepsilon) \cap P = \emptyset$  ce qui est impossible. Ainsi,  $\|x - a\| = \|x - b\| < \varepsilon$  et on en déduit que  $\frac{\|a-b\|}{2} < \varepsilon < \text{reach}(A)$ . Soit  $B_{ab}$  la plus petite boule contenant  $a$  et  $b$ . L'intersection  $A \cap B_{ab}$  est un intervalle qui intersecte  $P$  sur ses deux extrémités  $a$  et  $b$ . L'arête  $ab$  appartient donc à  $\text{Gabriel}(P, \varepsilon)$ . Réciproquement, supposons que  $ab$  soit une arête de  $\text{Gabriel}(P, \alpha)$ . Cela implique que  $\frac{\|a-b\|}{2} < \alpha < \text{reach}(A)$  et que  $A \cap B_{ab}$  est un intervalle qui intersecte  $P$  sur ses deux extrémités  $a$  et  $b$ . Ainsi  $a$  et  $b$  sont  $P$ -contigus sur  $A$ .  $\square$

Notons  $K(t) = \text{Rips}(P, t) \cup \text{Gabriel}(P, \alpha)$ . Clairement,  $K(0) = \text{Gabriel}(P, \alpha)$  et  $K(\alpha) = \text{Rips}(P, \alpha)$ . En augmentant continûment le paramètre  $t$  de 0 vers  $\alpha$ , on obtient une famille entrelacée de  $m + 1$  complexes simpliciaux :

$$\text{Gabriel}(P, \alpha) = K(0) = K_0 \subset K_1 \subset \dots \subset K_m = K(\alpha) = \text{Rips}(P, \alpha)$$

La proposition suivante montre que le passage entre deux complexes simpliciaux successifs de cette famille se fait par un collapse étendu.

**Proposition 5.2.** *Si  $2\alpha < \text{reach}(A)$  alors pour tout  $0 < i \leq m$ , l'opération réduisant  $K_i$  vers  $K_{i-1}$  est un collapse étendu.*

*Démonstration.* Pour  $0 < i < m$ , notons  $t_i$  la plus petite valeur de  $t$  telle que  $K_i = K(t)$ . Comme  $P$  est en position générale, tous les simplexes de  $K_i \setminus K_{i-1}$  sont des cofaces d'une arête  $ab$  vérifiant  $\|a - b\| = 2t_i$ . Nous allons montrer que le *link* de  $ab$  dans  $K_i$  est un cône, ce qui assurera que l'opération réduisant  $K_i$  en  $K_{i-1}$  est un collapse étendu.

Commençons par remarquer qu'aucune des cofaces de  $ab$  dans  $K_i$  n'appartient à  $\text{Gabriel}(P, \alpha)$ . Ainsi, toutes les cofaces de  $ab$  sont contenues dans  $\text{Rips}(P, t_i)$ . Les sommets de  $\text{Lk}_{K_i}(ab)$  sont donc précisément les points de  $P$  appartenant à  $L_{ab} = (B(a, 2t_i) \cap B(b, 2t_i)) \setminus \{a, b\}$  (voir la figure 5.5). Puisque  $ab \notin \text{Gabriel}(P, \alpha)$  et  $\|a - b\| = 2t_i \leq 2\alpha < R$ , le lemme 5.5 implique que  $E_{ab} \cap P \neq \{a, b\}$  et qu'il existe un point  $o \in P$  tel que  $o \in E_{ab} \setminus \{a, b\} \subset L_{ab}$ .

Ainsi, le *link* de  $ab$  dans  $K_i$  est non vide et contient au moins  $o$ . De la même manière que pour la preuve du théorème 5.4, on montre que ce complexe est un cône d'apex  $o$ .  $\square$

Nous pouvons maintenant donner des conditions assurant que l'algorithme de simplification 5 retourne un complexe homéomorphe à  $A$  lorsque son entrée est constituée des arêtes d'un complexe  $\text{Rips}(P, \alpha)$ . On suppose pour cela que la fonction REDUCIBLE dans l'algorithme 5 est vraie dès lors que son entrée est un cône comme c'est le cas pour (S1), (S2) et (S3)<sup>5</sup>.

**Théorème 5.6.** *Soit  $A$  une 1-variété à reach positif. Soit  $P$  un  $\varepsilon$ -échantillon de  $A$  et soit  $\alpha$  tel que  $\varepsilon < \alpha < \text{reach}(A)/2$ . Le complexe retourné par  $\text{SIMPLIFIE}(K, \text{Edges}(K))$  est homéomorphe à  $A$  si  $K = \text{Rips}(P, \alpha)$ .*

*Démonstration.* En initialisant l'algorithme  $\text{SIMPLIFIE}(K, \text{Edges}(K))$  avec  $K = \text{Rips}(P, \alpha)$ , ce dernier suit la même séquence de *collapses* que la séquence amenant  $K_m$  à  $K_0 = \text{Gabriel}(P, \alpha)$ . En effet, lorsque l'on extrait l'arête  $ab$  de longueur maximale de  $Q$ , deux cas peuvent se produire. Si  $ab \in \text{Gabriel}(P, \alpha)$  alors, d'après le lemme 5.5, le *link* de cette arête est vide et son *star* n'est donc pas retiré. Si  $ab \notin \text{Gabriel}(P, \alpha)$ , alors la proposition précédente montre que le *link* de  $ab$  est un cône, dès lors, la fonction REDUCIBLE peut réduire le *link* de  $ab$  en un point et l'arête  $ab$  est *collapsée*. Le complexe obtenu après simplification est donc  $\text{Gabriel}(P, \alpha)$  qui est homéomorphe à  $A$  d'après la proposition 5.1.  $\square$

## 5.4 Résultats expérimentaux

Dans cette section, nous appliquons notre algorithme de reconstruction sur les jeux de données introduits dans la section 4.4.1. Nous commençons par comparer l'efficacité des différentes stratégies que nous avons introduites afin de réduire un complexe en un point. Nous montrons ensuite des résultats expérimentaux obtenus avec notre algorithme de simplification sur les complexes de Rips construits à partir de notre jeu de données.

### 5.4.1 Comparaison des différentes stratégies pour REDUCTIBLE

Afin de comparer l'efficacité des quatre stratégies que nous avons introduites pour REDUCTIBLE, nous utilisons notre algorithme de simplification (algorithme 5), en utilisant pour REDUCTIBLE la fonction qui retourne vrai si et seulement si une des quatre stratégies retourne vrai. On dit qu'une stratégie (Sx) *échoue* à l'étape  $i$  si la stratégie retourne faux alors que REDUCTIBLE renvoie vrai. Autrement dit, on dit qu'une stratégie échoue lorsqu'elle ne parvient pas à réduire un complexe en un point alors qu'une autre stratégie y parvient. On note  $s_{\text{sommet}}^x(i)$  le nombre d'échecs de la stratégie (Sx) lors des  $i$  premières itérations de la première étape de l'algorithme de simplification (correspondant aux *collapses* de sommets). De

5. Si une stratégie (S) implémente REDUCIBLE et ne vérifie pas cette condition, on peut obtenir une stratégie (S') qui la vérifie facilement en renvoyant vrai si la stratégie (S) ou si la stratégie (S1) renvoie vrai.

$P$	Dimension intrinsèque $d$	Dimension de $K_{\text{out}}$	Homéomorphe	Temps de calcul
Cat	1	1	OUI	2 s
C2	2	2	OUI	1 min
S2	2	2	OUI	9 min
C3	3	4	NON	205 min
S3	3	4	NON	171 min
C4	4	5	NON	144 min
Tore	2	2	OUI	10 s
Pillow	2	2	OUI	1 min
Ramses	2	2	OUI	150 min
S03	3	3	NON	7 min

TABLE 5.1 – Pour chaque ensemble  $P$ , on indique la dimension du complexe  $K_{\text{out}}$ , le fait qu’il soit homéomorphe ou non à la variété  $A$  ainsi que le temps d’exécution. Tous les complexes de Rips sont construits avec les mêmes paramètres  $\alpha$  que dans le chapitre 3 (voir la table 4.1 pour des statistiques sur ces complexes) hormis le complexe construit avec l’ensemble de points S3. En effet, pour ce dernier, la valeur du paramètre de Rips a été abaissée en raison du temps de calcul trop élevé pour la simplification. Les calculs sont faits grâce à un processeur AMD cadencé à 3.4 GHz disposant de 8 Go de mémoire vive.

la même manière, on note  $s_{\text{arête}}^x(i)$  le nombre d’échecs de la stratégie  $(Sx)$  lors des  $i$  premières itérations de l’algorithme de simplification durant la phase de collapsés d’arêtes.

Sur la figure 5.6, on trace  $s_{\text{sommet}}^x$  pour  $x \in [1, 4]$  (c’est-à-dire pour toutes les stratégies), pour un sous-ensemble représentatif de notre jeu de données constitué de **Cat**, **S2**, **Ramses** et **S03**. Lorsque l’on collapse des arêtes, toutes les stratégies donnent la même réponse pour les deux premiers jeux de données, ainsi on trace seulement  $s_{\text{arête}}^x$  pour **Ramses** et **S03**, voir la figure 5.6. On observe que (S4) est la stratégie la plus efficace dans cette configuration. Elle trouve une séquence de réduction dès lors qu’une autre stratégie en trouve une lorsqu’on simplifie les quatre ensembles de points. Cependant, contrairement aux stratégies (S2) et (S3), nous n’avons pas encore pu prouver que la stratégie (S4) puisse réduire un complexe de Rips approximant un convexe. Pour la suite, nous utilisons la stratégie (S4) pour implémenter la fonction REDUCTIBLE.

#### 5.4.2 Résultats des simplifications sur le jeu de données

Nous utilisons maintenant notre algorithme de simplification sur les complexes de Rips construits sur nos jeux de données. Dans la table 5.1, on décrit le complexe  $K_{\text{out}}$  obtenu après simplification. Lorsque  $A$  est une courbe ou une surface, le résultat de la simplification  $K_{\text{out}}$  est un *flag-complexe* homéomorphe à la variété échantillonnée  $A$ . Dans les cas où  $A$  est une  $d$ -variété et  $d > 3$ , le complexe obtenu a une dimension proche de  $A$ .



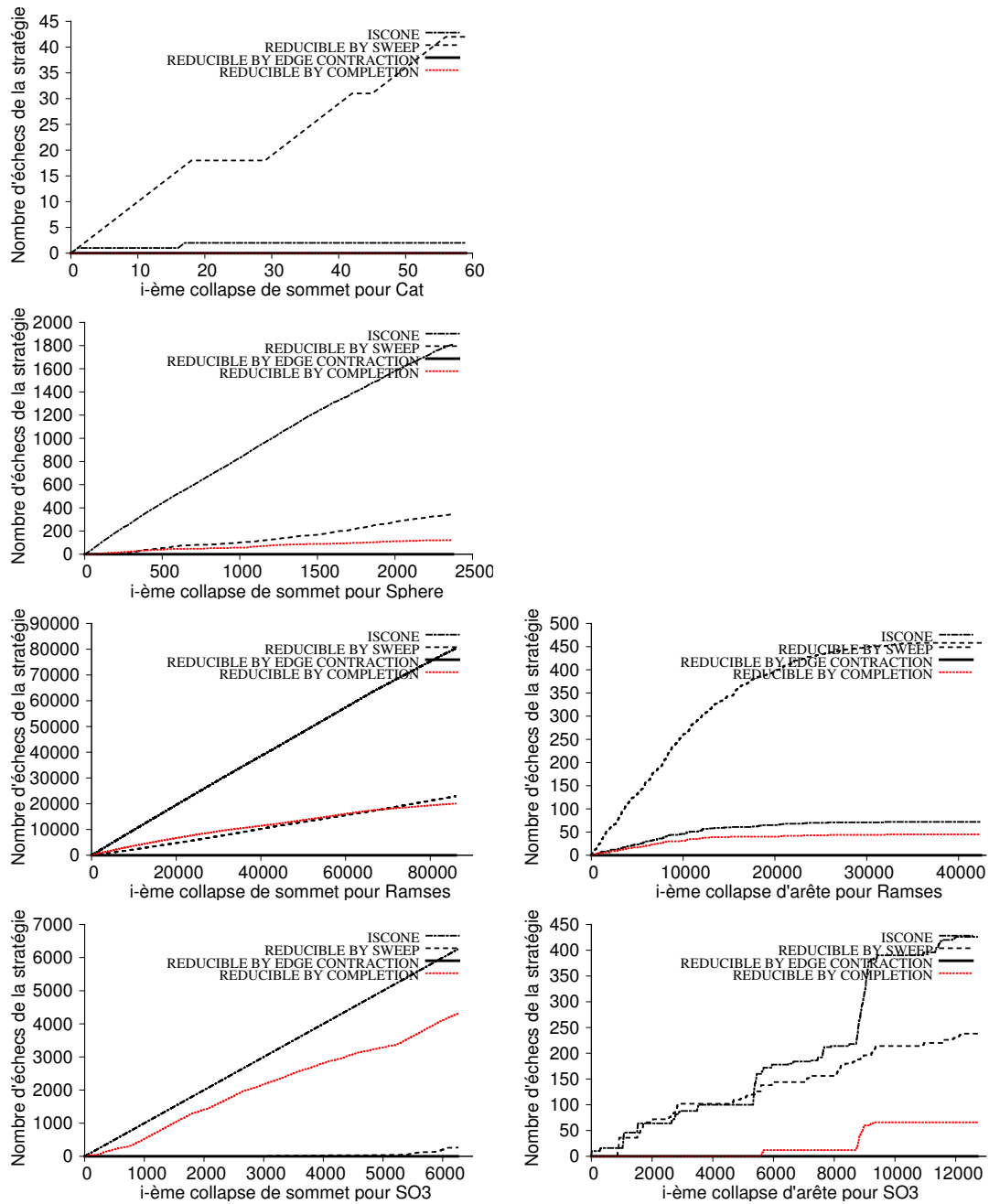


FIGURE 5.6 – Nombre d'échecs pour chaque stratégie lorsque l'on collapse les sommets (gauche) et lorsque l'on collapse des arêtes (droite) pour les ensembles de points (de haut en bas) Cat, S2, Ramses, SO3.

