



HAL
open science

Apport à l'analyse des paysages de fitness pour l'optimisation mono-objective et multiobjective

Sébastien Verel

► **To cite this version:**

Sébastien Verel. Apport à l'analyse des paysages de fitness pour l'optimisation mono-objective et multiobjective. Intelligence artificielle [cs.AI]. Université du Littoral Côte d'Opale, 2016. tel-01425127

HAL Id: tel-01425127

<https://hal.science/tel-01425127v1>

Submitted on 3 Jan 2017

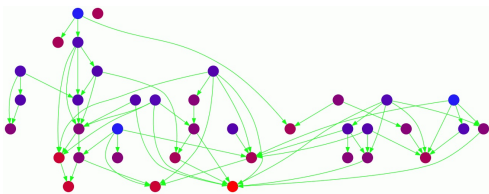
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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.

APPORT À L'ANALYSE DES PAYSAGES DE FITNESS POUR L'OPTIMISATION MONO-OBJECTIVE ET MULTIOBJECTIVE

Thèse d'Habilitation à Diriger des Recherches
Spécialité : Informatique

SÉBASTIEN VEREL



SCIENCE DES SYSTÈMES COMPLEXES
POUR L'OPTIMISATION
PAR MÉTHODES STOCHASTIQUES

Université du Littoral Côte d'Opale

<http://www-lisic.univ-littoral.fr/~verel/>

Habilitation à Diriger des Recherches (HDR), été/automne 2016

Soutenue le 12 décembre 2016 à Calais

Jury composé de :

- D.R. Évelyne LUTTON, INRA-AgroParisTech, *rapporteur*,
- Pr. Adnan YASSINE, Institut Supérieur d'Études Logistiques du Havre, *rapporteur*,
- Pr. Frédéric SAUBION, Université d'Angers, *rapporteur*,
- Pr. Philippe COLLARD, Université Nice Sophia-Antipolis/CNRS, *examineur*,
- Pr. Nouredine MELAB, Université Lille 1/CNRS/inria, *examineur*,
- Pr. Cyril FONLUPT, Université du Littoral Côte d'Opale, *examineur*.

Sébastien Verel : *Apport à l'analyse des paysages de fitness pour l'optimisation mono-objective et multiobjective*, Science des systèmes complexes pour l'optimisation par méthodes stochastiques, © Soutenue le **12 décembre 2016** à Calais.

PAGE WEB :

<http://www-lisic.univ-littoral.fr/~verel/>

COURRIEL :

verel@lisic.univ-littoral.fr

ABSTRACT

This manuscript presents some of my research activities I have conducted as assistant researcher ("maitre de conférences") at the university Nice Sophia Antipolis and at the university Littoral Côte d'Opale since 2006. The synthesis of the presented works is in the field of single- and multi-objective combinatorial optimization using stochastic algorithms (meta-heuristics such as evolutionary algorithms, local search algorithms, etc.). In particular, my research work focuses on the model of fitness landscapes, originally designed in the field of complex systems theory, which allows to study the dynamics of optimization algorithms. The main motivation of this work is to understand the relationship between an optimization algorithm and the problem to solve in order to explain and to predict the algorithms performance, and to design new more efficient algorithms based on this understanding.

This document is divided in two main parts. The first one is on the *local optima network* which is a new model of fitness landscapes for the single-objective optimization. New properties of optimization problems are shown, as well as the prediction capacities of performance based on the estimation of these features. The second one is on the *multi-objective optimization*. The properties of such fitness landscapes are defined, and precisely analyzed with respect to algorithms performance. New multi-objective optimization methods are also proposed. Finally, this document concludes with broad perspectives for the research domain of fitness landscapes and more generally in the field of stochastic optimization.

RÉSUMÉ

Ce mémoire présente quelques-unes de mes activités de recherche que j'ai conduites depuis 2006 en tant que maître de conférences à l'Université Nice Sophia Antipolis puis à l'Université du Littoral Côte d'Opale. La synthèse des travaux présentés s'inscrit dans le domaine de l'optimisation mono-objective et multiobjective de problèmes combinatoires par des algorithmes stochastiques du type métaheuristique (algorithmes évolutionnaires, recherches locales, etc.). En particulier, mes travaux portent sur le modèle des paysages de fitness, modèle issu des sciences de la complexité, qui permet étudier la dynamique des algorithmes d'optimisation. La principale motivation de ces travaux est de comprendre la relation entre un algorithme d'optimisation et le problème à résoudre afin d'en expliquer et d'en prédire les performances et de concevoir de nouveaux algorithmes plus efficaces à partir de cette compréhension.

Ce manuscrit se compose en deux parties principales. L'une porte sur le *réseau des optima locaux* qui est un nouveau modèle pour les paysages de fitness en optimisation mono-objective. De nouvelles propriétés des problèmes d'optimisation sont montrées, ainsi que les capacités de prédiction de performance à partir de l'estimation de ces propriétés. L'autre partie porte sur l'*optimisation multiobjective*. Les propriétés des paysages de fitness sont définies et précisément analysées en rapport avec les performances des algorithmes. De nouvelles méthodes d'optimisation multiobjective sont également proposées. Enfin, ce mémoire se termine par des perspectives pour le domaine des paysages de fitness et plus généralement en optimisation stochastique.

REMERCIEMENTS

Tout d'abord, je tiens à remercier les rapporteurs et les membres du jury qui ont accepté de lire ce manuscrit et d'assister à la soutenance. Qu'Évelyne Lutton, Frédéric Saubion, Adnan Yassine, Philippe Collard, Nouredine Melab et Cyril Fonlupt soient ici remerciés. Leurs questions et leurs remarques enthousiastes vont m'aider et m'encouragent à poursuivre mes travaux.

À la suite de ce travail il est difficile de ne pas porter un regard rétrospectif du chemin parcouru. En premier lieu au sens propre, ma chère équipe de Nice, Philippe (de nouveau lui!), Manuel et Michael qui m'ont donné tout ce qu'il fallait pour réussir et me construire dans le bon sens : l'attention, l'écoute, les conseils, les idées, les critiques introspectives, le recul et le plus important, l'amitié.

Les doctorants tiennent une place importante dans ce travail. Je remercie David Simoncini, et plus particulièrement Fabio Daolio avec qui j'ai vécu bien plus qu'une thèse. Ses qualités à la pointe scientifique et sa gentillesse (sans oublier quelques prouesses culinaires) s'impriment dans ce document. Merci également aux docteurs en devenir que sont Christopher Jankee et Mathieu Muniglia.

Le deuxième lieu où j'ai pu énormément développer ce travail est Lille. Je remercie l'équipe DOLPHIN et tout particulièrement Clarisse et Laetitia grâce à qui j'ai pu entrer dans le monde nordique (de la France) et surtout dans le monde du multiobjectif de la meilleure façon possible. Comment montrer la présence d'Arnaud et de Bilel au plus profond de ces travaux? Je pourrais dresser la liste de nos travaux communs, de nos innombrables discussions communes, de nos multiples lieux partagés, de nos affinités communes, de nos projets communs qui se développent et se développeront, de nos rires et nos selfies... Je les remercie de tout mon cœur pour tout cela et plus encore.

Le troisième lieu marquant de mon parcours est Calais. Je tiens à remercier toute l'équipe OSMOSE qui a su accueillir ce «mutant» que j'étais. Je m'y suis rapidement senti très bien parmi ses membres et les sciences qu'ils élaborent. Je remercie le sympathique laboratoire LISIC où les projets peuvent se développer avec le riche environnement de ce littoral.

Tout au long de ces lieux, les collaborations n'ont cessé de prendre une place importante dans mon parcours. Je remercie toutes les personnes avec qui j'ai pu collaborer. Merci à Marco Tomassini dont la fougue et la jeunesse marquent toujours mes travaux. Merci à Leonardo Vanneshi avec qui nous avons partagé en outre les idées de Marco. Évidemment, je remercie Gabriela Ochoa dont la patte a permis à ces travaux de prendre corps. À un gros quart de tour de Terre d'ici, je remercie humblement le professeur Kiyoshi Tanaka et Hernan Aguirre, des personnes précieuses avec qui je partage maintenant beaucoup de sentiments qui bizarrement ont pris naissance autour d'un problème d'optimisation commun et une façon de le regarder.

Hors du cadre de ce travail, mais au combien présente, je tiens à remercier Carole qui illumine l'autre face cachée de ce manuscrit.

Les prochains travaux à mener sont toujours ceux qui piquent le plus ma curiosité. En cela, je remercie tout ceux qui tout au long de mon parcours m'ont apporté de la motivation et ont enrichi mon envie de connaître.

Enfin, une spéciale dédicace, comme on peut le dire parfois, aux Coffee Makers qui m'ont accueilli dans leur lieu avec leur bon café tout au long de la rédaction de ce manuscrit.