# Conclusion et perspectives

L'objectif de cette thèse était de proposer des algorithmes capables de construire une approximation d'une variété connue seulement à partir d'un nuage de points de grandes dimensions. Une attention particulière a été consacrée afin de proposer des algorithmes efficaces et de fournir des conditions d'échantillonnage garantissant que les approximations reconstruites aient la même topologie que la variété échantillonnée. Le fil rouge de notre approche aura été l'utilisation des *flag-complexes* et tout particulièrement des complexes de Rips. Dans un premier temps, nous avons donné des conditions d'échantillonnage sous lesquelles le complexe de Rips capture le type d'homotopie d'une variété. Dans un second temps, nous avons introduit une structure de données pour complexes simpliciaux, particulièrement compacte lorsque le complexe encodé est proche d'un *flag-complexe*. Nous avons ensuite étudié deux opérations de simplification de complexes simpliciaux : la contraction d'arête et le collapse simplicial. Nous avons à chaque fois donné une condition locale permettant de s'assurer que ces opérations préservent le type d'homotopie. Expérimentalement, les contractions d'arêtes se sont révélées très efficaces pour réduire le nombre de sommets d'un complexe de Rips approchant une variété. Une fois le complexe de Rips simplifié, outre l'avantage d'avoir une représentation compacte ayant la même topologie, on peut également calculer efficacement des invariants topologiques comme ses groupes d'homologie. Quant à l'opération de collapse, elle s'est révélée efficace pour réduire la dimension d'un complexe de Rips voire pour extraire un complexe homéomorphe à celui-ci. Remarquons le rôle salutaire de la géométrie dans nos résultats. Beaucoup de problèmes abordés dans cette thèse sont des problèmes classiques de topologie algorithmique et ont des complexités rédhibitoires pour des complexes simpliciaux abstraits. Par exemple, reconnaître un complexe contractile est un problème indécidable. Le calcul des groupes d'homologie d'un *flag-complexe* à partir de son 1-squelette est quant-à-lui un problème NP-dur. En nous aidant de la géométrie des points d'un complexe, nous avons pu traiter ces problèmes sous certaines hypothèses. Lorsque les sommets d'un complexe de Rips sont proches d'un convexe, nous avons en particulier montré que l'on peut trouver efficacement une séquence de réduction du complexe vers un point. Concernant le calcul des groupes d'homologie, en nous aidant de la métrique des points, nous avons pu expérimentalement réduire drastiquement le nombre de sommets du complexe de Rips ; le calcul des groupes d'homologie pouvait alors être fait de manière efficace. L'analyse théorique de la simplification de complexes de Rips par les opérations de contraction d'arête et de collapse soulève des questions difficiles mais ouvre également des perspectives intéressantes en topologie algorithmique et en analyse de données. Une meilleure compréhension de ces deux opérations pourrait permettre de calculer les groupes d'homologies d'un complexe de Rips approchant une variété ou encore d'en extraire un complexe homéomorphe à la variété en temps polynomial en la taille de son 1-squelette.

Cette thèse soulève également des pistes de recherche sur la représentation de données de grandes dimensions ou sur la reconstruction d'approximation avec garantie topologique. Au niveau de la représentation, la structure de données squelette/bloqueurs s'est révélée particulièrement appropriée car compacte pour représenter les complexes simpliciaux obtenus après

contractions d'arêtes en partant de complexes de Rips. Reste à savoir si d'autres types de complexes simpliciaux utilisés en analyse de données peuvent être également encodés efficacement dans cette structure de données. Un objectif serait de trouver des conditions géométriques garantissant que la représentation squelette/bloqueurs de complexes classiques tels que les complexes de Čech, les complexes de témoins ou encore les complexes de Delaunay possède peu de bloqueurs. Une autre problématique serait de proposer un analogue au complexe de Rips dont on puisse garantir la topologie pour des conditions d'échantillonnage plus souples et plus réalistes. Nous avons vu que le complexe de Rips possède la topologie d'un cercle lorsque ses points sont des images prises autour d'un objet. Néanmoins, ce cas est très spécifique puisque les points sont distribués régulièrement (en particulier, il n'y a pas de points aberrants) autour d'une variété sans bords. Il serait intéressant de proposer un *flag*-complexe analogue au complexe de Rips dont on puisse garantir la topologie même lorsque l'ensemble de points échantillonne une variété avec bords de manière non uniforme voire même avec des points aberrants. En présence de points aberrants, une piste serait d'adapter les preuves du chapitre 2 dans le cadre d'une distance à la mesure [GMM11], robuste aux points aberrants, ainsi qu'aux complexes de Rips pondérés associés. Quant au problème de reconstruire une forme échantillonnée de manière non-uniforme, une idée naturelle serait d'adapter la stratégie proposée par Chazal et Lieutier qui consiste à considérer le nerf d'une famille de boules dont le rayon varie localement [CL08].

# Bibliographie

- [AAD07] Nina Amenta, Dominique Attali, and Olivier Devillers. Complexity of delaunay triangulation for points on lower-dimensional polyhedra. In Nikhil Bansal, Kirk Pruhs, and Clifford Stein, editors, *SODA*, pages 1106–1113. SIAM, 2007. (Cit  en page 3.)
- [AB99] Nina Amenta and Marshall W. Bern. Surface reconstruction by voronoi filtering. *Discrete & Computational Geometry*, 22(4) :481–504, 1999. (Cit  en pages 2, 14 et 82.)
- [AB11] Karim Adiprasito and Bruno Benedetti. Metric geometry and collapsibility. Technical Report arXiv :1107.5789, Jul 2011. Comments : 27 pages, 5 figures. (Cit  en page 70.)
- [AL10] Dominique Attali and Andr  Lieutier. Reconstructing shapes with guarantees by unions of convex sets. In *Proceedings of the 2010 annual symposium on Computational geometry*, SoCG '10, pages 344–353, New York, NY, USA, 2010. ACM. (Cit  en pages 14, 27 et 29.)
- [ALS11a] Dominique Attali, Andr  Lieutier, and David Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. In *Proceedings of the 27th annual ACM symposium on Computational geometry*, SoCG '11, pages 501–509, 2011. (Cit  en pages 37 et 52.)
- [ALS11b] Dominique Attali, Andr  Lieutier, and David Salinas. Vietoris-rips complexes also provide topologically correct reconstructions of sampled shapes. In *Proceedings of the 27th annual ACM symposium on Computational geometry*, pages 491–500, 2011. (Cit  en page 14.)
- [ALS12a] Dominique Attali, Andr  Lieutier, and David Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. *Int. J. Comput. Geometry Appl.*, 22(4) :279–304, 2012. (Cit  en pages 13, 37 et 52.)
- [ALS12b] Dominique Attali, Andr  Lieutier, and David Salinas. Vietoris-Rips Complexes also Provide Topologically Correct Reconstructions of Sampled Shapes. *Computational Geometry*, 46(4, special issue) :448–465, 2012. (Cit  en page 14.)
- [ALS13] Dominique Attali, Andr  Lieutier, and David Salinas. Collapsing rips complexes. *29th European Workshop on Computational Geometry (EuroCG)*, 2013. (Cit  en page 70.)
- [AS12] Michal Adamaszek and Juraj Stacho. Algorithmic complexity of finding cross-cycles in flag complexes. In *Proceedings of the 2012 symposium on Computational Geometry*, SoCG '12, pages 51–60, New York, NY, USA, 2012. ACM. (Cit  en page 65.)
- [Att98] Dominique Attali. r-regular shape reconstruction from unorganized points. *Comput. Geom.*, 10(4) :239–247, 1998. (Cit  en page 14.)

- [Bar10] Jonathan Ariel Barmak. On quillen's theorem a for posets. 2010. (Cit  en page 54.)
- [Bau72] Bruce G. Baumgart. Winged edge polyhedron representation. Technical report, Stanford, CA, USA, 1972. (Cit  en page 35.)
- [BC02] Jean-Daniel Boissonnat and Fr d ric Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Comput. Geom.*, 22(1-3) :185–203, 2002. (Cit  en page 14.)
- [BE89] D.W. Barnette and Allan L. Edelson. All 2-manifolds have finitely many minimal triangulations. *Israel Journal of Mathematics*, 67 :123–128, 1989. (Cit  en page 65.)
- [BG10] Jean-Daniel Boissonnat and Arijit Ghosh. Manifold reconstruction using tangential delaunay complexes. In Jack Snoeyink, Mark de Berg, Joseph S. B. Mitchell, G nter Rote, and Monique Teillaud, editors, *Symposium on Computational Geometry*, pages 324–333. ACM, 2010. (Cit  en pages 2, 3, 70 et 78.)
- [BGO07] Jean-Daniel Boissonnat, Leonidas J. Guibas, and Steve Y. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. In *Proceedings of the twenty-third annual symposium on Computational geometry*, SCG '07, pages 194–203, New York, NY, USA, 2007. ACM. (Cit  en pages 3, 70 et 78.)
- [Bjo95] A. Bjorner. Handbook of combinatorics (vol. 2). chapter Topological methods, pages 1819–1872. MIT Press, Cambridge, MA, USA, 1995. (Cit  en pages 10 et 53.)
- [BM12a] Jonathan Ariel Barmak and Elias Gabriel Minian. Strong homotopy types, nerves and collapses. *Discrete & Computational Geometry*, 47(2) :301–328, 2012. (Cit  en pages 11 et 71.)
- [BM12b] Jean-Daniel Boissonnat and Cl ment Maria. The simplex tree : an efficient data structure for general simplicial complexes. In *Proceedings of the 20th Annual European conference on Algorithms*, ESA'12, pages 731–742, Berlin, Heidelberg, 2012. Springer-Verlag. (Cit  en pages 36 et 48.)
- [BN01] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems 14*, pages 585–591. MIT Press, 2001. (Cit  en page 1.)
- [Bri89] E. Brisson. Representing geometric structures in d dimensions : topology and order. In *Proceedings of the fifth annual symposium on Computational geometry*, SCG '89, pages 218–227, New York, NY, USA, 1989. ACM. (Cit  en page 36.)
- [Car08] Gunnar Carlsson. Topology and data. Technical report, 2008. (Cit  en page 1.)
- [CCS03] Erik Carlsson, Gunnar Carlsson, and Vin De Silva. An algebraic topological method for feature identification. *Intl. J. Computational Geometry and Applications*, 16 :291–314, 2003. (Cit  en page 3.)
- [CCSL09] Fr d ric Chazal, David Cohen-Steiner, and Andr  Lieutier. A sampling theory for compact sets in euclidean space. *Discrete & Computational Geometry*, 41(3) :461–479, 2009. (Cit  en pages 2, 3, 14, 28, 29, 30, 31 et 98.)

- [CDR05] Siu-Wing Cheng, Tamal K. Dey, and Edgar A. Ramos. Manifold reconstruction from point samples. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '05, pages 1018–1027, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics. (Cité en pages 2, 3 et 78.)
- [CdSEG10] Erin W. Chambers, Vin de Silva, Jeff Erickson, and Robert Ghrist. Vietoris-rips complexes of planar point sets. *Discrete & Computational Geometry*, 44(1) :75–90, 2010. (Cité en pages 21 et 70.)
- [CG06] Frédéric Cazals and Joachim Giesen. Delaunay triangulation based surface reconstruction : Ideas and algorithms. In *EFFECTIVE COMPUTATIONAL GEOMETRY FOR CURVES AND SURFACES*, pages 231–273. Springer, 2006. (Cité en page 2.)
- [cga] CGAL, Computational Geometry Algorithms Library. <http://www.cgal.org>. (Cité en page 35.)
- [Chi67] D. R. J. Chillingworth. Collapsing three-dimensional convex polyhedra. *Mathematical Proceedings of the Cambridge Philosophical Society*, 63 :353–357, 3 1967. (Cité en page 70.)
- [CL05] Frédéric Chazal and André Lieutier. The " $\lambda$ -medial axis". *Graphical Models*, 67(4) :304–331, 2005. (Cité en page 104.)
- [CL08] Frédéric Chazal and André Lieutier. Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees. *Comput. Geom.*, 40(2) :156–170, 2008. (Cité en pages 14 et 88.)
- [Cla06a] Kenneth L. Clarkson. Building triangulations using  $\epsilon$ -nets. In *STOC 2006 : Proceedings of the Thirty-eighth Annual SIGACT Symposium*, 2006. (Cité en page 81.)
- [Cla06b] Kenneth L. Clarkson. Nearest-neighbor searching and metric space dimensions. In *In Nearest-Neighbor Methods for Learning and Vision : Theory and Practice*. MIT Press, 2006. (Cité en page 39.)
- [CO08] Frédéric Chazal and Steve Oudot. Towards persistence-based reconstruction in euclidean spaces. In Monique Teillaud, editor, *Symposium on Computational Geometry*, pages 232–241. ACM, 2008. (Cité en pages 1, 3, 14, 65 et 66.)
- [CS03] Gunnar Carlsson and Vin De Silva. Topological approximation by small simplicial complexes. Technical report, MISCHAIKOW, AND T. WANNER, 2003. (Cité en pages 3, 14 et 82.)
- [DEG99] T. K. Dey, H. Edelsbrunner, and S. Guha. Computational topology. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*. AMS, Providence, 1999. (Cité en pages 11, 57 et 71.)
- [DEGN98] Tamal K. Dey, Herbert Edelsbrunner, Sumanta Guha, and Dmitry V. Nekhayev. Topology preserving edge contraction. *Publ. Inst. Math. (Beograd) (N.S)*, 66 :23–45, 1998. (Cité en pages 5, 52 et 65.)