TABLE DES MATIÈRES

1	PAYSAGE DE FITNESS ET OPTIMISATION BOITE NOIRE	1
1.1	Optimisation par méthodes stochastiques	1
1.1.1	Optimisation combinatoire mono-objective	2
1.1.2	Optimisation combinatoire multiobjective	3
1.1.3	Méthodes stochastiques d'optimisation mono-objective	4
1.1.4	Méthodes stochastiques d'optimisation multiobjective	6
1.2	Paysage de fitness	7
1.2.1	Origine et définition	7
1.2.2	Buts de l'analyse des paysages de fitness	9
1.3	Aperçu du document	13
2	RÉSEAU DES OPTIMA LOCAUX	15
2.1	Introduction	16
2.1.1	Analyse de paysage de fitness	16
2.1.2	Un point de vue sur l'analyse de paysage de fitness	16
2.1.3	Science des réseaux complexes	19
2.2	Modélisation des paysages de fitness par le réseau des optima locaux	20
2.2.1	Réseau des optima locaux	21
2.2.2	Nouvelles propriétés des paysages de fitness	23
2.2.3	Analyses basées sur le réseau des optima locaux	28
2.2.4	Explication de performance des métaheuristiques	31
2.2.5	Prédiction de performance pour les problèmes de grande dimension	36
2.3	Conclusion et perspectives	42
3	ESPACES DE RECHERCHE DE L'OPTIMISATION MULTIOBJECTIVE	45
3.1	Introduction	46
3.1.1	Source de difficulté en optimisation multiobjective	46
3.1.2	Paysage de fitness et optimisation multiobjective	47
3.2	Structure des espaces de recherche de l'optimisation multiobjective	48
3.2.1	Portrait des problèmes d'optimisation multiobjective	48
3.2.2	Explication des performances des heuristiques	53
3.2.3	Prédiction de performance pour les problèmes de grande dimension	58
3.2.4	Paysage de fitness multiobjectif basé sur les ensembles	63
3.2.5	Approches scalaires par décomposition	68
3.3	Conclusion et perspectives	77
4	PERSPECTIVES GÉNÉRALES	81
A	NOTICE BIOGRAPHIQUE	85
A.1	Informations personnelles	85
A.2	Situation professionnelle	85
A.3	Cursus Universitaire	86
A.4	Activités scientifiques	86
A.5	Encadrement doctoral et scientifique	86
A.5.1	Co-encadrement doctoral en cours	86
A.5.2	Co-encadrement doctoral	86
A.5.3	Co-encadrement Master 2	87
A.5.4	Encadrement stage doctoral	87
A.5.5	Co-encadrement Post-docdoral	87

A.6	Diffusion scientifique	88
A.7	Rapporteur de thèse et instance d'évaluation	89
A.8	Administration et animation de la science	90
A.9	Activités d'enseignement	92
B	BIBLIOGRAPHIE PERSONNELLE CLASSÉE DEPUIS LA THÈSE	95
B.1	Livre	95
B.2	Chapitre de livre	95
B.3	Articles de journaux internationaux avec comité de lecture	95
B.4	Articles de conférences internationales avec comité de lecture	96
C	BIBLIOGRAPHIE	103

1

PAYSAGE DE FITNESS ET OPTIMISATION BOITE NOIRE

CONTENU

1.1	Optimisation par méthodes stochastiques	1
1.1.1	Optimisation combinatoire mono-objective	2
1.1.2	Optimisation combinatoire multiobjective	3
1.1.3	Méthodes stochastiques d'optimisation mono-objective	4
1.1.4	Méthodes stochastiques d'optimisation multiobjective	6
1.2	Paysage de fitness	7
1.2.1	Origine et définition	7
1.2.2	Buts de l'analyse des paysages de fitness	9
1.3	Aperçu du document	13

RÉSUMÉ

Ce chapitre introductif pose le cadre préalable de mes recherches à travers principalement une revue bibliographique du domaine, il expose les points suivants :

- une introduction générale du cadre de l'optimisation combinatoire mono-objective et multiobjective dite *boite noire*, en particulier en définissant l'optimalité au sens Pareto,
- un aperçu des méthodes d'optimisation mono-objective,
- les principales approches en optimisation multiobjective avec des algorithmes représentatifs,
- une introduction au modèle de paysage de fitness,
- une proposition de classification des finalités de l'analyse des paysages de fitness.

1.1 OPTIMISATION PAR MÉTHODES STOCHASTIQUES

Intuitivement, la résolution d'un problème d'optimisation consiste à trouver le ou les meilleures solutions qui maximisent (ou de manière équivalente minimisent) un, dans le cas de l'optimisation mono-objective, ou plusieurs, dans le cas de l'optimisation multiobjective, critères donnés.

Si aucune hypothèse et information supplémentaire sur le ou les critères ne sont soit connues, fournies, ou utilisées pendant la résolution, alors on parle d'optimisation *boite noire*. En optimisation *boite noire*, une méthode de résolution ne peut alors utiliser aucune information supplémentaire à part la connaissance des solutions potentielles et de la valeur du ou des critères qui est, en quelque sorte, donnée par un oracle. Par opposition, l'optimisation *boite grise* suppose connaître au moins partiellement la définition du problème d'optimisation. La modélisation d'un problème en un problème d'optimisation *boite noire* est très souple et n'exige de la part du modélisateur que la définition de l'ensemble des

solutions possibles au problème et du ou des critères qui relatent la valeur de ces solutions potentielles. Dans tout ce manuscrit, nous adoptons le point de vue de l'optimisation boîte noire aussi bien pour les méthodes de résolution que pour les méthodes d'analyse de ces problèmes.

De même, tous les travaux de ce manuscrit portent sur l'*optimisation combinatoire*, aussi appelée optimisation discrète, c'est-à-dire lorsqu'il s'agit de trouver une ou plusieurs solutions dans un ensemble discret qui est très souvent de cardinalité finie.

Dans ce contexte, les méthodes de résolution consistent à énumérer et tester un certain nombre de solutions potentielles au problème. Il existe de nombreuses façons d'énumérer ces solutions. Aussi après avoir rappelé les définitions de l'optimisation mono- et multiobjective, cette section introduit les *méthodes stochastiques* de résolution et plus particulièrement les métaheuristiques.

1.1.1 Optimisation combinatoire mono-objective

Un problème d'optimisation mono-objectif est un couple (\mathcal{X}, f) où \mathcal{X} est un ensemble de solution potentielle au problème et $f : \mathcal{X} \rightarrow \mathbb{R}$ est une fonction qui associe à chaque solution potentielle un nombre réel. L'ensemble \mathcal{X} est appelé *espace de recherche* et la fonction f *fonction d'évaluation* ou encore *fonction de fitness*. Lorsque l'ensemble \mathcal{X} est discret, le problème d'optimisation est dit *combinatoire*.

Résoudre un problème d'optimisation mono-objectif consiste à trouver la ou les meilleures solutions selon l'unique fonction d'évaluation f , c'est-à-dire pour un problème de maximisation, déterminer l'ensemble optimal \mathcal{X}^* :

$$\mathcal{X}^* = \operatorname{argmax}_{x \in \mathcal{X}} f(x)$$

Dans la suite, sauf mention contraire, les problèmes d'optimisation sont des problèmes de maximisation. Pour un problème de minimisation, on pourra de manière équivalente maximiser la fonction opposée $-f$. La recherche de l'ensemble optimal est parfois trop «stricte» et souvent une «bonne» solution au problème peut être attendue, c'est-à-dire une solution proche de l'optimalité, la qualité de l'approximation désirée étant jugée par l'utilisateur final.

La formulation d'un problème en un problème d'optimisation est en soi un processus délicat, appelé *modélisation* dans ce contexte, qui nécessite à la fois une bonne connaissance du problème à résoudre et des méthodes de résolution, ou au minimum de leur formalisme. Mes travaux ne s'intéressent pas à ce domaine de recherche en soi. Dans ce document, nous nous plaçons dans le cadre de l'optimisation *boîte noire* qui ne requiert aucune connaissance particulière sur le problème d'optimisation, seule la connaissance de $x \in \mathcal{X}$ et $f(x)$ sont *a priori* utilisées par les méthodes de résolution et de leur étude. Ce cadre offre une grande liberté dans la modélisation elle-même. À la fois, les problèmes NP-difficiles (SAT, TSP, etc.) peuvent être abordés dans le cadre boîte noire en prenant en compte les propriétés particulières de ces problèmes¹. Surtout l'optimisation boîte noire permet aussi de modéliser les problèmes où la fonction d'évaluation est issue d'un calcul extérieur comme un simulateur avec de nombreuses applications en conception industrielle [YTS10], contrôle de systèmes complexes [Sch+13], etc. Néanmoins, même si l'optimisation boîte noire n'exige aucune connaissance supplémentaire sur le problème à résoudre, très souvent les méthodes d'optimisation formulent des hypothèses implicites. En effet, intuitivement, une permutation aléatoire des valeurs de fitness des solutions d'un espace de recherche \mathcal{X} conduit à définir un problème d'optimisation aléatoire et sans structure où il y a peu de chance de concevoir une méthode plus efficace qu'une méthode de recherche aléatoire. Tout l'art des méthodes

1. Pour certains problèmes comme SAT, QAP, etc., on notera tout de même qu'afin de réduire la complexité du calcul de la fonction d'évaluation, on utilise la forme additive de ces fonctions d'évaluation (technique d'évaluation incrémentale).

d'optimisation consiste à exploiter la structure conjointe de l'espace de recherche \mathcal{X} et de la fonction d'évaluation f .

1.1.2 Optimisation combinatoire multiobjective

L'optimisation multiobjective propose une méthode de modélisation différente à l'optimisation mono-objective. Il n'est pas rare qu'un problème réel comporte en réalité plusieurs critères à satisfaire. Par exemple, on peut imaginer qu'il faut améliorer l'efficacité aérodynamique et le poids des ailes d'un avion et cela sûrement conjointement avec l'efficacité d'une autre partie de l'avion; ou encore diminuer le temps de parcours des voitures dans le centre d'une ville et même temps celui de la périphérie conjointement à la réduction des émissions polluantes dans la journée et aux heures de pointe, etc. Une modélisation d'un problème réel en problème d'optimisation mono-objective, même boîte noire, devra proposer, *a priori* avant toute résolution, une agrégation pertinente de ces critères en une fonction objectif qui associe un seul nombre réel à chaque solution potentielle. La modélisation en problème d'optimisation multiobjectif offre une souplesse de modélisation supplémentaire en différant le moment où l'on décide de l'importance relative des objectifs.

L'optimisation multiobjective s'inscrit dans un processus de décision. Un décideur modélise le problème en exprimant plusieurs critères, aussi appelés objectifs, souvent contradictoires dont, *a priori*, il ne peut pas ou ne veut pas donner une classification, une hiérarchisation ou une pondération afin de ne pas, par exemple, introduire de biais dans les solutions finales trouvées. Typiquement, les objectifs sont souvent relatifs à des notions de qualité et de coût. Ensuite, la résolution du problème multiobjectif apporte non plus une solution mais un ensemble de solutions qui optimisent conjointement l'ensemble des objectifs. Le décideur peut alors étudier les solutions obtenues *a posteriori* de la résolution, et selon le contexte sélectionner une solution particulière. La modélisation multiobjective apporte une souplesse dans la modélisation, une meilleure connaissance sur le problème par l'analyse *a posteriori* d'un ensemble de solutions possibles, et donc renforce la confiance dans le choix d'une solution particulière. A contrario, la modélisation oblige le décideur à une étape supplémentaire d'analyse et l'optimisation multiobjective est souvent d'une plus grande complexité (temps de résolution plus long par exemple). Toutefois, aussi bien académiquement², que dans de nombreux domaines d'application, dont il est désormais impossible d'établir une liste exhaustive, la modélisation multiobjective devient de plus en plus utilisée.

Formellement, un problème d'optimisation multiobjective se distingue de l'optimisation mono-objective par l'espace d'arrivée de la fonction d'évaluation. Un problème d'optimisation multiobjectif est un couple $(\mathcal{X}, \mathcal{F})$ où \mathcal{X} est l'ensemble des solutions potentielles, l'espace de recherche aussi appelé *espace de décision* dans ce cas, et $\mathcal{F} : \mathcal{X} \rightarrow \mathbb{R}^m$ est la fonction qui associe à chaque solution un vecteur de dimension $m \geq 2$ de nombres réels correspondant aux valeurs des objectifs. Lorsque l'espace \mathcal{X} est discret, l'optimisation est dite combinatoire et l'optimisation peut être aussi boîte noire lorsqu'aucune information supplémentaire n'est fournie lors de la résolution. Les composantes f_i de la fonction \mathcal{F} sont appelées les objectifs ou encore critères, chacun peut être maximisé ou minimisé. Dans la suite, les objectifs seront maximisés. L'image de \mathcal{X} par \mathcal{F} est nommé *l'espace objectif* par opposition à l'espace de décision. Comme pour l'optimisation mono-objective, nous nous intéressons dans ce travail à l'optimisation combinatoire boîte noire.

En optimisation multiobjective, la notion d'optimalité repose sur la relation de *dominance de Pareto* entre les solutions définie à partir d'un ordre (strict) partiel dans l'espace objectif :

Définition 1 Une solution $x \in \mathcal{X}$ **domine** une solution $x' \in \mathcal{X}$, noté alors $x' \prec x$, si et seulement si $\forall i \in \{1, \dots, m\} f_i(x') \leq f_i(x)$ et $\exists i \in \{1, \dots, m\}, f_i(x') < f_i(x)$.

2. Par exemple, le livre sur le sujet de K. Deb datant de 2001 [Debo1] est cité près de 2000 fois, sans compter les nombreux tutoriels et autres workshops sur le sujet.

Il est alors possible de définir une notion d’optimalisé au niveau d’une solution comme étant une solution non dominée par aucune autre solution :

Définition 2 Une solution $x \in \mathcal{X}$ est **Pareto optimale** ou encore **non dominée** si et seulement si $\forall x' \in \mathcal{X}, x \not\prec x'$.

Sauf cas particulier, contrairement au cas mono-objectif, il n’existe pas une seule solution optimale pour un problème d’optimisation multiobjective, il existe plusieurs solutions non dominées, ou Pareto optimales. Chaque solution Pareto optimale exprime un compromis particulier entre les objectifs. Ceci constitue la principale différence avec l’optimalité mono-objective, l’optimalité d’un problème d’optimisation multiobjective est un ensemble de solutions :

Définition 3 L’ensemble des solutions non dominées (Pareto optimales) est appelé l’ensemble **Pareto optimal** : $POS = \{x \in \mathcal{X} : \forall x' \in \mathcal{X}, x \not\prec x'\}$. L’image de l’ensemble Pareto POS par la fonction d’évaluation \mathcal{F} est appelée le **front de Pareto**.

La Figure 1 illustre les définitions précédentes. La résolution d’un problème multiobjectif consiste à trouver l’ensemble Pareto optimal. Néanmoins, cet ensemble est généralement de taille exponentielle en fonction de la dimension du problème [Ehro6]. Ainsi, il s’agit le plus souvent de trouver une approximation de cet ensemble de cardinalité plus petite et dont la qualité en termes d’approximation du front de Pareto, c’est-à-dire de proximité de l’image des solutions trouvées au front de Pareto, et en termes de diversité, c’est-à-dire la répartition des images des solutions trouvées sur le front de Pareto, sont définies par le décideur.

On distingue aussi souvent deux types de difficulté selon le nombre d’objectifs. L’optimisation multiobjective est parfois réservée aux problèmes dont le nombre d’objectifs est 2 ou 3, alors que les problèmes dont le nombre d’objectifs est 4 ou plus sont appelés optimisation *many-objective*³ [ITNo8]. Cette distinction se fonde surtout sur les buts de chaque type d’optimisation. En effet, lorsque le nombre d’objectifs devient important la probabilité qu’une solution soit Pareto optimale devient grande et l’ensemble Pareto optimal devient de taille très importante. La principale difficulté dans ce cas est de trouver un ensemble diversifié sur le front de Pareto et moins de trouver des solutions proches du front de Pareto qui constitue la principale difficulté dans le cas multiobjectif avec 2 ou 3 objectifs. La conception des algorithmes d’optimisation multiobjective ou au moins leur performance change selon la typologie multi- ou many-objective.

1.1.3 Méthodes stochastiques d’optimisation mono-objective

De manière schématique, on peut distinguer deux grandes catégories de méthodes de résolution de problème d’optimisation : les méthodes *exactes* (ou *complètes*) et les méthodes *approximées* (ou *heuristiques*). Les méthodes exactes garantissent de trouver la solution optimale au problème, mais souvent le temps d’exécution ou l’espace mémoire nécessaire augmentent rapidement avec la dimension du problème, en particulier pour les problèmes NP-difficiles. Et par conséquent, elles sont réservées aux instances de taille modérée. Parmi les méthodes exactes, citons Branch-and-Bound, la programmation dynamique, les méthodes de relaxation Lagrangienne, les méthodes basées sur la programmation linéaire en nombres entiers, de générations de colonnes, etc. Ces méthodes ne sont pas étudiées dans ce travail et l’on pourra par exemple se référer aux livres suivants [NW88] ou [Jün+09].

Dans le cas des grandes instances, on peut sacrifier l’optimalité pour un temps d’exécution plus raisonnable en employant des heuristiques. Les heuristiques sont généralement spécifiques à un problème. Les *méthodes à garantie de performance* (voir [Vaz13] pour une revue) sont des algorithmes de complexité polynomiale garantissant de trouver une solution

3. On pourrait tenter de traduire «many-objectif» à l’aide d’un adverbe latin par *magna-objectif*, mais nous garderons la dénomination usitée de many-objectif.