- [DFGH04] Leila De Floriani, David Greenfieldboyce, and Annie Hui. A data structure for non-manifold simplicial d-complexes. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, SGP '04, pages 83–92, New York, NY, USA, 2004. ACM. (Cité en page 36.)
- [DGGZ02] Tamal K. Dey, Joachim Giesen, Samrat Goswami, and Wulue Zhao. Shape dimension and approximation from samples. In *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '02, pages 772–780, Philadelphia, PA, USA, 2002. Society for Industrial and Applied Mathematics. (Cité en page 70.)
- [DI12] Stefan Dantchev and Ioannis Ivrissimtzis. Technical section : Efficient construction of the čech complex. *Comput. Graph.*, 36(6) :708–713, October 2012. (Cité en pages 3 et 14.)
- [DL87] D. P. Dobkin and M. J. Laszlo. Primitives for the manipulation of three-dimensional subdivisions. In *Proceedings of the third annual symposium on Computational geometry*, SCG '87, pages 86–99, New York, NY, USA, 1987. ACM. (Cité en page 35.)
- [DSC04] Vin De Silva and Gunnar Carlsson. Topological estimation using witness complexes. In *Proceedings of the First Eurographics conference on Point-Based Graphics*, SPBG'04, pages 157–166, Aire-la-Ville, Switzerland, Switzerland, 2004. Eurographics Association. (Cité en pages 3 et 81.)
- [dSG07] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7 :339–358, 2007. (Cité en pages 3, 7 et 11.)
- [Ede87] Herbert Edelsbrunner. *Algorithms in combinatorial geometry*. Springer-Verlag New York, Inc., New York, NY, USA, 1987. (Cité en page 36.)
- [EH10] Herbert Edelsbrunner and John Harer. *Computational Topology - an Introduction*. American Mathematical Society, 2010. (Cité en page 44.)
- [ES97] Herbert Edelsbrunner and Nimish R. Shah. Triangulating topological spaces. *Int. J. Comput. Geometry Appl.*, 7(4) :365–378, 1997. (Cité en page 10.)
- [GH97] Michael Garland and Paul S. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 209–216, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co. (Cité en pages 52 et 59.)
- [GMM11] Leonidas J. Guibas, Quentin Mérigot, and Dmitriy Morozov. Witnessed k-distance. In Ferran Hurtado and Marc J. Van Kreveld, editors, *27th Annual Symposium on Computational Geometry, SoCG 2011, June, 2011*, pages 57–64, Paris, France, June 2011. ACM, ACM. (Cité en pages 2 et 88.)
- [Gro93] K. Grove. Critical point theory for distance functions. In *Proc. of Symposia in Pure Mathematics*, volume 54, 1993. (Cité en pages 17 et 20.)
- [GS85] Leonidas Guibas and Jorge Stolfi. Primitives for the manipulation of general subdivisions and the computation of voronoi. *ACM Trans. Graph.*, 4(2) :74–123, April 1985. (Cité en page 35.)

- [GW03] Joachim Giesen and Uli Wagner. Shape dimension and intrinsic metric from samples of manifolds with high co-dimension. In *Discrete and Computational Geometry*, pages 329–337. Press, 2003. (Cit  en pages 1 et 70.)
- [Hat02] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002. (Cit  en pages 11 et 25.)
- [Hau95] J.C. Hausmann. *On the Vietoris-Rips Complexes and a Cohomology Theory for Metric Spaces*. Prospects in topology : proceedings of a conference in honor of William Browder. Princeton Univ Pr, 1995. (Cit  en pages 14 et 33.)
- [HDD<sup>+</sup>93] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Mesh optimization. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques, SIGGRAPH '93*, pages 19–26, New York, NY, USA, 1993. ACM. (Cit  en page 52.)
- [JW10] Gwena l Joret and David R. Wood. Irreducible triangulations are small. *J. Comb. Theory Ser. B*, 100(5) :446–455, September 2010. (Cit  en page 65.)
- [KMS98] T. Kaczynski, M. Mrozek, and M. Slusarek. Homology computation by reduction of chain complexes. *Computers & Mathematics with Applications*, 35(4) :59–70, February 1998. (Cit  en page 51.)
- [Lat01] Janko Latschev. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6) :522–528, 2001. (Cit  en pages 14 et 33.)
- [Law87] Serge Lawrencenko. Irreducible triangulations of a torus. In *Ukrain. Geom. Sb.*, volume 30, page 52–62, 1987. (Cit  en page 65.)
- [Ler45] J. Leray. Sur la forme des espaces topologiques et sur les points fixes des repr sentations . *J. Math. Pures Appl.*, 24 :95–167, 1945. (Cit  en page 53.)
- [Lie91] Pascal Lienhardt. Topological models for boundary representation : a comparison with n-dimensional generalized maps. *Computer-Aided Design*, 23(1) :59 – 82, 1991. (Cit  en page 36.)
- [Lie04] Andr  Lieutier. Any open bounded subset of  $\mathbb{R}^n$  has the same homotopy type as its medial axis. *Computer-Aided Design*, 36(11) :1029–1046, 2004. (Cit  en pages 103 et 104.)
- [LT97] H lio Lopes and Geovan Tavares. Structural operators for modeling 3-manifolds. In *Proceedings of the fourth ACM symposium on Solid modeling and applications, SMA '97*, pages 10–18, New York, NY, USA, 1997. ACM. (Cit  en page 35.)
- [Mel11] Sergey A. Melikhov. Combinatorics of embeddings. 2011. (Cit  en pages 55 et 56.)
- [MP78] David E. Muller and Franco P. Preparata. Finding the intersection of two convex polyhedra. *Theor. Comput. Sci.*, 7 :217–236, 1978. (Cit  en pages 35 et 45.)
- [Mun93] J.R. Munkres. *Elements of Algebraic Topology*. Perseus Book, 1993. (Cit  en page 8.)
- [Nag08] Uwe Nagel. Empty simplices of polytopes and graded betti numbers. *Discrete & Computational Geometry*, 39(1-3) :389–410, 2008. (Cit  en pages 36 et 38.)



- [Nev06] Eran Nevo. Higher minors and van kampen's obstruction. 2006. (Cité en pages 52 et 57.)
- [NNM96] S. Nene, S. Nayar, and H. Murase. Columbia Object Image Library (COIL-100). Technical report, Columbia University, 1996. (Cité en page 62.)
- [NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1) :419–441, March 2008. (Cité en pages 3, 13, 14 et 31.)
- [RS00] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *SCIENCE*, 290 :2323–2326, 2000. (Cité en page 1.)
- [Rub95] Joachim H. Rubinstein. An algorithm to recognize the  $S^3$ -sphere. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 601–611, Basel, 1995. Birkhäuser. (Cité en page 65.)
- [She12] Donald R. Sheehy. Linear-size approximations to the vietoris-rips filtration. In *Proceedings of the 2012 symposium on Computational Geometry*, SoCG '12, pages 239–248, New York, NY, USA, 2012. ACM. (Cité en page 59.)
- [Sil] Vin Silva. A weak characterisation of the Delaunay triangulation. *Geometriae Dedicata*, 135(1) :39–64. (Cité en page 3.)
- [Smi11] Stephen D. Smith. *Subgroup Complexes*. Progress in Mathematics. American Mathematical Society, 2011. (Cité en page 12.)
- [SR34] E. Steinitz and H. Rademacher. *Vorlesungen über die Theorie der Polyeder : unter Einschluss der Elements der Topologie*. Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Julius Springer, 1934. (Cité en page 65.)
- [Tan12] Martin Tancer. Recognition of collapsible complexes is np-complete. *CoRR*, abs/1211.6254, 2012. (Cité en pages 70 et 72.)
- [Tho94] Abigail Thompson. Thin position and the recognition problem for  $S^2 \times S^3$ . *Math. Res. Lett.*, 1(5) :613–630, 1994. (Cité en page 65.)
- [TSL00] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500) :2319–2323, 2000. (Cité en page 1.)
- [YJLM10] Anna Yershova, Swati Jain, Steven M. Lavalley, and Julie C. Mitchell. Generating uniform incremental grids on  $so(3)$  using the hopf fibration. *Int. J. Rob. Res.*, 29(7) :801–812, June 2010. (Cité en page 63.)
- [Zom10a] Afra Zomorodian. Fast construction of the vietoris-rips complex. *Computers & Graphics*, 34(3) :263–271, 2010. (Cité en pages 3 et 13.)
- [Zom10b] Afra Zomorodian. The tidy set : a minimal simplicial set for computing homology of clique complexes. In *Proceedings of the 2010 annual symposium on Computational geometry*, SoCG '10, pages 257–266, New York, NY, USA, 2010. ACM. (Cité en pages 3, 13, 36, 48 et 51.)

# Index

- $(\varepsilon, \delta)$ -échantillon, 59
- Centers( $X, t$ ), 15
- Hull( $X, t$ ), 15
- $\mu$ -reach, 29
- $\varepsilon$ -cover, 40
- $\varepsilon$ -net, 59
- $\varepsilon$ -packing, 40
- $\varepsilon$ -échantillon, 82
- Flag-complexe, 8
- Join, 9
- Link, 9, 53
- Offset, 7
- Shadow, 8
- Star, 9
  
- Axe médian, 2
  
- Bloqueur, 37
- Bloqueur (simplement) redondant, 48
  
- Cône, 11, 44
- Centre, 7
- Collapsable, 11
- Collapse élémentaire, 11, 44, 71
- Collapse étendu, 11, 71
- Collapse classique, 11
- Collapse généralisé, 11, 71
- Complexe de Čech, 10
- Complexe de Rips, 11
- Contractile, 9
- Contraction d'arête, 46, 53
  
- Défaut de convexité global  $\lambda_P$ , 71
- Diamètre, 7
- Distance de Hausdorff, 7
  
- Enveloppe convexe, 7
- Expansion, 8
  
- Face, Coface, 8
- Fonction critique, 28
  
- Fonctions de défaut de convexité  $c_P$  et  $h_P$ , 15
- Graphe de Gabriel, 82
- Homéomorphe, 9
- Inégalité de Jung, 7
- Nerf, 53
- Nombres de *covering* et de *packing*, 40
  
- Points et valeurs critiques de la fonction de distance, 17
  
- Quasi complexe de Rips, 21
  
- Réalisation géométrique, 8
- Rétracte par déformation, 9
- Rayon, 7
  
- Simplexe abstrait, 8
  
- Type d'homotopie, 9
  
- Valeur inerte, 21



# Table des figures

1.1	Les deux premiers espaces topologiques sont homéomorphes contrairement aux deux derniers. Les trois espaces ont le même type d'homotopie. . . . .	9
1.2	A gauche : le complexe de Čech avec le paramètre $\alpha$ comprend six triangles et a le type d'homotopie d'un cercle. Au milieu : le complexe de Rips de paramètre $\alpha$ contient deux triangles en plus et est homéomorphe à une 2-sphère. Son shadow est un disque topologique. A droite : le complexe de Čech de paramètre $\vartheta_2 \alpha$ contient toutes les faces d'un 5-simplexe et est homéomorphe à une 5-boule. . .	10
1.3	Illustration d'un collapse classique (gauche), étendu (milieu) et généralisé (droite). Le <i>link</i> de $\sigma$ , représenté en vert, est respectivement la fermeture d'un simplexe, un cône et un complexe contractile. . . . .	12
2.1	A gauche, plus petit <i>offset</i> de $X$ contenant $\text{Centers}(X, t)$ . A droite, plus petit <i>offset</i> de $X$ contenant $\text{Hull}(X, t)$ . . . . .	16
2.2	Notations pour la preuve du lemme 2.1. . . . .	17
2.3	Différents ensembles de points $P$ sur la gauche et tracé des fonctions de défaut de convexité $c_P$ respectives sur la droite. . . . .	18
2.4	Les trois premières figures représentent $\text{Centers}(P, t)$ pour $t \in \{0, 3; 0, 5; 0, 72\}$ . La dernière figure donne le graphe de la fonction $c_P(t)$ . . . . .	19
2.5	Les points noirs sont les points critiques de la fonction de distance à la courbe $X$ . . . . .	19
2.6	Notations pour la preuve du théorème 2.3. A gauche : $\tau_i$ est le simplexe dont les sommets sont les points de $P$ dans la boule $B(z_i, t_i)$ . $\sigma_i$ est la face obtenue en gardant les sommets sur le bord de $B(z_i, t_i)$ . A droite : représentation schématique des simplexes de $\mathcal{C}_i \setminus \mathcal{C}_{i-1}$ . . . . .	24
2.7	A gauche : le complexe de Rips et le complexe de Čech sont contractiles lorsque $\gamma = (2 - \vartheta_D)\alpha$ et $D = 2$ . A droite : si on augmente infinitésimalement la valeur de $\gamma$ , le complexe de Rips est une sphère tandis que le complexe de Čech est toujours contractile. . . . .	26

2.8	A gauche : pour $i \in \{2.3, 2.4, 2.5\}$ , les hypothèses du théorème $i$ sont représentées par les régions $Z_i$ qui ne croisent pas le graphe de la fonction de défaut de convexité. Plus précisément, si $c_P \cap Z_7 = \emptyset$ , le théorème 2.3 implique que $\text{Rips}(P, \alpha) \simeq P^\alpha$ . Si $h_X \cap Z_9 = \emptyset$ , le théorème 2.4 implique que $\text{Cech}(P, \alpha) \simeq X^{\alpha-2\varepsilon}$ . Si $h_X \cap Z_{10} = \emptyset$ , le théorème 2.5 implique que $\text{Rips}(P, \alpha) \simeq X^{2\alpha-\vartheta_D\alpha-2\varepsilon}$ . A droite : borne supérieure sur $h_X$ pour $\mu \in \{0, \frac{1}{3}, \frac{1}{2}, 1\}$ fournie par le lemme 2.6.	27
2.9	Notations pour la preuve du lemme 2.5.	29
2.10	(a) Meilleurs ratios $\frac{\varepsilon}{R}$ que l'on obtient pour une reconstruction correcte d'une forme à $\mu$ -reach positif avec le complexe de Čech ou le complexe de Rips pour $D \in \{2, 3, 4, 5, +\infty\}$ ; comparaison avec le ratio $\frac{\mu^2}{5\mu^2+12}$ obtenu dans [CCSL09]. (b) et (c) $\mu_D^*$ et $\lambda_D^{\text{rips}}(1)$ tracées en fonction de $D$ .	30
2.11	Tracés des deux fonctions ( $\alpha_D^{\min}(\mu; \cdot; 1), \alpha_D^{\max}(\mu; \cdot; 1)$ ) pour $\mu = 0,92$ (gauche) et $\mu = 1$ (droite) lorsque $D \rightarrow \infty$ .	33
2.12	Tracé de $\alpha_D^{\min}(\mu, \varepsilon, 1)$ et $\alpha_D^{\max}(\mu, \varepsilon, 1)$ en fonction de $\mu$ et $\varepsilon$ lorsque $D \rightarrow \infty$ . Les deux coupes selon $\mu = 0,92$ et $\mu = 1$ de la figure 2.11 sont représentées en rouge. Etant donné $\mu$ et $\varepsilon$ , il faut que $\alpha$ soit entre ces deux surfaces pour pouvoir garantir le type d'homotopie du complexe de Rips.	33
3.1	Diagramme de Hasse de $K$ .	38
3.2	Gauche : Un complexe simplicial constitué de six sommets, dix arêtes et quatre triangles. Milieu : le même complexe simplicial représenté cette fois-ci par son 1-squelette et l'ensemble de ses bloqueurs d'ordre 1 $\{cdf, bcd\}$ . Droite : Diagramme de Hasse de l'expansion du 1-squelette. Les nœuds en violet appartiennent au complexe simplicial. Les bloqueurs et les simplexes maximaux pour l'inclusion sont représentés par des nœuds encadrés.	39
3.3	Représentation de la structure de données pour le complexe simplicial de la Figure 3.2.	40
3.4	Nombre de $k$ -simplexes dans $\text{Rips}(P, \alpha)$ en fonction de $k$ pour un ensemble de points échantillonnant $C_2$ . Une échelle logarithmique est utilisée pour l'axe $y$ . Expérimentalement, nous observons que le nombre de $k$ -simplexes est inférieur mais proche de $\frac{n}{k+1} \binom{N}{k}$ pour $N = 25$ . Le nombre de simplexes est donc au moins $\frac{n}{N+1} 2^N \approx 3.4 \cdot 10^9$ . Ce nombre est à comparer avec notre structure de données qui a juste besoin de stocker un graphe constitué de 2646 sommets et 80304 arêtes pour représenter le même complexe.	42

- 3.5 Le triangle  $ayz$  est un 2-bloqueur. Les *links* de  $a$ , de  $b$  et de  $ab$  valent respectivement  $\text{Lk}(a) = \{x, y, z, xy, xz, b, by, bz\}$ ,  $\text{Lk}(b) = \{y, z, yz, a, ay, az\}$  et  $\text{Lk}(ab) = \{y, z\}$ . Observons que  $\sigma = \alpha\beta$  avec  $\alpha = x$  et  $\beta = yz$  vérifie les conditions (i) et (ii) du lemme 3.6. Ainsi, la contraction d'arête  $ab \mapsto c$  engendre la création du 3-bloqueur  $xyz$  et la destruction du 2-bloqueur  $ayz$ . . . . . 47
- 4.1 De gauche à droite : le *link* de  $a$ , le *link* de  $b$ , le *link* de  $ab$  et le complexe simplicial obtenu après la contraction d'arête  $ab$ . On remarque que  $ab$  vérifie la condition du *link*. De manière équivalente, aucun bloqueur ne contient  $ab$  (voir le lemme 4.2). La contraction d'arête  $ab$  préserve donc le type d'homotopie. Notons que cette contraction d'arête crée un bloqueur  $cxy$  et que  $\alpha = x$  and  $\beta = y$  vérifient les hypothèses du lemme 3.6 du chapitre 3 caractérisant les bloqueurs d'un complexe après une contraction d'arête. La contraction de n'importe quelle arête parmi  $cv$ ,  $xv$  ou  $yv$  enlève le bloqueur  $cxy$ . . . . . 54
- 4.2 En haut et en bas : deux complexes simpliciaux de dimension 2 avec un bloqueur passant par  $ab$ . A droite, la contraction d'arête  $ab \mapsto c$  peut changer le type d'homotopie (en bas) ou le préserver (en haut). . . . . 55
- 4.3 Illustration de la simplification par contractions d'arêtes. En haut à gauche, l'ensemble de points **Pillow**. En haut à droite, Le complexe de Rips construit avec ces points. En bas à gauche, le complexe obtenu après 5807 contractions d'arêtes (il reste alors 1000 sommets sur les 6807 sommets initiaux); chaque sommet du complexe est lié à tous les points de ses préimages par des segments verts clairs. En bas à droite, le complexe obtenu après 6707 contractions, il ne reste alors que 100 sommets. . . . . 60
- 4.4 Illustration de la simplification par contractions d'arêtes. En haut à gauche, le complexe de Rips d'un ensemble de 2000 points obtenus en échantillonnant uniformément  $S_2$ . En haut à droite, le complexe obtenu après 1500 contractions d'arêtes (il reste alors 500 sommets sur les 2000 sommets initiaux); chaque sommet du complexe est lié à tous les points de ses préimages par des segments verts clairs. En bas à gauche, le complexe obtenu après 1900 contractions, il ne reste alors que 100 sommets. En bas à droite, le complexe obtenu après 1900 contractions ainsi que les points initiaux. . . . . 61
- 4.5 Douze points parmi les 72 points de l'ensemble **Cat**. Chaque point représente une image de  $128 \times 128$  pixels d'une statue en forme de chat prise sous différentes poses durant sa rotation autour d'un axe fixe. . . . . 62
- 4.6 Gauche : Complexe de Rips construit avec les points de **Cat**. Droite : Après simplification, on obtient un complexe simplicial constitué de trois sommets, trois arêtes et un unique bloqueur. . . . . 64

4.7	Nous traçons en fonction de $i$ , le nombre de bloqueurs dans $K_i$ à gauche, le nombre cumulé de bloqueurs redondants enlevés de la structure de données aux étapes antérieures à $i$ au milieu, et le degré moyen de chaque sommet dans $K_i$ à droite. De haut en bas, le jeu de données est respectivement <b>C2</b> , <b>S2</b> , <b>C3</b> , <b>S2</b> , <b>S03</b> , et <b>Ramses</b> . . . . .	67
5.1	Notation pour la preuve du théorème 5.1. A gauche : les quatre boules unité ouvertes contiennent l'enveloppe convexe de $a, b, c$ et $d$ . De gauche à droite : la preuve consiste à balayer l'espace avec une sphère centrée en $o \in \text{Hull}(\{a, b, c, d\})$ dont le rayon $t$ décroît continûment de $+\infty$ vers 0. On déduit du balayage une séquence de collapses réduisant $\text{Cech}(P, \alpha)$ en un sommet. . . . .	73
5.2	Notation pour la preuve du théorème 5.3. Deux types d'événements peuvent se produire : un collapse de sommet (sur la gauche) ou un collapse d'arête (sur la droite). Le cas d'un collapse d'arête est illustré lorsque le triangle $oab$ est équilatéral. . . . .	75
5.3	Notation pour la preuve du théorème 5.4. . . . .	76
5.4	Notation pour la preuve du lemme 5.2 dans le cas 1 (gauche) et dans le cas 2 (droite). . . . .	80
5.5	Notation pour la preuve de la proposition 5.2 . . . . .	83
5.6	Nombre d'échecs pour chaque stratégie lorsque l'on collapse les sommets (gauche) et lorsque l'on collapse des arêtes (droite) pour les ensembles de points (de haut en bas) <b>Cat</b> , <b>S2</b> , <b>Ramses</b> , <b>S03</b> . . . . .	86
A.1	A gauche : la borne sur $\  \text{Center}(\sigma) - \text{Center}(\sigma') \ $ du lemme A.1 est optimale. A droite : notation de la preuve du lemme A.1. . . . .	101
A.2	Notations pour la preuve du lemme 2.6. . . . .	103
B.1	Notations pour la preuve du lemme B.1. A gauche : lorsque $\gamma$ est proche de $r$ . A droite : lorsque $\gamma$ est proche de $R$ . . . . .	106
E.1	Notation pour la preuve du lemme E.2. . . . .	117
E.2	Notation pour la preuve du lemme 5.1. . . . .	119

# Lemmes supplémentaires pour le chapitre 2

## A.1 Stabilité de la plus petite boule englobante

Dans cette annexe, nous prouvons que le rayon et le centre de la plus petite boule englobante sont stables par perturbation de  $\sigma$ .

**Lemme A.1.** *Pour tous ensembles  $\sigma, \sigma' \subset \mathbb{R}^n$  non-vides et bornés tels que  $d_H(\sigma, \sigma') \leq \varepsilon$ , les deux inégalités suivantes sont vérifiées :  $|\text{Rad}(\sigma) - \text{Rad}(\sigma')| \leq \varepsilon$  et  $\|\text{Center}(\sigma) - \text{Center}(\sigma')\| \leq \sqrt{2\varepsilon \text{Rad}(\sigma) + \varepsilon^2}$ .*

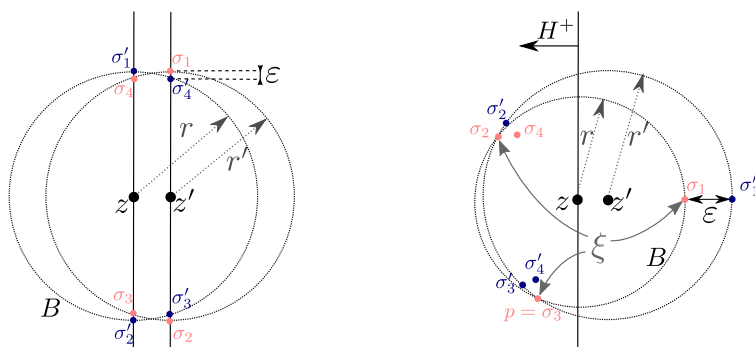


FIGURE A.1 – A gauche : la borne sur  $\|\text{Center}(\sigma) - \text{Center}(\sigma')\|$  du lemme A.1 est optimale. A droite : notation de la preuve du lemme A.1.

*Démonstration.* Soit  $B$  la plus petite boule englobante de  $\sigma$ , nous avons  $\sigma' \subset \sigma^\varepsilon \subset B^\varepsilon$ , ce qui montre que  $\text{Rad}(\sigma') \leq \text{Rad}(\sigma) + \varepsilon$ . Pour la deuxième affirmation du lemme, notons  $z = \text{Center}(\sigma)$ ,  $z' = \text{Center}(\sigma')$ ,  $r = \text{Rad}(\sigma)$  et  $r' = \text{Rad}(\sigma')$ ; voir la figure A.1. Supposons que  $z \neq z'$  (le résultat est évident sinon) et considérons l'hyperplan  $H$  passant par  $z$  et orthogonal au segment  $zz'$ . Soit  $\xi = \text{Cl}(\sigma) \cap \partial B$ . Par construction,  $\xi$  est fermé et a la même plus petite boule englobante que  $\sigma$ . Ainsi,  $z \in \text{Hull}(\xi)$  et le demi-espace fermé  $H^+$  borné par  $H$  et qui évite que  $z'$  intersecte  $\xi$ . Soit  $p$  un point dans cette intersection. Par choix de  $p$  dans  $H^+$ , le triangle  $pzz'$  est obtus sur le sommet  $z$  et  $\|z' - p\|^2 \geq \|z - p\|^2 + \|z' - z\|^2$ . En utilisant le fait que  $\|z - p\| = r$  et  $\|z' - p\| \leq r' + \varepsilon$  on obtient que  $\|z' - z\|^2 \leq (r' + \varepsilon)^2 - r^2$ . En permutant



$\sigma$  et  $\sigma'$  on obtient :

$$\|z' - z\|^2 \leq \min\{(r' + \varepsilon)^2 - r^2, (r + \varepsilon)^2 - r'^2\}.$$

Notons que  $(r' + \varepsilon)^2 - r^2 \leq (r + \varepsilon)^2 - r'^2$  si et seulement si  $r \geq r'$ . En considérant les deux cas  $r \leq r'$  et  $r' \leq r$ , nous obtenons l'inégalité recherchée.  $\square$

## A.2 Stabilité des hypothèses du théorème 2.3

Dans cette annexe, nous établissons la stabilité des hypothèses du théorème 2.3 sous des petites perturbations de l'ensemble de points. Etant donné un ensemble de points  $P \subset \mathbb{R}^n$ , on dit qu'une application  $f : P \rightarrow \mathbb{R}^n$  est une  $\varepsilon$ -perturbation de  $P$  si  $f$  est injective et  $\|p - f(p)\| \leq \varepsilon$  pour tout point  $p \in P$ . Etant donné un complexe simplicial  $K$ , on définit le complexe simplicial  $f(K) = \{f(\sigma) \mid \sigma \in K\}$ .