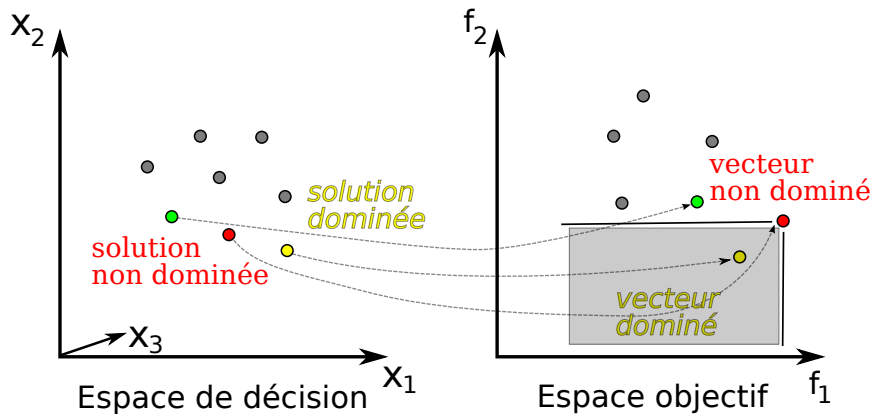


FIGURE 1 – Illustration des définitions de Pareto dominance pour un problème de maximisation biobjectif. Les solutions verte et rouge sont mutuellement *non dominées* ou encore nommées *incomparables*, alors que la solution jaune est dominée par la solution rouge, car meilleure selon les deux objectifs. La solution rouge n'étant dominée par aucune solution de l'espace de décision, elle est Pareto optimale ou non dominée.

à un facteur près de l'optimalité. Malgré tout, ces algorithmes sont spécifiques à un problème et peu adaptés au contexte boîte noire. Les *heuristiques constructives* [SJ13] sont aussi des méthodes spécifiques à un problème. Elles construisent pas à pas les solutions selon une heuristique tenant compte de la définition de celui-ci en s'appuyant sur une fonction d'évaluation partielle. Par exemple, en choisissant l'élément de la solution qui améliore le plus une fonction d'évaluation partielle.

Enfin, la dernière catégorie d'heuristique dont l'étude fait partie de ce travail est celle des *métaheuristiques*. Ce sont des méthodes conçues pour être non spécifiques à un problème. Les métaheuristiques sont des méthodes d'optimisation stochastiques qui s'inscrivent dans ce cadre de l'optimisation boîte noire. Le principe général est d'explorer itérativement de manière stochastique une ou plusieurs solutions à la fois en se guidant uniquement à l'aide de la fonction d'évaluation. Parmi les nombreuses métaheuristiques, citons les plus représentatives que l'on peut classer en deux sous-catégories : les méthodes qui manipulent une seule solution à la fois, métaheuristiques à solution unique aussi appelées parfois métaheuristiques de parcours, comme le recuit simulé [KGV83], la recherche taboue [Glo86], la recherche locale itérée [LMS10], ou les méthodes qui manipulent une population (ensemble) de solutions, métaheuristiques à population, dont les algorithmes évolutionnaires comme les algorithmes génétiques [Hol75] [Gol89]⁴, les stratégies d'évolution [Rec73], la programmation génétique [Koz90], les algorithmes à estimation de distribution [LL02], etc. ou encore des algorithmes bio-inspirés comme les algorithmes par essaims particuliers [KE95] ou de colonie de fourmis [Dor92], etc. Plusieurs ouvrages donnent un large panorama complet sur la conception de ces métaheuristiques : [GKo6] [Dre+06] [Talo9].

De manière générale et à grand trait, les performances des métaheuristiques dépendent de la gestion entre l'exploration de solutions dans une nouvelle région de l'espace de recherche et l'exploitation autour des meilleures solutions déjà rencontrées. Toutes les métaheuristiques ont une façon particulière de gérer ce compromis soit par un concept de mémoire explicite, de taux d'exploration, de maintien de diversité dans une population de solutions, etc. D'ailleurs, bien que non présentés dans le manuscrit, les travaux de thèse de D. Simoncini que j'ai co-encadré ont porté sur une l'étude du compromis entre exploration et exploitation dans les algorithmes évolutionnaires parallèles que sont les algorithmes évolutionnaires cellulaires [Sim+06a ; Sim+06b ; Sim+07 ; Sim+09].

4. Le livre de D. Goldberg reste l'un des ouvrages les plus cités en informatique, 67,600 fois au dernier recensement

1.1.4 Méthodes stochastiques d'optimisation multiobjective

Résoudre un problème d'optimisation multiobjectif consiste à trouver l'ensemble des solutions Pareto optimales (POS), ou plus exactement une approximation de cet ensemble (voir section 3.1). Pour trouver une bonne approximation, les métaheuristiques, et plus particulièrement les algorithmes évolutionnaires, sont largement mises en avant depuis plusieurs dizaines d'années, et ces algorithmes continuent d'être un sujet de recherche de plus en plus actif. À ce jour, on peut distinguer trois types de méthodes heuristiques de résolution.

Les *approches Pareto* améliorent progressivement un ensemble de solutions en se basant sur la relation de dominance Pareto. Les solutions de l'ensemble courant de recherche, ensemble encore appelé *archive*, sont remplacées préférentiellement par de nouvelles solutions qui les dominent. On peut trouver dans ce type d'approche le célèbre algorithme NSGA-II [Deb+02], ou encore des méthodes qui dérivent de recherche locale [Lie+09] comme la Pareto Local Search (PLS) [PCSo4] ou le Global-Simple EMO (GSEMO) [LTZ04]. Le principe de ces algorithmes est illustré avec celui de PLS (algorithme 1) : des solutions voisines d'une des solutions de l'archive servent à mettre à jour l'archive courante à l'aide de la relation de dominance Pareto.

Algorithme 1 La Pareto Local Search (PLS) met à jour l'archive avec les solutions non dominées selon la relation de Pareto créées à partir d'une des solutions non visitées de l'archive.

```

1: Générer une solution aléatoire  $x_0 \in \mathcal{X}$ 
2:  $A \leftarrow \{x_0\}$ 
3: repeat
4:   Sélectionner une solution non-visitée  $x \in A$ 
5:   Créer l'ensemble des solutions voisines  $\mathcal{N}(x)$ 
6:   Marquer  $x$  comme visitée
7:    $A \leftarrow$  solutions non-dominées de  $A \cup \mathcal{N}(x)$ 
8: until toutes les solutions sont visitées

```

Les *approches basées indicateur* ont pour but explicite d'améliorer un indicateur qui mesure la qualité d'un ensemble de solutions. L'exemple fondateur de ce genre d'approche est l'Indicator-Based Evolutionary algorithm (IBEA) de Zitzler [ZK04] dont l'algorithme est donné par l'algorithme 2. SMS-MOEA [BNE07] est un autre exemple d'un tel algorithme. Le schéma d'algorithme respecte celui des algorithmes évolutionnaires avec sélection, variation aléatoire et remplacement. La fitness d'une solution est basée sur la variation de l'indicateur qu'elle engendre lorsque celle-ci est supprimée de l'ensemble. Initialement, les indicateurs de qualité [Fon+05] sont des mesures de performance pour comparer les algorithmes multiobjectif entre eux. Ils associent une valeur numérique à un ensemble de solutions dont la valeur optimale doit être atteinte par l'ensemble Pareto optimale. Les principaux indicateurs sont l'hypervolume qui mesure le volume dans l'espace objectif couvert par l'ensemble et l'indicateur- ϵ qui mesure la translation minimale dans l'espace objectif à réaliser pour que chaque élément d'un ensemble de référence soit dominé par l'une des solutions de l'ensemble. La plupart des indicateurs de qualité respectent la relation de dominance Pareto : les indicateurs sont des fonctions croissantes selon la relation de dominance Pareto.

Les *approches scalaires*, ou encore appelées *approches à décomposition* sont basées sur la scalarisatation du problème multiobjectif en de multiples problèmes mono-objectifs. Les optimisations de chaque sous-problème mono-objectif ne sont pas indépendantes, mais coordonnées pour améliorer l'efficacité globale de résolution du problème multiobjectif. Le représentant le plus marquant de cette approche est l'algorithme MOEA/D [ZL07] («D» pour décomposition) dont les performances sont maintenant reconnues. L'algorithme 3 en donne une description. Chaque sous-problème mono-objectif est défini à partir d'une direction de recherche et d'un point de référence. L'idée est que chaque sous-problème «recherche» dans une partie de l'espace objectif en coopération avec les recherches des problèmes voisins. La fonction scalaire à optimiser est une combinaison des objectifs qui guide la recherche et

Algorithm 2 Indicator-Based Evolutionary Algorithm (IBEA). La fitness d'une solution est définie par la comparaison paire à paire en termes d'indicateur $I : f(x) = \sum_{x' \in P \setminus \{x\}} (-e^{-I(x',x)/\kappa})$. κ est une constante de normalisation.

```

1: P ← initialisation()                                ▷ Crée une population initiale de solutions
2: repeat
3:   P' ← sélection(P)
4:   Q ← variation_aleatoire(P')
5:   Évaluation des solutions de Q à l'aide de f
6:   P ← remplacement(P, Q)
7: until le nombre maximum d'évaluations est atteint

```

permet d'atteindre certains points de l'espace objectif. Classiquement, cette fonction scalaire est une somme pondérée des objectifs : $g(x|\lambda, z_i^*) = \sum_i \lambda_i |z_i^* - f_i(x)|$, ou encore une norme de Tchebychev $g(x|\lambda, z_i^*) = \max_i \lambda_i |z_i^* - f_i(x)|$.

Algorithm 3 MOEA/D. μ sous-problèmes scalaires $g(\cdot|\lambda_i, z_i^*)$ définis par μ directions λ_i et points de référence z_i^* .

```

1: ( $\lambda^1, \dots, \lambda^\mu$ ) ← initialization_direction()
2: Initialize  $\forall i = 1.. \mu$  B(i) the neighboring sub-problems of sub-problem i
3: ( $x^1, \dots, x^\mu$ ) ← initialization_solution()
4: repeat
5:   for i = 1.. $\mu$  do
6:     Select x and x' randomly in { $x_j : j \in B(i)$ }
7:     y ← mutation_crossover(x, x')
8:     for j  $\in B(i)$  do
9:       if  $g(x_j|\lambda_j, z_j^*) < g(y|\lambda_j, z_j^*)$  then
10:         $x_j \leftarrow y$ 
11:       end if
12:     end for
13:   end for
14: until maximum number of evaluations is reached

```

1.2 PAYSAGE DE FITNESS

Le paysage de fitness est le concept central des travaux présentés. Cette section expose les origines du concept et propose une classification des méthodes et des finalités de l'analyse des paysages de fitness.

1.2.1 Origine et définition

Le concept de paysage de fitness a été introduit par S. Wright [Wri32] en 1932 dans le domaine de la biologie. L'idée de Wright est de donner une idée intuitive et cependant riche de l'adaptation des espèces. Le modèle est un espace abstrait constitué de l'espace des génotypes et d'une dimension supplémentaire qui représente la valeur d'adaptation (la valeur de fitness). Wright projette l'espace des génotypes dans un espace à deux dimensions ce qui lui permet de représenter à la fois la géométrie des surfaces possibles (pics, vallées, etc.) et les dynamiques possibles d'adaptation sur le paysage de fitness : marche adaptative vers un pic, transition entre deux pics selon le taux de mutation et la pression de sélection, etc. La figure 2 montre les premières représentations proposées originellement par Wright.

Le concept de paysage de fitness est utilisé dans de nombreux domaines scientifiques tels que la physique (verre de spins [Par86]), biophysique (protéine folding [CD91]), chimie (surface d'énergie [Mez82]), en gestion (étude des organisations [RS02b]), etc. En optimisation

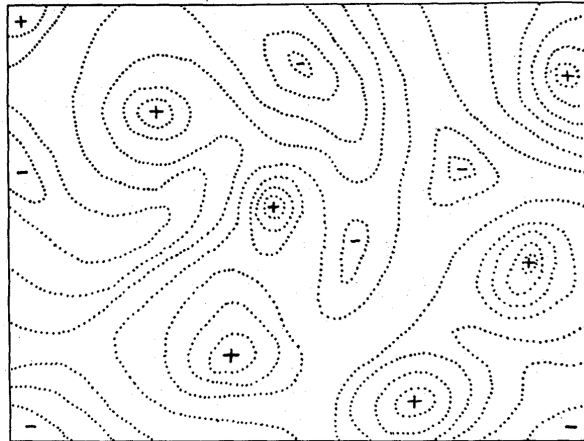


FIGURE 2.—Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness.

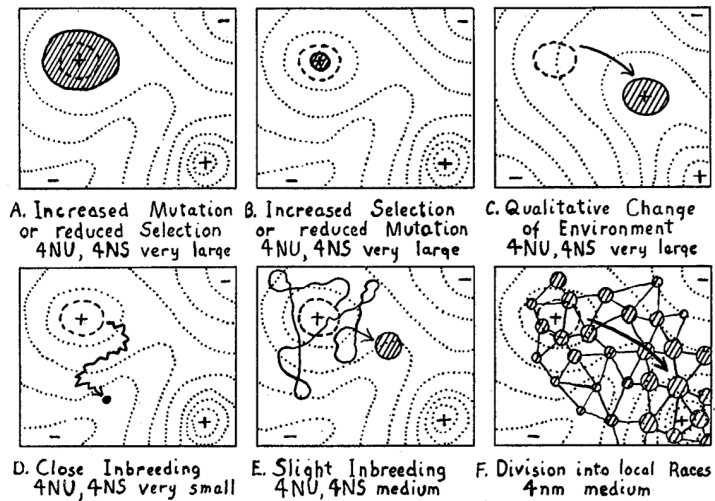


FIGURE 4.—Field of gene combinations occupied by a population within the general field of possible combinations. Type of history under specified conditions indicated by relation to initial field (heavy broken contour) and arrow.

FIGURE 2 – Représentations originales de paysages de fitness issues de l'article [Wri32] de Wright publié en 1932.

stochastique, initialement en évolution artificielle, les paysages de fitness ont été formalisés et étudiés principalement sous l'impulsion des travaux de T. Jones et P. Stadler [MWS91; Jong95a; Sta95]. Les génotypes possibles sont remplacés par les solutions potentielles au problème d'optimisation et la valeur d'adaptation par la fonction de fitness. Un paysage de fitness est un triplet $(\mathcal{X}, \mathcal{N}, f)$ où \mathcal{X} est l'ensemble des solutions potentielles au problème d'optimisation appelé espace de recherche, $\mathcal{N} : \mathcal{X} \rightarrow 2^{\mathcal{X}}$ est une relation de voisinage entre les solutions qui associe à chaque solution un ensemble de solutions dites voisines et $f : \mathcal{X} \rightarrow \mathbb{R}$ est la fonction d'évaluation à optimiser, aussi appelée fonction de fitness dans ce contexte. Dans la suite, sans perte de généralité, on supposera que la fonction objectif doit être maximisée. Les définitions sont données dans le cas mono-objectif. Pour le cas multiobjectif, on peut changer la fonction d'évaluation pour étendre la définition comme le suggèrent certains travaux [GD07; Gar09b], nous verrons dans la section 3.2.4 qu'il est possible de proposer une autre définition finalement plus proche de l'optimisation mono-objective.

Chaque élément du triplet constituant le paysage de fitness doit être défini en fonction du problème à résoudre et de la classe d'algorithmes que l'on désire étudier. Naturellement, l'espace de recherche \mathcal{X} et la fonction de fitness f doivent être ceux qui définissent le problème d'optimisation et qui sont utilisés par les algorithmes de recherche. Pour un même problème d'optimisation, il peut exister plusieurs espaces de recherche, chacun reposant sur un choix de représentation des solutions [Roto6; RG03]. D'ailleurs, bien que ces travaux ne soient pas exposés dans ce manuscrit, l'étude des représentations possibles a fait le sujet de la thèse de M. Bercachi [Ber+07; Ber+08; Ber+09]. Il peut aussi exister plusieurs fonctions à optimiser possibles, en particulier pour les problèmes avec contraintes [MS96; Coe02; Goto1] où la fonction objectif est l'une des relaxations possibles au problème. Toutefois, l'espace de recherche et la fonction objectif sont plus relatifs au problème à résoudre qu'à l'algorithme utilisé. En revanche, la relation de voisinage \mathcal{N} est définie en fonction de la classe d'algorithmes utilisée : «one operator, one landscape» selon la formule de T. Jones [Jong95b]. La relation de voisinage doit être définie selon l'opérateur utilisé dans les algorithmes d'optimisation, ou selon une distance liée aux opérateurs. Typiquement, lorsque l'espace de recherche est l'ensemble des chaînes binaires d'une longueur donnée, le voisinage est l'ensemble des solutions à une distance de Hamming inférieure à un entier k où k est le nombre de bits changés par l'opérateur. Lorsqu'un opérateur stochastique op est utilisé dans un algorithme, le voisinage $x \in \mathcal{X}$ peut se définir par $\mathcal{N}(x) = \{y \in \mathcal{X} \mid \Pr(y = op(x)) > \epsilon\}$ où ϵ est une constante positive ou nulle donnée. La figure 3 illustre en deux dimensions de la même manière que Wright la définition d'un paysage de fitness pour l'optimisation mono-objective.

La spécificité d'un paysage de fitness peut être perçue comme une limitation : l'étude d'un paysage de fitness pour un problème et un algorithme d'optimisation ne pouvant pas directement apporter d'information pour un autre algorithme d'optimisation. Cependant, de même que la difficulté d'optimisation d'un problème dépend de l'algorithme considéré, la structure du paysage de fitness dépend également de certaines propriétés particulières de l'algorithme et du problème. Par exemple, tous les algorithmes n'ont pas le même temps de résolution sur une classe particulière de problèmes [WM95; WM97]. L'étude d'un paysage de fitness permet d'analyser et de comparer différentes représentations, différentes fonctions objectif ou différents opérateurs de recherche locale. De manière générale, les buts sont alors de mieux comprendre l'interaction entre le problème et l'algorithme et de sélectionner les composants d'algorithme les plus adaptés selon la classe de problèmes à résoudre.