**Lemme A.2.** *Soit  $P \subset \mathbb{R}^n$  un ensemble fini de points. Considérons deux réels  $\beta \geq \alpha \geq 0$  tels que*

$$c_P(\vartheta_n \beta) < 2\alpha - \vartheta_n \beta$$

*et supposons de plus que  $\alpha$  est une valeur inerte de  $P$ . Alors, il existe  $\varepsilon > 0$  et  $\beta' > \beta$  tel que pour toute  $\varepsilon$ -perturbation  $f$  de  $P$ ,*

- (i)  $c_{f(P)}(\vartheta_n \beta') < 2\alpha - \vartheta_n \beta'$  ;
- (ii)  $\text{Cech}(f(P), \alpha) = f(\text{Cech}(P, \alpha))$  ;
- (iii) *si  $\text{Flag}(G)$  est un  $(\alpha, \beta)$ -quasi-complexe de Rips de  $P$ , alors  $\text{Flag}(f(G))$  est un  $(\alpha, \beta')$ -quasi complexe de Rips de  $f(P)$ .*

*Démonstration.* Montrons que (i) est vérifiée. Posons  $t = \vartheta_n \beta$  et définissons  $\bar{t} = \min\{\text{Rad}(\sigma) \mid \emptyset \neq \sigma \subset P \text{ et } \text{Rad}(\sigma) > t\}$ . Par construction,  $\bar{t} > t$ . Le lemme 2.3 garantit que pour tout sous-ensemble  $P' \subset \mathbb{R}^n$  dont la distance de Hausdorff à  $P$  est  $\varepsilon$  et pour tout  $t' \geq 0$  :

$$c_{P'}(t') < c_P(t' + \varepsilon) + \sqrt{2t'\varepsilon + \varepsilon^2} + \varepsilon.$$

Par hypothèse, nous avons que  $2\alpha - t - c_P(t) > 0$ . En choisissant  $\varepsilon > 0$  suffisamment petit, on peut toujours trouver  $t' > t$  tel que (1)  $t' + \varepsilon < \bar{t}$ , (2)  $2\alpha - t' - c_P(t) > 0$  et (3)  $\sqrt{2t'\varepsilon + \varepsilon^2} + \varepsilon \leq \frac{2\alpha - t' - c_P(t)}{2}$ . Puisque  $c_P(t' + \varepsilon) = c_P(t)$ , on en déduit que

$$c_{P'}(t') < c_P(t) + \frac{2\alpha - t' - c_P(t)}{2} < 2\alpha - t'$$

et (i) est prouvée pour  $\beta' = t' / \vartheta_n$ . En choisissant  $\varepsilon > 0$  suffisamment petit, on peut toujours supposer, en plus des conditions (1), (2) et (3), la condition (4)  $\varepsilon < \beta' - \beta$  et (5)  $\text{Rad}(\sigma) \notin [\alpha - \varepsilon, \alpha + \varepsilon]$  pour tout  $\emptyset \neq \sigma \subset P$ . Soit  $f$  une  $\varepsilon$ -perturbation de  $P$ . En utilisant le lemme A.1 et la condition (5), on obtient

$$\begin{aligned} \sigma \in \text{Cech}(P, \alpha) &\Leftrightarrow \text{Rad}(\sigma) \leq \alpha \Leftrightarrow \text{Rad}(\sigma) \leq \alpha - \varepsilon \\ &\Rightarrow \text{Rad}(f(\sigma)) \leq \alpha \Leftrightarrow f(\sigma) \in \text{Cech}(f(P), \alpha) \end{aligned}$$

et

$$\begin{aligned} f(\sigma) \in \text{Cech}(f(P), \alpha) &\Leftrightarrow \text{Rad}(f(\sigma)) \leq \alpha \Rightarrow \text{Rad}(\sigma) \leq \alpha + \varepsilon \\ &\Leftrightarrow \text{Rad}(\sigma) \leq \alpha \Leftrightarrow \sigma \in \text{Cech}(P, \alpha), \end{aligned}$$

ce qui prouve (ii). Soit  $G$  un graphe dont le *flag*-complexe est un  $(\alpha, \beta)$ -quasi complexe de Rips et soient  $p$  et  $q$  deux points de  $P$  tels que  $\|f(p) - f(q)\| \leq 2\alpha$ . On a  $\|p - q\| \leq 2\alpha + 2\varepsilon$  et ainsi, en utilisant la condition (5),  $\|p - q\| \leq 2\alpha$ . On en déduit que l'arête  $\{p, q\}$  appartient à  $G$  et donc l'arête  $\{f(p), f(q)\}$  appartient à  $f(G)$ . Similairement, supposons que  $\|f(p) - f(q)\| > 2\beta'$ . Cela implique que  $\|p - q\| > 2\beta' - 2\varepsilon > 2\beta$  avec la condition (4) et ainsi l'arête  $\{p, q\}$  n'appartient pas à  $G$ . L'arête  $\{f(p), f(q)\}$  n'appartient donc pas à  $f(G)$ , ce qui montre (iii).  $\square$

### A.3 Preuve du lemme 2.6

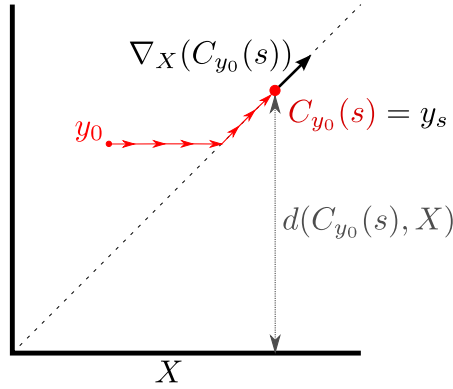


FIGURE A.2 – Notations pour la preuve du lemme 2.6.

*Preuve du lemme 2.6.* Etant donné  $\sigma \subset X$  tel que  $\text{Rad}(\sigma) \leq R$  et  $y_0 \in \text{Hull}(\sigma)$ , nous établissons une borne supérieure sur  $d(y_0, X)$  que l'on exprime comme une fonction de  $\text{Rad}(\sigma)$ .

La première étape de cette preuve est de trouver un point  $y_T$  qui soit "suffisamment" loin de  $X$  en suivant une courbe intégrale du gradient généralisé  $\nabla_X$  en partant de  $y_0$ . Soit  $X^c = \mathbb{R}^n \setminus X$ . Les auteurs de [Lie04] ont montré qu'il existe une application continue  $\Phi_X : \mathbb{R}_+ \times X^c \rightarrow X^c$  telle que :

$$\frac{d}{dt^+} \Phi_X(t, y) = \nabla_X(\Phi_X(t, y))$$

où  $\frac{d}{dt^+}$  dénote la dérivée à droite. Ainsi,  $\Phi_X$  est un flot et la courbe intégrale  $t \mapsto \Phi_X(t, y_0)$  est une courbe rectifiable commençant en  $y_0$ . Cette courbe est soit de longueur finie soit elle termine sur un point critique de la fonction de distance de  $X$ . Si pour  $T > 0$  l'ensemble  $\Phi_X([0, T], y_0)$  ne contient pas de point critique alors elle peut être paramétré par une fonction continue  $\mathcal{C}_{y_0} : [0, L] \rightarrow \mathbb{R}^n$  telle que  $\mathcal{C}_{y_0}(0) = y_0$  et telle que la longueur de  $\mathcal{C}_{y_0}([0, s])$  est  $s$ ; voir la partie droite de la figure A.2. Montrons que sous les conditions du lemme, on peut choisir  $T = R - d(y_0, X)$ . Pour tout  $s < R - d(y_0, X)$ , on note  $d(y_s, X) \leq d(y_0, X) + \|y_s - y_0\| \leq$

$d(y_0, X) + s < R$  et ainsi  $\chi_X(d(y_s, X)) \geq \mu$  ce qui implique que  $\|\nabla_X(y_s)\| \geq \mu$ . On en déduit que la courbe intégrale  $\mathcal{C}_{y_0}$  n'atteint pas de point critique tant que  $s < R - d(y_0, X)$  et que  $\mathcal{C}_{y_0}$  peut être paramétré sur l'intervalle  $[0, R - d(y_0, X)]$ . Dès lors, on peut poser  $T = R - d(y_0, X)$ . Il a été établi dans [Lie04, CL05] que :

$$\forall s \in [0, L), \quad \frac{d}{ds^+} d(\mathcal{C}_{y_0}(s), X) = \|\nabla_X(\mathcal{C}_{y_0}(s))\|$$

En intégrant sur l'intervalle  $[0, T]$ , on obtient :

$$\frac{d(y_T, X) - d(y_0, X)}{T} \geq \mu.$$

En appliquant le lemme 2.1 avec  $x = y_T$  et  $y = y_0$ , on a que  $d(y_T, \sigma)^2 \leq T^2 + \text{Rad}(\sigma)^2$ . On peut alors en déduire la suite d'inégalités suivante :

$$(d(y_0, X) + \mu T)^2 \leq d(y_T, X)^2 \leq d(y_T, \sigma)^2 \leq T^2 + \text{Rad}(\sigma)^2.$$

En posant  $T = R - d(y_0, X)$ ,  $\delta = \frac{d(y_0, X)}{R}$ ,  $\rho = \frac{\text{Rad}(\sigma)}{R}$  et en réarrangeant cette inégalité on obtient :

$$\mu(2 - \mu)\delta^2 - 2(1 + \mu - \mu^2)\delta + 1 - \mu^2 + \rho^2 \geq 0.$$

Puisque  $\delta \leq 1$ , on a finalement :  $\delta \leq \frac{1 + \mu(1 - \mu) - \sqrt{1 - \rho^2 \mu(2 - \mu)}}{\mu(2 - \mu)}$ , ce qui prouve le lemme.

□

# Optimalité de l'hypothèse du théorème 2.3

Rappelons la définition de  $P_\gamma$ . Pour cela, notons  $\{v_0, \dots, v_D\}$  les sommets d'un simplexe régulier  $\sigma$  de diamètre  $2\alpha$ . Etant donné  $\gamma > 0$ , notons également  $\bar{v}_i$  le point de la sphère  $S(\text{Center}(\sigma), \gamma)$  le plus éloigné de  $v_i$ . Nous définissons alors l'ensemble des points  $P_\gamma = \{v_0, \dots, v_D, \bar{v}_0, \dots, \bar{v}_D\}$ ; voir la figure B.1. Le lemme suivant donne la valeur des fonctions de défauts de convexité de  $P_\gamma$  en  $\vartheta_D\alpha$  pour  $D = 2$ .

**Lemme B.1.** *Si  $D = 2$  et  $\frac{2-\vartheta_D^2}{\vartheta_D}\alpha \leq \gamma \leq \vartheta_D\alpha$  alors  $c_{P_\gamma}(\vartheta_D\alpha) = h_{P_\gamma}(\vartheta_D\alpha) = \gamma$ .*

*Démonstration.* Notons  $p = \operatorname{argmax}_{p \in \text{Hull}(P_\gamma)} d(p, P_\gamma)$  le point dans l'enveloppe convexe de  $P_\gamma$  le plus loin de  $P_\gamma$  et notons également  $c = \text{Center}(\{v_0, v_1, v_2\})$ . Nous allons montrer que  $p = c$ . Considérons  $\text{Vor}(P_\gamma)$  le diagramme de Voronoi de  $P_\gamma$  et notons  $V_x$  la cellule de Voronoi associée à un point  $x \in P_\gamma$ . Supposons que  $\frac{2-\vartheta_D^2}{\vartheta_D}\alpha \leq \gamma \leq \vartheta_D\alpha$ . En notant  $r$  le rayon de la plus petite boule inscrite dans  $\text{Hull}(\{v_0, v_1, v_2\})$  et  $R = \text{Rad}\{v_0, v_1, v_2\}$ , observons que  $r = \frac{2-\vartheta_D^2}{\vartheta_D}\alpha$  et que  $R = \vartheta_D\alpha$ . En effet, en notant  $\text{Height } \sigma$  la hauteur de  $\sigma$ , on a que  $r = \text{Height } \sigma - \text{Rad } \sigma = (\frac{2}{\vartheta_D} - \vartheta_D)\alpha$ . Ainsi  $r \leq \gamma \leq R$  (voir la figure B.1). Dans ce cas, les intersections  $V_{v_0} \cap V_{\bar{v}_1} \cap V_{\bar{v}_2}$ ,  $V_{\bar{v}_0} \cap V_{v_1} \cap V_{\bar{v}_2}$  et  $V_{\bar{v}_0} \cap V_{\bar{v}_1} \cap V_{v_2}$  sont non vides et on note  $p_0, p_1$  et  $p_2$  les points associés à chacune de ces intersections. Il est facile de voir que  $p$  appartient à l'intérieur de  $\text{Hull}(P_\gamma)$  ainsi  $p$  est un sommet de  $\text{Vor}(P_\gamma)$ . Comme les sommets de  $\text{Vor}(P_\gamma)$  sont réduits à  $\{c, p_0, p_1, p_2\}$ , il suffit de montrer que  $d(c, P_\gamma) \geq d(p_i, P_\gamma)$  pour  $i \in \{0, 1, 2\}$ . Observons que, pour  $r \leq \gamma \leq R$ , le point  $p_i$  appartient à la boule  $B(c, r) \cap [cv_i]$  (voir la figure B.1). Or pour tout  $x \in B(c, r) \cap [cv_i]$ , on a que  $d(x, P_\gamma) \leq d(c, P_\gamma)$  ainsi  $d(c, P_\gamma) \geq d(p_i, P_\gamma)$ . Ainsi  $c$  est le point dans l'enveloppe convexe de  $P_\gamma$  le plus loin de  $P_\gamma$ . De plus,  $c = \text{Center}(\{\bar{v}_0, \bar{v}_1, \bar{v}_2\})$  et  $\text{Rad}(\{\bar{v}_0, \bar{v}_1, \bar{v}_2\}) = \gamma \leq \vartheta_D\alpha$  ainsi  $c \in \text{Centers}(P_\gamma, \vartheta_D\alpha) \subset \text{Hull}(P_\gamma, \vartheta_D\alpha)$ . On obtient alors finalement que :

$$h_{P_\gamma}(\vartheta_D\alpha) = c_{P_\gamma}(\vartheta_D\alpha) = d(c, P_\gamma) = \gamma$$