1.2.2 Buts de l'analyse des paysages de fitness

Le but principal de l'analyse de paysages de fitness est la compréhension de la structure de l'espace de recherche du point de vue de l'algorithme d'optimisation. Cette compréhension permet de repérer les écueils et les avantages des algorithmes d'optimisation, typi-

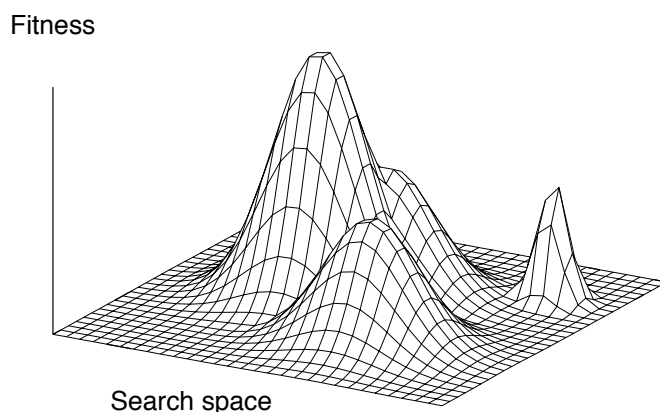


FIGURE 3 – Représentation d’un paysage de fitness en optimisation combinatoire mono-objective.

quement des métaheuristiques, afin de sélectionner ou d’améliorer ensuite ces algorithmes d’optimisation selon le problème d’optimisation à résoudre.

Il existe deux types d’analyse de paysage de fitness. L’une est algébrique s’inscrivant dans un scénario d’optimisation boîte grise. À partir d’une classe connue de problèmes d’optimisation comme le problème du voyageur de commerce (TSP) ou un problème de satisfiabilité (SAT), des propriétés algébriques du problème sont étudiées. Les travaux de Stadler [Stao2] s’inscrivent dans cette démarche. Parmi d’autres, l’exemple le plus représentatif de ses travaux est celui sur les paysages élémentaires [Sta96 ; RSo2a] qui étudie le Laplacien du graphe induit par le paysage de fitness. Ce type d’analyse permet d’obtenir des propriétés universelles sur la structure du paysage de fitness pour une classe de problèmes d’optimisation. Cette démarche et ces résultats sont prolongés encore aujourd’hui dans les travaux de D. Withley et ses coauteurs [Whi15] : à partir de la connaissance précise du problème (de type NK-landscapes, max-SAT), ils développent des techniques à faible complexité de calcul afin de trouver des optima locaux à l’aide des opérateurs de mutation [Whi+14] ou de croisement [CWA14].

L’autre type d’analyse de paysage de fitness est statistique et s’inscrit dans un scénario d’optimisation boîte noire où seuls le codage et la qualité des solutions sont accessibles et connus. À partir d’un échantillonnage du paysage de fitness constitué de solutions de l’espace de recherche et de la valeur correspondante de la fonction objectif, des propriétés statistiques sont extraites et utilisées pour caractériser le paysage de fitness. Ce type d’approche est de loin celui le plus utilisé en pratique. En effet il s’inscrit parfaitement dans la problématique de la sélection d’algorithmes de Rice [Ric76] qui a été repris dans le contexte du choix d’algorithmes en optimisation [Smio9 ; ME14a ; Muñ+15]. La figure 4 schématise le principe du modèle de la sélection d’algorithmes. Tout d’abord, les caractéristiques d’un problème d’optimisation sont extraites, puis à partir de celles-ci, un algorithme est sélectionné pour en mesurer ces performances. Le schéma de K. Malan [ME14a] ajoute qu’à partir des caractéristiques du problème et de l’algorithme sélectionné, il soit possible de prédire les performances de l’algorithme en utilisant une technique d’apprentissage automatique et sans exécuter celui-ci.

L’analyse de paysage de fitness prend une part essentielle dans la sélection d’algorithmes. Dans un contexte d’optimisation boîte noire, l’analyse de paysage de fitness consiste en

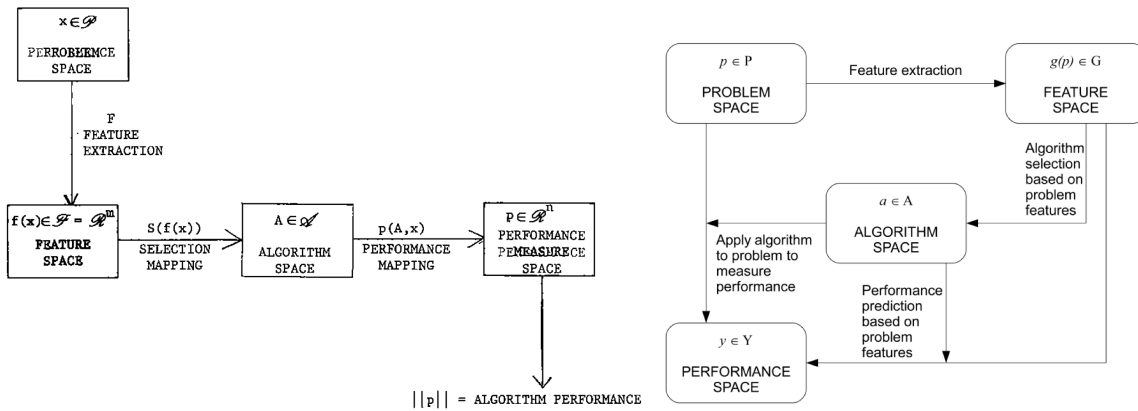


FIGURE 4 – Schémas généraux du choix d’algorithmes selon le modèle de Rice [Ric76] et selon le modèle de Malan [ME14a] qui raffine le modèle original pour l’optimisation en ajoutant la prédiction de performances.

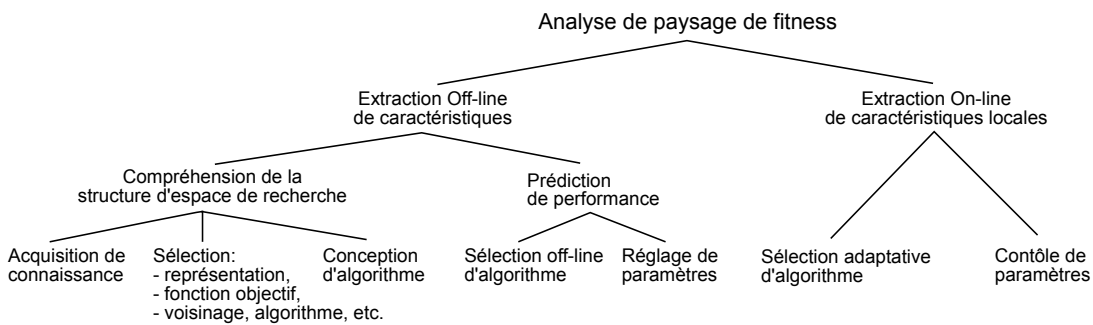


FIGURE 5 – Finalités de l’analyse de paysage de fitness : contexte (première décision de l’arbre), méthodologie (deuxième décision de l’arbre) et buts (feuilles de l’arbre) de l’analyse de paysage de fitness.

l’**extraction de métriques** de problème (sans connaissance *a priori* du problème) - association entre l’espace des problèmes et l’espace des caractéristiques - et dans la **prédiction de performance** - association entre l’espace des algorithmes et l’espace des performances. Le challenge dans le domaine des paysages de fitness est de proposer les métriques pertinentes qui permettent de décrire la structure de l’espace de recherche en lien avec la complexité d’un algorithme d’optimisation de type métaheuristique.

De manière effective, les finalités de l’analyse de paysage de fitness telles que l’on peut les trouver dans la littérature peuvent se classer en plusieurs catégories selon le contexte, les moyens et les buts de l’analyse. La figure 5 en donne une synthèse. Comme pour le réglage des paramètres [Eib+07], la classification distingue le contexte off-line avant le processus d’optimisation et du contexte on-line pendant le processus d’optimisation. Dans le contexte off-line, il s’agit de comprendre la dynamique d’une classe d’algorithmes ou de concevoir différents composants d’un algorithme pour d’une utilisation ultérieure. De nombreuses métriques (mesurant les caractéristiques de multimodalité, de rugosité, de neutralité, etc.) peuvent être calculées. Ces métriques peuvent être globale au paysage de fitness comme la corrélation fitness distance [JF95], ou locale c’est-à-dire au niveau du voisinage des solutions. Dans tous les cas, souvent une mesure ou une série de mesures résume une caractéristique du paysage. Certaines de ces caractéristiques seront détaillées dans le chapitre suivant et l’on pourra par exemple lire les références suivantes [ME13; Ver13; Ver15]. À partir des caractéristiques extraites du paysage de fitness, deux méthodologies sont possibles. L’une tente de comprendre la structure du paysage en dressant un portrait général du paysage

ou en comparant plusieurs paysages de fitness. L'autre méthode fait appel à des techniques d'apprentissage automatique pour sélectionner ou pour régler un algorithme adapté au problème.

Les finalités relatives à la compréhension de la structure de l'espace de recherche sont classées de gauche à droite sur la Fig. 5 des plus fondamentales aux plus proches de la conception et de l'amélioration des algorithmes.