□

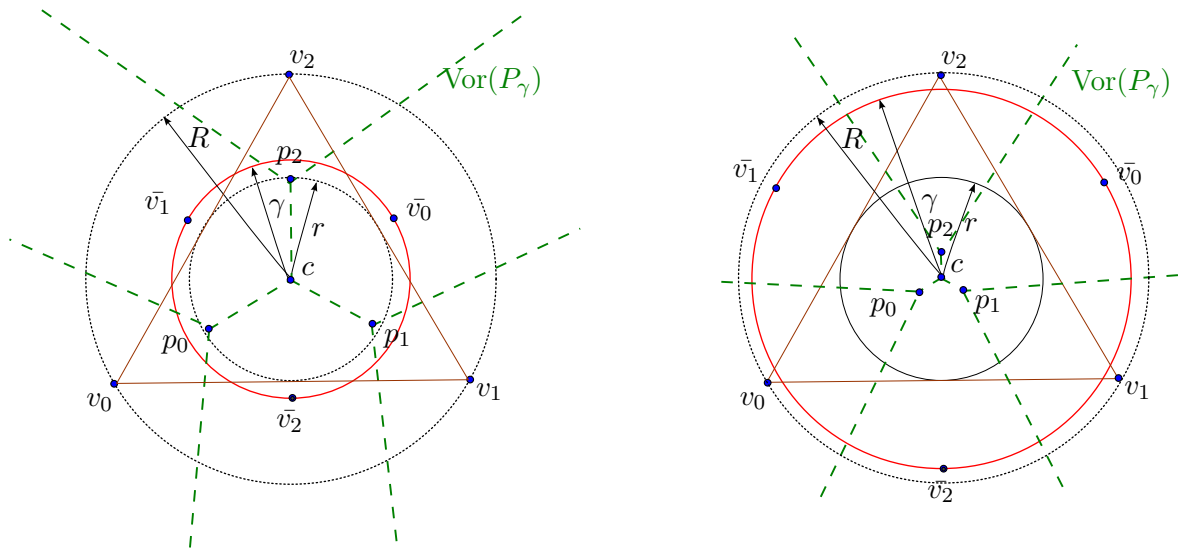


FIGURE B.1 – Notations pour la preuve du lemme B.1. A gauche : lorsque  $\gamma$  est proche de  $r$ . A droite : lorsque  $\gamma$  est proche de  $R$ .

# Constantes pour la reconstruction de formes à *reach* positif

Dans cette annexe, nous refaisons les calculs de la section 2.4.3 en posant  $\mu = 1$ ,  $R = 1$ ,  $n = +\infty$ . Nous considérons donc une forme  $X$  dont le *reach* est supérieur ou égal à 1 et un ensemble fini de points  $P$  tel que  $d_H(P, X) \leq \varepsilon$ .

**Reconstruction avec le complexe de Čech.** En combinant le théorème 2.4 et le lemme 2.6, on obtient que le complexe de Čech  $\check{Cech}(P, \alpha)$  a le même type d'homotopie que  $X^\eta$  pour  $0 < \eta < 1$  dès lors que

$$1 - \sqrt{1 - t^2} < t - 3\varepsilon, \quad \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon]$$

ce qui peut être réécrit comme

$$2t^2 - 2t(1 + 3\varepsilon) + 9\varepsilon^2 + 6\varepsilon < 0, \quad \forall t \in [\alpha + \varepsilon, \alpha + 3\varepsilon].$$

Cette condition est vérifiée dès lors que la valeur absolue de la différence entre les deux racines  $t^1(\varepsilon) \leq t^2(\varepsilon)$  du polynôme en  $t$  est supérieure à  $2\varepsilon$ , c'est-à-dire, lorsque  $0 > 13\varepsilon^2 + 6\varepsilon - 1$ . La borne supérieure pour la valeur  $\varepsilon$  telle que l'équation précédente tient est  $\lambda^{\text{cech}}(1) = \frac{-3 + \sqrt{22}}{13} \approx 0.13$ . Pour  $\varepsilon < \lambda^{\text{cech}}(1)$ , les valeurs admissibles pour  $\alpha$  sont situées dans l'intervalle  $I_1(\varepsilon) = [t^1(\varepsilon) - \varepsilon, t^2(\varepsilon) - 3\varepsilon]$ . En insérant l'expression des racines dans les deux points extrêmes de cet intervalle, on obtient

$$I_1(\varepsilon) = \left[ \frac{1}{2} + \varepsilon - \frac{\sqrt{1 - 6\varepsilon - 9\varepsilon^2}}{2}, \frac{1}{2} - \frac{3\varepsilon}{2} + \frac{\sqrt{1 - 6\varepsilon - 9\varepsilon^2}}{2} \right].$$

Notons que quand  $\varepsilon$  tend vers  $\lambda^{\text{cech}}(1)$ , l'intervalle  $I_1(\varepsilon)$  tend vers un singleton  $\frac{8}{13} - \frac{\sqrt{22}}{26} \approx 0.44$ .

**Reconstruction avec le complexe de Rips.** En combinant le théorème 2.5 et le lemme 2.6, on obtient que le complexe de Rips  $\text{Rips}(P, \alpha)$  a le même type d'homotopie que  $X^\eta$  pour tout  $0 < \eta < R$  dès lors que

$$1 - \sqrt{1 - (\sqrt{2}\alpha + \varepsilon)^2} < 2\alpha - \sqrt{2}\alpha - 2\varepsilon.$$

ce qu'on peut réécrire comme

$$5\varepsilon^2 + 4(2 - \sqrt{2})\alpha^2 - 2(4 - 3\sqrt{2})\alpha\varepsilon + 4\varepsilon - 2(2 - \sqrt{2})\alpha < 0$$

Comme nous cherchons la plus grande valeur de  $\varepsilon$  pour laquelle l'équation précédente tient, nous pouvons supposer que la dérivée partielle selon  $\alpha$  du membre gauche s'annule et obtenir l'équation suivante  $4(2 - \sqrt{2})\alpha - (4 - 3\sqrt{2})\varepsilon - (2 - \sqrt{2}) = 0$ . En insérant  $\alpha = ((1 - \sqrt{2})\varepsilon + 1)/4$  dans l'inégalité précédente, on obtient

$$(10 + 7\sqrt{2})\varepsilon^2 + (8 + 6\sqrt{2})\varepsilon + \sqrt{2} - 2 < 0$$

Le membre de gauche est un polynôme de degré 2 en  $\varepsilon$  dont la plus grande racine  $\varepsilon_{+\infty}^{\text{rips}}(1) = \frac{2\sqrt{2-\sqrt{2}-\sqrt{2}}}{2+\sqrt{2}} \approx 0.034$  donne le supremum de  $\varepsilon$  pour lequel l'inégalité précédente tient. La valeur correspondante pour  $\alpha$  est

$$\alpha_{+\infty}^{\text{rips}}(1) = 1 - \frac{\sqrt{2}}{2} + \left(1 - \frac{3\sqrt{2}}{4}\right) \sqrt{2 - \sqrt{2}} \approx 0.25.$$

# Pseudo-codes

---

Dans un premier temps, nous donnons les pseudo-codes permettant de réaliser plusieurs opérations sur la structure de données squelette/bloqueurs comme calculer le *link* d'un simplexe  $\alpha$ , tester si un simplexe appartient au complexe, tester si un complexe est un cône, calculer les bloqueurs passant par  $c$  après une contraction d'arête  $ab \mapsto c$  ou encore tester si la condition du *link* est vérifiée pour une arête  $ab$ . Nous donnons ensuite les pseudo-codes des algorithmes du chapitre 5 permettant de réduire un complexe en un point.

Pour exprimer les complexités, nous notons  $N_\sigma = \max_{v \in \sigma} \#\mathcal{N}(v)$ ,  $B_\sigma = \max_{v \in \sigma} \#\mathcal{B}(v)$  et  $Q_\sigma = \max_{v \in \sigma} \max_{\beta \in \mathcal{B}(v)} \#\beta$ . Nous notons également  $A_\sigma$  le nombre de bloqueurs inclus dans  $\sigma \cup \mathcal{N}(\sigma)$ . Nous supposons que nous pouvons décider si  $v \in \sigma$  en  $O(\log \#\sigma)$  et décider si  $\tau \subset \sigma$  en  $O(\#\sigma \log \#\sigma)$ . Finalement, nous supposons qu'enlever un sommet de  $\sigma$  se fait en  $O(\log \#\sigma)$ .

## D.1 Calcul du *link* d'un simplexe

Pour calculer la complexité de l'Algorithme 6, on suppose que  $\alpha$  est de taille constante. Calculer les sommets du *link* coûte  $O(N_\alpha \log N_\alpha + \#\mathcal{N}(\alpha)B_{\mathcal{N}(\alpha)})$ . Calculer les arêtes coûte  $O((\#V)^2[\log N_V + B_V \log Q_V])$ . Quant aux bloqueurs, leur calcul coûte  $O((\#V)B_V[\log Q_V + (\#V) \log(\#V)] + A_\alpha B_\alpha (\#V) \log(\#V))$ . En sommant ces trois coûts et en utilisant le fait que  $V \subseteq \mathcal{N}(\alpha)$  et  $\#\mathcal{N}(\alpha) \leq N_\alpha$ , on obtient la complexité de cet algorithme qui est  $O(l(\alpha))$  où :

$$l(\alpha) = (N_\alpha)^2 [\log N_{\mathcal{N}(\alpha)} + B_{\mathcal{N}(\alpha)}(\log N_\alpha + \log Q_{\mathcal{N}(\alpha)})] + A_\alpha B_\alpha N_\alpha \log N_\alpha$$

## D.2 Tester l'appartenance d'un simplexe au complexe

Pour  $L = K$ , la complexité de l'Algorithme 7 est  $O(g(\#\sigma, \sigma))$  où  $g(x, y) = x^2(\log N_y + B_y \log x)$ . Si  $L$  est le *link* d'un sommet  $v \in K$ , cette complexité est aussi bornée par  $O(g(\#\sigma, \sigma))$  puisque pour chaque sommet  $x$  du *link* de  $v$ , on a  $\#\mathcal{N}_{\text{Lk}(v)}(x) \leq \#\mathcal{N}_K(x)$  et  $\#\mathcal{B}_{\text{Lk}(v)}(x) \leq \#\mathcal{B}_K(x)$ .



---

**Algorithme 6** Calcule le 1-squelette  $(V, E)$  et les bloqueurs  $B$  du *link* d'un simplexe  $\alpha$ .

---

```

 $V \leftarrow \emptyset$ 
 $\mathcal{N}(\alpha) \leftarrow \bigcap_{u \in \alpha} \mathcal{N}(u)$ 
for all  $v \in \mathcal{N}(\alpha)$  do
   $newvertex \leftarrow \mathbf{true}$ 
  for all  $\beta \in \mathcal{B}(v)$  do
    if  $\beta \setminus \{v\} \subset \alpha$  then  $newvertex \leftarrow \mathbf{false}$ ; break; end if
  if  $newvertex$  then  $V \leftarrow V \cup \{v\}$  end if
 $E \leftarrow \emptyset$ 
for all  $x \in V$  do
  for all  $y \in V$  such that  $x < y$  and  $y \in \mathcal{N}(x)$  do
     $newedge \leftarrow \mathbf{true}$ 
    for all  $\beta \in \mathcal{B}(x)$  such that  $y \in \beta$  do
      if  $\beta \setminus \{x, y\} \subset \alpha$  then  $newedge \leftarrow \mathbf{false}$ ; break; end if
    if  $newedge$  then  $E \leftarrow E \cup \{x, y\}$  end if
 $B \leftarrow \emptyset$ 
for all  $x \in V$  do
  for all  $\beta \in \mathcal{B}(x)$  do
     $\sigma \leftarrow \beta \setminus \alpha$ 
    if  $\dim \sigma \geq 2$  and  $x$  first vertex of  $\sigma$  and  $\sigma \subset V$  then
       $newblocker \leftarrow \mathbf{true}$ 
      for all  $a \in \alpha$  do
        for all  $\eta \in \mathcal{B}(a)$  such that  $\eta \subset (\sigma \cup \alpha)$  do
          if  $(\eta \setminus \alpha) \subsetneq \sigma$  then  $newblocker \leftarrow \mathbf{false}$ ; break; end if
        if  $newblocker$  then  $B \leftarrow B \cup \{\sigma\}$  end if

```

---



---

**Algorithme 7** Retourne vrai si et seulement si le simplexe  $\sigma \subset \text{Vert}(L)$  appartient au sous-complexe  $L \subset K$ .

---

```

for all  $v \in \sigma$  do
  for all  $w \in \sigma$  such that  $v < w$  do
    if  $w \notin \mathcal{N}_L(v)$  then return false end if
for all  $v \in \sigma$  do
  for all  $\tau \in \mathcal{B}_L(v)$  do
    if  $\tau \subset \sigma$  then return false end if
return true

```

---

## D.3 Calculer les bloqueurs créés par une contraction d'arête

Avant de donner notre algorithme pour énumérer les bloqueurs après une contraction d'arête, nous avons besoin de quelques remarques. Nous supposons que  $\sigma = \alpha\beta$  avec  $\dim \sigma \geq 1$  et  $ab \cap \sigma = \emptyset$  satisfasse les conditions (i) et (ii) du lemme 3.6. Puisque les bloqueurs d'ordre 0 ont une dimension supérieure ou égale à 1, les deux ensembles  $\alpha$  et  $\beta$  sont non vides. En notant  $d_i(v)$  la dimension maximale pour tous les bloqueurs d'ordre  $i$  passant par  $v$ , on déduit directement du lemme que la dimension maximale des bloqueurs d'ordre 1 passant par  $c$  vérifie  $d_1(c) \leq d_0(a) + d_0(b)$  (en particulier la Figure 3.5 est un cas où l'égalité est atteinte). Finalement, on montre que si  $\alpha$  est un sommet, alors  $\alpha \in \mathcal{N}(a)$ . Supposons que  $\alpha$  soit un sommet. Puisque  $\dim \sigma \geq 1$ ,  $\alpha$  est une face propre de  $\sigma$ . Ainsi, (i) implique  $\alpha \in \text{Lk}(a) \cup \text{Lk}(b)$  et (ii) implique  $\alpha \notin \text{Lk}(b)$ , on a alors  $\alpha \in \text{Lk}(a)$ .

Nous pouvons maintenant exprimer l'ensemble des bloqueurs passant par  $c$  dans notre structure de données. En premier lieu, notons que

$$\text{Blockers}_0(K) = \text{Blockers}_1(K) \cup \{xy \mid x \in K^{(0)}, y \in K^{(0)}, xy \notin K^{(1)}\}.$$

On en déduit que  $b\alpha \in \text{Blockers}_0(K)$  si et seulement si  $b\alpha \in \text{Blockers}_1(K)$  ou  $\alpha \in K^{(0)} \setminus \mathcal{N}(b)$ . Ainsi, pour parcourir les simplexes  $\sigma = \alpha\beta$  avec  $\dim \sigma \geq 1$ ,  $ab \cap \sigma = \emptyset$  satisfaisant (i) et (ii), il suffit de prendre  $\alpha$  dans l'ensemble de simplexes suivant :

$$Z_a(b) = \{\alpha \mid b\alpha \in \mathcal{B}(b), ab \cap \alpha = \emptyset\} \cup (\mathcal{N}(a) \setminus (\mathcal{N}(b) \cup \{b\})).$$

En permutant  $a$  et  $b$ , on définit  $Z_b(a)$  symétriquement et on obtient :

$$\mathcal{B}_{K'}(c) = \{c\alpha\beta \mid \alpha \in Z_a(b), \beta \in Z_b(a), \alpha\beta \in K, \forall \tau \subsetneq \alpha\beta, \tau \in \text{Lk}(a) \cup \text{Lk}(b)\}.$$

On peut dériver un algorithme pour calculer  $\mathcal{B}_{K'}(c)$  directement de cette formule. Le pseudo-code de cet algorithme est donné dans l'algorithme 8. Soit  $N_\sigma = \max_{v \in \sigma} \#\mathcal{N}(v)$ . La mise à jour de notre structure de données après une contraction d'arête a une complexité qui augmente avec le nombre et la dimension des bloqueurs du voisinage de  $a$  et  $b$ . En supposant qu'il n'y a pas de bloqueurs dans le voisinage de  $a$  et de  $b$ , cette mise à jour peut être faite efficacement en  $O(N_a N_b \log N_{\mathcal{N}(a)})$ .

## D.4 Tester si un complexe est un cône

L'algorithme 9 a pour complexité  $O(\#\text{Vert}(L))$ .

## D.5 Tester la condition du *link*

L'algorithme 10 a comme complexité  $O(B_a \log Q_a)$  si l'ensemble des bloqueurs passant par  $a$  vérifie  $\mathcal{B}(a) \neq \emptyset$  et  $O(1)$  autrement.

---

**Algorithme 8** Calcule les bloqueurs  $B$  passant par  $c$  après une contraction d'arête  $ab \mapsto c$ .

---

```

 $B \leftarrow \emptyset$ ;  $L_a \leftarrow \text{Lk}(a)$ ;  $L_b \leftarrow \text{Lk}(b)$ 
for all  $\alpha \in Z_a(b)$  do
  for all  $\beta \in Z_b(a)$  do
    if  $\alpha\beta \in K$  then
       $\text{newblocker} \leftarrow \text{true}$ 
      for all  $\tau \subsetneq \alpha\beta$  dont la dimension est inférieure de 1 à  $\alpha\beta$  do
        if  $\tau \notin L_a$  and  $\tau \notin L_b$  then
           $\text{newblocker} \leftarrow \text{false}$ ; break
      if  $\text{newblocker}$  then  $B \leftarrow B \cup \{c\alpha\beta\}$  end if

```

---

**Algorithme 9** Renvoie vrai si et seulement si le complexe  $L$  est un cône.

---

```

for all  $x \in \text{Vert}(L)$  do
  if  $\mathcal{B}_L(x) = \emptyset$  then
    if  $\#\mathcal{N}_L(x) = \#\text{Vert}(L) - 1$  then
      return true
return false

```

---

**Algorithme 10** Renvoie vrai si et seulement si  $ab$  vérifie la condition du *link*  $\text{Lk}(ab) = \text{Lk}(a) \cap \text{Lk}(b)$ .

---

```

for all  $\sigma \in \mathcal{B}(a)$  do
  if  $b \in \sigma$  then return false end if
return true

```

---

## D.6 Enumérer les simplexes du $k$ -squelette

Nous supposons la donnée d'un complexe simplicial  $K$  représenté dans la structure squelette/bloqueur à partir duquel on cherche à retrouver l'ensemble des simplexes du  $k$ -squelette du complexe. L'algorithme incrémental introduit par Zomorodian attribue un ordre arbitraire aux sommets du complexe  $K$ . Il ajoute ensuite chaque sommet incrémentalement. A chaque fois qu'un sommet  $v$  est ajouté, il ajoute toutes les cofaces de  $v$  pour lesquelles le sommet  $v$  est maximum selon l'ordre des sommets. Nous notons  $\mathcal{N}^-(u) = \{v \in \text{Vert}(K) \mid u > v, uv \in \text{Edges}(K)\}$  les voisins de  $u$  qui lui sont inférieurs avec l'ordre qu'on a fixé. Pour un simplexe  $\tau \in K$ , nous notons  $\mathcal{N}^-(\tau) = \bigcap_{x \in \text{Vert}(\tau)} \mathcal{N}^-(x)$  l'ensemble des voisins inférieurs aux sommets de  $\tau$ . L'algorithme 11 suivant est similaire à celui proposé par Zomorodian qui permet de calculer le  $k$ -squelette d'un complexe dans notre structure de données.

Les deux modifications par rapport à l'algorithme original sont l'affectation  $M \leftarrow N \cap \mathcal{N}^-(v) \cap \text{Vert}(\text{Lk}_K(\sigma))$  qui remplace l'affectation  $M \leftarrow N \cap \mathcal{N}^-(v)$  à la ligne 8 de ADD-COFACES ainsi que l'ajout du test  $\sigma \notin \text{Blockers}_1(K)$  à la ligne 7. Lorsque le complexe est un *flag-complexe*, il est clair que l'Algorithme 11 a le même comportement que l'algorithme INCREMENTAL-VR puisque qu'on a alors  $N \cap \mathcal{N}^-(v) = N \cap \mathcal{N}^-(v) \cap \text{Vert}(\text{Lk}_K(\sigma))$  et la

**Algorithme 11** INCREMENTAL-NEARLY-FLAG( $K, k$ )

---

```

 $S \leftarrow \emptyset$ 
for all  $u \in \text{Vert}(K)$  do
   $N = \mathcal{N}^-(u)$ 
  ADD-COFACES( $K, k, \{u\}, N, S$ )
return  $S$ 

```

---

**Algorithme 12** ADD-COFACES( $K, k, \tau, N, S$ )

---

```

1:  $S \leftarrow S \cup \{\tau\}$ 
2: if  $\dim(\tau) \geq k$  then
3:   return
4: else
5:   for all  $v \in N$  do
6:      $\sigma \leftarrow \tau \cup \{v\}$ 
7:     if  $\sigma \notin \text{Blockers}_1(K)$  then
8:        $M \leftarrow N \cap \mathcal{N}^-(v) \cap \text{Vert}(\text{Lk}_K(\sigma))$ 
9:       ADD-COFACES( $K, k, \sigma, M, S$ )

```

---

condition du test  $\sigma \notin \text{Blockers}_1(K)$  est toujours vérifiée. Lorsque le complexe n'est pas un *flag-complexe* (i.e. lorsque le complexe comporte des bloqueurs) on peut modifier légèrement le théorème de Zomorodian et sa preuve pour obtenir le résultat suivant.

**Théorème D.1.** *INCREMENTAL-NEARLY-FLAG( $K, k$ ) renvoie le  $k$ -squelette de  $K$*

*Démonstration.* Au début de chaque itération dans la boucle de INCREMENTAL-NEARLY-FLAG,  $S$  contient tous les simplexes du  $k$ -squelette dont le sommet maximal (selon l'ordre maximal que l'on a fixé sur les sommets) précède le sommet courant  $u$ . Cet invariant est maintenu dans le corps de la boucle : ADD-COFACES ajoute tous les simplexes du  $k$ -squelette dont le sommet maximal est  $u$ . Cette procédure récursive requiert pour ses entrées les deux conditions suivantes :

1.  $\tau$  est un simplexe du  $k$ -squelette
2.  $N$  est l'ensemble des voisins inférieurs des sommets  $\tau$  i.e.  $N = \mathcal{N}^-(\tau)$

Ces conditions sont initialement remplies puisque INCREMENTAL-NEARLY-FLAG appelle cette procédure en lui passant pour argument un sommet et ses voisins inférieurs qui appartiennent tous à  $S$ . Le corps de ADD-COFACES maintient cet invariant. A la ligne 1, cette procédure ajoute  $\tau$  à  $S$  qui est simplexe du squelette d'après notre invariant. La ligne 2 est le cas de base pour la récursion. Au départ la boucle ligne 4,  $\tau$  est un simplexe de  $S$  et chaque sommet  $v \in N$  est un voisin des sommets de  $\tau$ . Montrons que toutes les arêtes de  $\sigma$  appartiennent alors à  $\text{Edges}(K)$ . Soit  $u \neq w$  deux sommets distincts de  $\sigma$ . Si  $u$  et  $w$  appartiennent tous les deux à  $\tau$  alors  $uw \in \text{Edges}(K)$  car  $\tau$  est un simplexe de  $S$ . Sinon, supposons sans perte de généralité que seul  $w \in \tau$  et que  $u = v$ . Puisque  $v \in N$ , l'arête  $uw$  appartient à  $\text{Edges}(K)$ . Ainsi toutes les arêtes de  $\sigma$  appartiennent à  $\text{Edges}(K)$ . Comme de plus  $\sigma$  n'est pas un bloqueur de  $K$  d'après le test ligne 7, on en déduit que  $\sigma = \tau \cup \{v\}$  est également un

simplexe de  $S$ . De plus,  $\dim(\sigma) \leq k$  d'après notre cas de base. Ainsi, nous calculons les voisins inférieurs de  $\sigma$  dans  $M$  à la ligne 6 et nous ajoutons les cofaces de  $\sigma$  à la ligne 7 *via* un appel récursif à ADD-COFACES. Notons que les deux invariants sont maintenus et la procédure ajoute récursivement toutes les cofaces de dimension inférieure ou égale à  $k$  de son entrée  $\tau$  comme demandé.

□

## D.7 Pseudo-codes pour réduire un complexe en un point

Afin de décrire l'algorithme REDUCTIBLE\_BY\_SWEEP qui reproduit la séquence de collapse de la preuve du théorème 5.3, nous introduisons une fonction INITIALIZE\_SWEEP( $L, o$ ) qui renvoie une liste de priorité de sommets et d'arêtes. A chaque sommet  $a$  de  $L$  différent de  $o$ , on associe une priorité  $\varphi(a) = \|o - a\| - \alpha$  et à chaque arête  $ab$  de  $L$ , on associe une priorité  $\varphi(ab) = d(o, B(a, \alpha) \cap B(b, \alpha))$ . En utilisant la priorité  $\varphi$ , INITIALIZE\_SWEEP( $L, o$ ) trie la liste des simplexes contenant tous les sommets différents de  $o$  et toutes les arêtes  $ab$  telles que  $\varphi(ab) < \min(\varphi(a), \varphi(b))$ .

---

**Algorithme 13** REDUCTIBLE\_BY\_SWEEP(Complexe simplicial  $L$ )

---

```

if  $L = \emptyset$  then
  return false
 $o =$  un sommet de  $L$ 
 $Q.$ INITIALIZE_SWEEP( $L, o$ ) {Voir l'explication ci-dessus pour cette fonction.}
while  $|Q| \neq \emptyset$  do
  Extraire de  $Q$  le simplexe  $\sigma$  de priorité maximale
  if not ( $\text{Lk}_L(\sigma)$  est un cône) then
    return false
  else
     $L = L \setminus \text{St}_L(\sigma)$ 
return true

```

---

---

**Algorithme 14** REDUCTIBLE\_BY\_COMPLETION( Flag-complexe  $L$ )

---

```

if  $L = \emptyset$  then
  return false
if  $L$  est un cône then
  return true
 $ab =$  Plus petite arête non présente dans  $L$  connectant deux sommets de  $L$ 
 $L = \text{Flag}(\text{Sk}^{(1)}(L) \cup \{ab\})$ 
if REDUCTIBLE_BY_COMPLETION( $\text{Lk}_L(ab)$ ) then
  return REDUCTIBLE_BY_COMPLETION( $L$ )
else
  return false

```

---



---

**Algorithme 15** REDUCTIBLE\_BY\_EDGE\_CONTRACTIONS(Complexe simplicial  $L$ )

---

```

if  $L = \emptyset$  then
  return false
 $Q = \text{Edges}(L)$ 
while  $Q \neq \emptyset$  do
   $ab =$  Plus petite arête de  $Q$ 
  Extraire  $ab$  de  $Q$ 
  if  $\text{Lk}_L(a) \cap \text{Lk}_L(b) = \text{Lk}_L(ab)$  then
    Contracter dans  $L$  l'arête  $ab$  en un nouveau sommet  $c = \frac{a+b}{2}$ 
return  $|\text{Vert}(L)| = 1$ 

```

---



# Lemmes géométriques du chapitre 5

Rappelons que  $B_p$  désigne la boule unité fermée centrée en  $p$ . Nous énonçons et prouvons ici plusieurs lemmes de ce chapitre.