L'une des finalités fondamentales de l'analyse de paysage de fitness est l'acquisition de nouvelles connaissances sur une classe de problèmes d'optimisation du point de vue de la métaheuristique. On peut trouver de très nombreux travaux avec cette finalité. Parmi bien d'autres, citons les premiers travaux [MWS91] [MW92] ou parmi les plus récents par Tayarani-N et Prugel-Bennett sur une série de problèmes d'optimisation combinatoire [TP14], comme les problèmes de satisfiabilité [PT12] ou de coloration de graphe [TP15], ou encore certains de mes travaux qui ne seront pas mentionnés dans la suite du document en programmation génétique [Col+06; Van+12], sur les paysages de fitness NK [Van+09], ou sur le problème de planification de type flow-shop [Mar+11b]. Le but de cette démarche est de comprendre les performances des algorithmes d'optimisation au-delà des performances obtenues sur un benchmark. Le portrait dressé par l'analyse de paysage de fitness permet d'analyser plus finement la relation de l'algorithme avec la structure du problème afin d'appréhender dans quel contexte un algorithme peut être performant ou non. Pour illustrer cette démarche citons la phrase d'un article de Y. Wu concernant un problème d'apprentissage d'un réseau bayésien [WMC11] où il étudie un paysage de fitness : «*Differences in algorithm performance are explained in terms of these landscapes, which in turn are related to properties of the bayesian network structures*». Cette phrase montre bien l'intention de nombreux auteurs lorsqu'ils mènent ce type de recherche : expliquer les performances en termes de structure de l'espace de recherche.

L'analyse de paysage de fitness est également utilisée pour sélectionner certains composants importants des algorithmes de recherche. Le principe consiste à comparer les métriques de deux paysages de fitness relatifs à un même problème d'optimisation où plusieurs choix de représentations, fonctions de fitness, ou de voisinages sont possibles. La sélection et sa justification d'une composante de l'algorithme sont alors effectuées avant l'optimisation et non pas seulement *a posteriori* par l'observation de performance sur un benchmark. Ce type de travaux est sans doute le cas le plus courant et le plus simple d'utilisation de l'analyse de paysage de fitness et l'on peut trouver de nombreux travaux de comparaison dont il n'est pas possible de dresser la liste exhaustive. Citons des exemples représentatifs pour chacune des comparaisons possibles : le choix de représentation pour le problème du sac à dos multidimensionnel [TPCo8], le choix de la fonction de fitness pour un problème de génération de test logiciel [Llo8] ou encore le choix de l'opérateur de mutation et de cross-over pour le problème d'affectation quadratique (QAP) [MFOoa], ou pour le problème de partitionnement [AR14].

Enfin, la compréhension de la structure du paysage de fitness peut servir à concevoir un algorithme adapté à la forme à ce paysage. On ne concevra pas le même type d'algorithme selon que le paysage est très multimodal avec de petits bassins d'attraction ou très «plat» avec beaucoup de grands plateaux qui dominent la forme générale. Dans cette approche nous trouvons par exemple des travaux sur le voyageur de commerce [Fon+99], le problème de partitionnement de graphe [MFOob] où un algorithme évolutionnaire hybridé (memetic) est basé sur l'analyse de paysage de fitness, ou encore un problème applicatif comme la conception d'antennes [AZSo2]. J'ai pu aussi également travaillé sur le problème de planification de type flow-shop où l'algorithme de type recherche locale itérée a été conçu en se basant sur les caractéristiques neutres du paysage de fitness [Mar+11c; Mar+11a; Mar+12].

Dans le contexte off-line, l'analyse de paysage de fitness permet aussi de prédire les performances des algorithmes d'optimisation à l'aide de techniques d'apprentissage automatique. On distingue alors deux finalités : l'une consiste à sélectionner un algorithme adéquat, l'autre à régler certains paramètres d'un algorithme donné (voir Fig. 5).

De nombreuses méthodes ont été proposées pour le réglage automatique off-line des algorithmes, souvent appelé tuning selon la classification proposée par Eiben *et al.* [Eib+07], dont le principe repose sur une optimisation de la valeur des paramètres qui peuvent être catégoriels, discrets ou continus. Les méthodes d'optimisation utilisées procèdent donc par essai/erreur en optimisant les paramètres eux-mêmes sur un ensemble d'instances d'un problème. Les meilleures méthodes sont REVAC++ [SE09], ParamILS [Hut+09], Gender-Based Genetic Algorithm [AST09], iterated F-Race [Bir+10]. Une autre approche repose sur le schéma de Rice (Fig. 4) de la sélection d'algorithmes. À partir de métriques du problème, la performance est prédite, puis soit un algorithme est sélectionné parmi un portefeuille d'algorithmes (technique de portfolio) [Smio9], soit les paramètres sont réglés comme dans la méthode SMAC [HHL11] utilisé dans le solveur SATzilla [Xu+08]. Bien sûr les métriques peuvent être en partie issues de l'analyse de paysage de fitness. Une fois le modèle prédictif établi, l'estimation des métriques du paysage doivent avoir une complexité raisonnable par rapport au budget global de l'optimisation. Les premiers travaux sur le réglage de paramètres basés sur l'analyse de paysage de fitness sont contemporains de l'essor du domaine de l'évolution artificielle, Grefenstette [Gre95] utilise la distribution de fitness des opérateurs et se poursuivent encore aujourd'hui comme dans [LLY11b; LLY11a] où là encore des mesures relatives à la probabilité d'amélioration des opérateurs sont utilisées. Les techniques de portfolio de métaheuristiques se développent rapidement de nos jours. Hutter et ces coauteurs proposent une étude des méthodes de prédictions de performance [Hut+14]. On peut aussi trouver une revue des méthodes portfolio dans le domaine de l'optimisation combinatoire dans [Kot14] et dans le domaine de l'optimisation continue dans [Mer+11; Muñ+15].

Dans le contexte de l'analyse de paysage de fitness on-line, au cours du processus d'optimisation, certaines caractéristiques du paysage de fitness sont mesurées et utilisées afin de modifier les paramètres de l'algorithme ou bien de sélectionner l'algorithme le plus adapté parmi un portefeuille d'algorithmes (branche de droite du schéma de la figure 5). L'estimation des métriques doit avoir un coût de calcul modéré pour être exploitée durant l'optimisation. Pour cette raison, seules les caractéristiques locales, c'est-à-dire relatives à la qualité des solutions visitées par l'algorithme ou à la qualité de leurs solutions voisines, sont utilisées. En optimisation combinatoire, l'idée de faire évoluer les paramètres d'un algorithme génétique en fonction sur l'état de la recherche n'est pas nouvelle [SP94] et une revue sur le sujet date même de 1997 [HME97]. Le sujet est toujours très actif. Par exemple, une méthode adaptative de contrôle de la taille de population est proposée basée sur la probabilité d'amélioration de l'algorithme [DD15]. Récemment, des techniques d'apprentissage automatique séquentiel sont utilisées pour sélectionner les opérateurs de manière on-line : poursuite adaptative [Thio5] ou Upper Confidence Bound (UCB) [Fia+10]. En optimisation continue, pour l'instant, peu de travaux tentent cette approche portfolio [YCZ13; BP14]. On trouve aussi cette problématique en environnement de calcul distribué avec le modèle en île des algorithmes évolutionnaires soit pour contrôler les paramètres de communication entre les îles [Can+12; Can+13; GLS16] soit pour régler les algorithmes qui s'exécutent sur chaque île [DV11; Jan+16]. Hormis la technique d'apprentissage utilisée qui permet de sélectionner, la principale difficulté de ces travaux est d'estimer une métrique pertinente du paysage de fitness qui mesure la qualité d'un opérateur ou d'un paramètre. La plupart des approches se basent sur l'amélioration de l'algorithme entre deux itérations [Fia+08; Mat+09], mais sûrement d'autres recherches sont à mener dans ce domaine.

1.3 APERÇU DU DOCUMENT

La suite du manuscrit expose certains de mes travaux de recherche que j'ai pu réaliser au cours d'encadrement de thèses et de travail de master ou en collaboration avec d'autres collègues. Pour plus de clarté, certains des travaux ont été mis de côté afin de montrer plus

clairement la démarche dans l'étude des paysages de fitness. Le manuscrit est donc divisé en deux chapitres principaux l'un traitant du réseau des optima locaux en optimisation mono-objective et l'autre du thème général des paysages de fitness en optimisation multiobjective.

Le chapitre 2 expose plusieurs de nos contributions sur le réseau des optima locaux qui est un nouveau modèle pour les paysages de fitness en optimisation mono-objective. La première section 2.1 introductive donne un aperçu des mesures utilisées dans l'analyse des paysages de fitness. Elle propose également un point de vue sur l'analyse des paysages de fitness par la technique d'agrégation de variables. Elle inscrit aussi nos travaux dans les sciences des systèmes complexes et rappelle les récents travaux sur les réseaux complexes qui sont utilisés par la suite. La section suivante 2.2.1 donne les principales définitions du réseau des optima locaux en montrant les origines du concept. Ensuite dans la section 2.2.2, la démarche consiste à d'abord confirmer la pertinence du concept sur les petites instances de la classe générique des paysages NK de Kaufman et en même temps à montrer de nouvelles propriétés de ces paysages de fitness. Puis, nous montrons dans la section 2.2.3 que l'on peut utiliser le réseau des optima locaux pour les finalités identifiées des paysages de fitness (comparaison de difficulté d'instances, d'opérateurs, etc.) sur des petites instances de problème de différents problèmes classiques d'optimisation combinatoire. Nous montrons dans la section 2.2.4 que les métriques mesurant les caractéristiques du réseau des optima locaux permettent d'expliquer les performances de métaheuristiques. Enfin dans la section 2.2.5, notre démarche montre que l'on peut également étendre l'étude du réseau des optima locaux sur les instances de problème de grande dimension en utilisant une méthode d'échantillonnage qui estime certaines métriques du réseau. En effet, nous montrons qu'il est possible de prédire les performances de métaheuristiques et de sélectionner la plus pertinente en fonctions de ces métriques. Le chapitre dans la section 2.3 conclue en résumant les principaux résultats et en proposant des perspectives à ces travaux.