**Lemme E.1.** *Soit  $o \in \mathbb{R}^d$  et un ensemble fini  $\sigma \subset \mathbb{R}^d$  tel que  $\bigcap_{p \in \sigma} B_p \neq \emptyset$ . Supposons que  $d(o, \bigcap_{p \in \sigma} B_p) = t$  pour  $t \in \mathbb{R}$  et soit  $x$  le point unique de  $\bigcap_{p \in \sigma} B_p$  dont la distance à  $o$  est  $t$ . Supposons que  $o \neq x$ . Si  $x$  se situe dans l'intérieur de  $B_u$  pour  $u \in \sigma$ , alors  $x$  est aussi le point de  $\bigcap_{p \in \sigma \setminus \{u\}} B_p$  le plus proche de  $o$ .*

*Démonstration.* Supposons que  $x$  appartienne à l'intérieur de  $B_u$  pour  $u \in \sigma$ . Puisque  $o \neq x$ , on ne peut pas avoir  $\sigma = \{u\}$  et ainsi  $\sigma' = \sigma \setminus \{u\}$  est non vide. Montrons que  $x$  est le point de  $\bigcap_{p \in \sigma'} B_p$  le plus proche de  $o$ . Supposons par contradiction qu'il existe un point  $x'$  dans  $\bigcap_{p \in \sigma'} B_p$  qui soit plus proche de  $o$  que  $x$ . Puisque la fonction  $x \mapsto \|x - o\|$  est convexe, la distance à  $o$  décroît sur le segment  $[xx']$  dans un voisinage de  $x$  et puisque ce segment, dans un voisinage de  $x$ , est contenu dans  $\bigcap_{p \in \sigma} B_p$  cela contredit le fait que  $x$  soit le plus proche point de  $o$  dans  $\bigcap_{p \in \sigma} B_p$ .  $\square$

Afin de prouver le lemme 5.1, nous introduisons le lemme E.2 qui énonce des propriétés sur trois disques s'intersectant dans le plan deux à deux mais n'ayant pas d'intersection commune. Il nous permet de déduire certaines propriétés sur la façon dont s'intersectent quatre boules dans  $\mathbb{R}^d$  ce qui nous permet de démontrer le lemme 5.1. Comme avant,  $B_x$  dénote la boule unité fermée centrée en  $x$ .

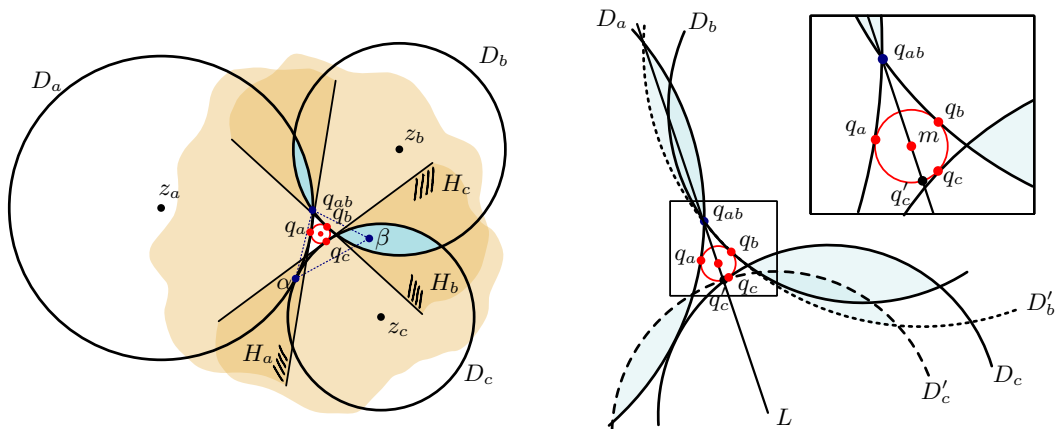


FIGURE E.1 – Notation pour la preuve du lemme E.2.



**Lemme E.2.** Soient  $D_a, D_b$  et  $D_c$  trois disques de rayons inférieurs ou égaux à 1 et tels que leurs intersections deux à deux soient non vides tandis que l'intersection commune des trois disques est vide. Soit  $q_{ab}$  le point de  $D_a \cap D_b$  le plus proche du centre de  $D_c$ . Il existe un point  $q_c \in D_c$  tel que :

- pour tout point  $\alpha \in D_a \cap D_c$  et  $\beta \in D_b \cap D_c$ , le point  $q_c$  est dans l'enveloppe convexe de  $\alpha, \beta$  et  $q_{ab}$  ;
- $\|q_c - q_{ab}\| \leq \sqrt{3} - 1$ .

*Preuve du lemme E.2.* Considérons le disque  $D_m$  dont la frontière est tangente aux frontières des trois disques  $D_a, D_b$  et  $D_c$  et dont l'intérieur n'intersecte aucun des trois disques  $D_a, D_b$  et  $D_c$ . Pour  $x \in \{a, b, c\}$ , les deux disques  $D_x$  et  $D_m$  s'intersectent sur le point  $q_x$  ; voir la partie gauche de la figure E.1. Soient  $\alpha \in D_a \cap D_c$  et  $\beta \in D_b \cap D_c$ . Le point  $q_c$  appartient à l'enveloppe convexe de  $\alpha, \beta$  et  $q_{ab}$ . En effet, pour  $x \in \{a, b, c\}$ , soit  $H_x$  le demi-plan contenant  $D_x$  et évitant l'intérieur de  $D_m$ . Nous avons  $\alpha \in H_a \cap H_c, \beta \in H_b \cap H_c$ , et  $q_{ab} \in H_a \cap H_b$ . Le triangle  $\alpha\beta q_{ab}$  couvre la fermeture de  $\mathbb{R}^2 \setminus (H_a \cup H_b \cup H_c)$  et ainsi  $q_c$  appartient à l'enveloppe convexe de  $\alpha, \beta$  et  $q_{ab}$ .

Montrons maintenant que  $\|q_c - q_{ab}\| \leq \sqrt{3} - 1$ . Pour  $x \in \{a, b, c\}$ , nous notons le centre de  $D_x$  par  $z_x$  et son rayon par  $r_x$ . Nous allons transformer les trois disques  $D_a, D_b$  et  $D_c$  de telle sorte qu'après la transformation :

- (i) les trois disques s'intersectent deux à deux mais n'ont pas d'intersection commune ;
- (ii) la distance entre  $q_c$  et  $q_{ab}$  ne diminue pas lors de la transformation ;
- (iii)  $r_x \leq 1$  pour  $x \in \{a, b, c\}$  ;
- (iv) les centres  $z_a, z_b$ , et  $z_c$  forment un triangle équilatéral de diamètre deux.

Soit  $q'_c$  le point le plus loin de  $q_{ab}$  sur le bord de  $D_m$  ; voir la partie droite de la figure E.1. Clairement,  $\|q'_c - q_{ab}\| \geq \|q_c - q_{ab}\|$ . Les deux points tangents  $q_a$  et  $q_b$  décomposent le bord de  $D_m$  en deux arcs et il est facile de voir que l'un d'entre eux contient  $q_c$  et  $q'_c$ . Considérons le disque  $D'_c$  obtenu par rotation de  $D_c$  autour de  $m$  jusqu'à ce qu'il rencontre  $q'_c$ . En effectuant cette rotation, le disque reste en contact avec au moins un des deux disques  $D_a$  ou  $D_b$ . Sans perte de généralité, on peut supposer que  $D_a \cap D'_c \neq \emptyset$ . Soit  $L$  la droite passant par  $q_{ab}, q'_c$  et  $m$  et soit  $D'_b$  le symétrique de  $D_a$  par rapport à  $L$ . Nous avons  $D'_b \cap D'_c \neq \emptyset$ . Les deux bords de  $D_a$  et  $D'_b$  s'intersectent en deux points, l'un étant  $q_{ab}$ . En remplaçant  $D_b$  par  $D'_b$  et  $D_c$  par  $D'_c$ , il est facile de voir que les trois disques  $D_a, D_b$  et  $D_c$  vérifient maintenant (i), (ii) et (iii) et leurs centres forment un triangle isocèle. On peut donc transformer les trois disques de telle manière qu'après la transformation, ils vérifient en plus (iv). Lorsque c'est le cas, on a clairement que  $\|q_c - q_{ab}\| = \sqrt{3} - 1$ .  $\square$

*Preuve du lemme 5.1.* Considérons  $c$  tel que  $B_c \cap B_a \cap B(o, t) \neq \emptyset$  et  $B_c \cap B_b \cap B(o, t) \neq \emptyset$ . Si  $x \in B_c$ , alors le lemme est trivialement vrai puisque  $x \in B(x, \sqrt{3} - 1) \cap B(o, t)$ . Supposons que  $x \notin B_c$ . Prenons  $\alpha \in B_c \cap B_a \cap B(o, t)$  et  $\beta \in B_c \cap B_b \cap B(o, t)$  et considérons un plan  $\Pi$  contenant les trois points  $x, \alpha$  et  $\beta$ . Ce plan intersecte les quatre boules  $B_a, B_b, B_c$  et  $B(o, t)$  en quatre disques que nous notons respectivement  $D_a, D_b, D_c$  et  $D_o$  ; voir la partie gauche de

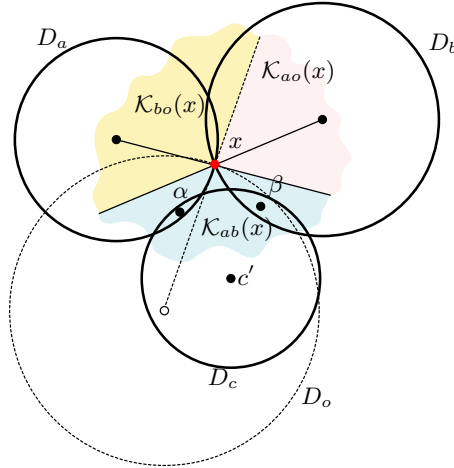


FIGURE E.2 – Notation pour la preuve du lemme 5.1.

la figure E.2. Les trois disques  $D_a$ ,  $D_b$  et  $D_o$  ont une intersection non vide réduite à un point  $x$ . Nous avons  $\alpha \in D_c \cap D_a \cap D_o \neq \emptyset$ ,  $\beta \in D_c \cap D_b \cap D_o \neq \emptyset$  et  $x \notin D_c$ .

Soit  $c'$  le centre de  $D_c$ . Prouvons que  $x$  est le point de  $D_a \cap D_b$  le plus proche de  $c'$ . Pour cela, définissons le cône extérieur de  $D_a \cap D_o$  en  $x$  comme l'ensemble de points :

$$\mathcal{K}_{ao}(x) = \{y \in \Pi \mid \forall z \in D_a \cap D_o, \langle y - x, z - x \rangle \leq 0\}.$$

De manière équivalente,  $\mathcal{K}_{ao}(x)$  est l'ensemble des points dont la distance à  $x$  est inférieure ou égale à la distance de n'importe quel autre point de  $D_a \cap D_o$ . Comme  $x \notin D_c$  et  $\alpha \in D_c$ , on a que  $\|c' - \alpha\| \leq \|c' - x\|$ . Ainsi,  $\alpha$  est un point dans l'intersection  $D_a \cap D_o$  plus proche de  $c'$  que de  $x$ . De manière équivalente,  $c' \notin \mathcal{K}_{ao}(x)$ . Similairement,  $c' \notin \mathcal{K}_{bo}(x)$ . Puisque  $\mathcal{K}_{ao}(x) \cup \mathcal{K}_{bo}(x) \cup \mathcal{K}_{ab}(x) = \Pi$ , on en déduit que  $c' \in \mathcal{K}_{ab}(x)$ . Autrement dit,  $x$  est le point de  $D_a \cap D_b$  le plus proche de  $c'$ .

Ainsi, le lemme E.2 s'applique et montre l'existence d'un point  $x' \in D_c$  dans l'enveloppe convexe de  $\alpha$ ,  $\beta$  et  $x$  tel que  $x' \in B(x, \sqrt{3}-1)$ . Puisque les trois points  $\alpha$ ,  $\beta$  et  $x$  appartiennent à  $B(o, t)$ , il s'ensuit que  $x' \in B(o, t)$ , ce qui montre le résultat.  $\square$



# Reconstruction de forme en grandes dimensions

David Salinas

---

**Résumé** — Dans cette thèse, nous cherchons à reconstruire une approximation d’une variété connue seulement à partir d’un nuage de points de grande dimension l’échantillonnant. Nous nous efforçons de trouver des méthodes de reconstructions efficaces et produisant des approximations ayant la même topologie que la variété échantillonnée. Une attention particulière est consacrée aux complexes de cliques et particulièrement aux complexes de Rips. Nous montrons que le complexe de Rips capture la topologie d’une variété échantillonnée en supposant de bonnes conditions d’échantillonnage. En tirant avantage de la compacité des complexes de cliques qui peuvent être représentés de manière compacte avec un graphe, nous présentons une structure de données appelée *squelette/bloqueurs* pour complexes simpliciaux. Nous étudions ensuite deux opérations de simplifications, la contraction d’arête et le collapse simplicial, qui s’avèrent utiles pour réduire un complexe simplicial sans en changer sa topologie.

**Mots clés** : complexe de Rips, grandes dimensions, collapse, contraction d’arête, structure de données, complexes simpliciaux, topologie algorithmique.

---

## Shape reconstruction in high dimensions

---

**Abstract** — In this thesis, we look for methods for reconstructing an approximation of a manifold known only through a high-dimensional point cloud. Especially, we are interested in efficient methods that produce approximations that share the same topology as the sampled manifold. A particular attention is devoted to flag-complexes and more specially to Rips complexes due to their compactedness. We show that the Rips complex shares the topology of a sampled manifold under good sampling conditions. By taking advantage of the compactedness of flag-complexes, we present a data structure for simplicial complexes called *skeleton/blockers*. We then study two simplification operations, the edge contraction and the simplicial collapse, that turn out to be useful for reducing a simplicial complex without changing its topology.

**Keywords** : Vietoris-Rips complex, high dimension, collapse, edge contraction, data structure, simplicial complex, computational topology.

---