Le chapitre 3 développe nos recherches sur la structure des problèmes d'optimisation multiobjective. La section introductive 3.1.1 fait le bilan des sources de difficultés potentielles des approches de résolution multiobjective. Ce bilan est nécessaire afin d'utiliser et de proposer de nouvelles métriques pertinentes de paysage de fitness pour un problème d'optimisation multiobjective. La démarche méthodologique est la même que pour l'optimisation mono-objective. Nous établissons d'abord dans la section 3.2.1 un portrait des espaces de recherche pour les problèmes énumérables de petite dimension, puis nous étendons ces métriques en les estimant, lorsque cela est possible, au cas des problèmes de grande dimension. De même, nous passons de résultats qualitatifs qui permettent de proposer des conseils de conceptions d'algorithmes à une démarche quantitative orientée données en expliquant, dans le chapitre 3.2.2, puis en prédisant les performances d'algorithmes multiobjectives, dans le chapitre 3.2.3. Les premières métriques de paysages de fitness multiobjectifs s'appuient sur une définition implicite de paysage qui considère seulement le changement de dimension de la fonction de fitness. Dans la section 3.2.4, nous proposons une nouvelle définition des paysages fitness qui envisage l'espace de recherche des ensembles de solutions et un indicateur de qualité sur ces ensembles comme fonction de fitness. Cette définition permet d'utiliser les outils standards d'analyse des paysages de fitness. Dans la section 3.2.5, nous étudions les spécificités des approches scalaires en particulier l'influence des directions de recherche sur la dynamique des algorithmes. Nous proposons un algorithme distribué adaptatif qui contrôle ces directions de recherche dans un environnement de calcul distribué. Enfin, pour résoudre une version multiobjective du problème quadratique binaire non contraint, nous proposons un algorithme d'optimisation hybride entre une approche scalaire et Pareto en partie inspiré par la structure du paysage de fitness de ce problème. La dernière section 3.3 établit une conclusion qui résume les principaux résultats du chapitre et propose quelques perspectives.

Le manuscrit se termine par un court chapitre 4 qui expose certaines perspectives au-delà des perspectives exprimées dans chaque chapitre.

2

RÉSEAU DES OPTIMA LOCAUX

CONTENU

2.1	Introduction	16
2.1.1	Analyse de paysage de fitness	16
2.1.2	Un point de vue sur l'analyse de paysage de fitness	16
2.1.3	Science des réseaux complexes	19
2.2	Modélisation des paysages de fitness par le réseau des optima locaux	20
2.2.1	Réseau des optima locaux	21
2.2.2	Nouvelles propriétés des paysages de fitness	23
2.2.3	Analyses basées sur le réseau des optima locaux	28
2.2.4	Explication de performance des métaheuristiques	31
2.2.5	Prédiction de performance pour les problèmes de grande dimension	36
2.3	Conclusion et perspectives	42

RÉSUMÉ

Ce chapitre contient les points suivants relatifs aux publications indiquées :

- Un aperçu de l'analyse des paysages de fitness avec les métriques existantes, suivi d'un point de vue personnel sur les techniques d'analyse en lien avec les systèmes complexes.
- Une brève revue sur les récents travaux concernant les réseaux complexes.
- La définition du réseau des optima locaux qui repose sur le principe d'agrégation de variables, technique utilisée dans les systèmes complexes. [Och+08; Ver+11a; Och+14]
- La mise en évidence de nouvelles propriétés des paysages de fitness à l'aide du réseau des optima locaux. [TVOo8; VOTo8; VOT10]
- Plusieurs analyses de paysage de fitness basées sur le réseau des optima locaux sur différents problèmes d'optimisation combinatoire. [OVT10; Dao+10; Dao+11; Och+11]
- L'explication de performance par des modèles de régression linéaire dont les prédicteurs sont les métriques du réseau des optima locaux. [Dao+12; Chi+12; Dao+13]
- La prédiction de performance de plusieurs métaheuristiques pour les problèmes mono-objectif de grande dimension.
- Une conclusion qui résume la démarche et les résultats obtenus, ainsi que des perspectives spécifiques pour ces travaux.

Principaux collaborateurs pour ces travaux : Fabio Daolio (PhD, université de Lausanne, Suisse), Marco Tomassini (université de Lausanne, Suisse), Gabriela Ochoa (Université de Stirling, UK).

2.1 INTRODUCTION

2.1.1 Analyse de paysage de fitness

Beaucoup de techniques ont été explorés dans la littérature pour analyser les paysages de fitness. Les outils de l'analyse des paysages de fitness permettant d'étudier les dynamiques complexes adaptatives sont issus de différentes disciplines telles que la physique théorique, la chimie, la biologie de l'évolution, l'optimisation par algorithme évolutionnaire, etc. Ce domaine s'inscrit ainsi pleinement dans les sciences de la complexité. Classiquement en évolution artificielle, pour élaborer une nouvelle technique d'analyse, une caractéristique du paysage de fitness est mise en avant. Elle est ensuite mise en relation avec la difficulté d'optimisation des métaheuristiques sur une classe de problèmes dont le degré de cette caractéristique peut être facilement contrôlé. Cette démarche a été fortement impulsée par les travaux pionniers de M. Mitchell, S. Forrest et J. Holland en 1992 [MFH92]. Le tableau 1 synthétise les principales caractéristiques de paysage de fitness qui ont été étudiées. Les détails de ces mesures ne sont pas donnés dans ce document, pour une revue plus complète, on pourra consulter les tutoriels [VO10; Ver15], l'article de K. Malan [ME13] ainsi que les articles originaux indiqués. Pour résumer brièvement, les principales mesures s'intéressent aux caractéristiques telles que les optima locaux (multimodalité) que sont les «pics» du paysage de fitness, l'interaction des variables codant les solutions (épistasie), la pertinence de la fonction de fitness dans le guidage vers l'optimum (déceptivité), la notion très importante de régularité et qualité du voisinage (évolubilité), ou encore la forme de plateaux des paysages de fitness lorsque beaucoup de solutions ont la même qualité (neutralité). Par ailleurs, j'ai pu développer les outils logiciels qui intègrent toutes ces mesures à travers la plateforme ParadisEO-MO [Hum+13a], module de la plateforme ParadisEO¹ [CMT04] dédiée aux métaheuristiques à solution unique, dont je suis le coresponsable depuis cette refonte.

2.1.2 Un point de vue sur l'analyse de paysage de fitness

À la suite de ces travaux sur l'analyse de paysage de fitness, on peut distinguer deux approches dans l'étude des paysages de fitness. La première approche étudie le paysage au niveau des solutions elles-mêmes. Les travaux remarquables de P. Stadler réalisent l'analyse spectrale du graphe induit par les solutions et la relation de voisinage, les nœuds étant valués par la fonction de fitness. Le graphe est de taille très importante puisque égale à la taille de l'espace de recherche, mais lorsque la fonction de fitness a des propriétés particulières comme l'additivité², le Laplacien du graphe (intuitivement, équivalent à la fonction dérivée sur un graphe) reste simple [Stao2] (paysages élémentaires). Ainsi, P. Stadler, ses coauteurs et d'autres auteurs à la suite [RSo2a; Whi+14] ont pu montrer de nombreuses propriétés algébriques des paysages de fitness en particulier dans un contexte d'optimisation boîte grise où la forme de la fonction de fitness est connue.

La seconde approche dans l'étude des paysages de fitness utilise implicitement sans jamais clairement de dénommer un principe d'*agrégation de variables* que l'on rencontre dans d'autres sciences de la complexité telle que la physique ou la biologie (voir [APoo] par exemple). Cette technique s'utilise lorsque le système à étudier est de grande dimension, avec beaucoup de variables et où le système d'équation qui décrit la dynamique au niveau des éléments du système est en général trop grand pour être étudié formellement. Par ailleurs, pour appliquer une méthode d'agrégation de variables, la dynamique du système doit montrer deux échelles caractéristiques, comme des deux échelles de temps ou deux échelles d'espace. Dans ce cas, une méthode d'agrégation de variables permet de construire un modèle réduit de la dynamique du système. Le principe consiste à regrouper ensemble

1. ParadisEO est un projet avec un code ouvert (open source) et libre : <http://paradiseo.gforge.inria.fr/>

2. Intuitivement, une fonction de fitness est additive lorsqu'elle s'écrit sous forme de somme de composantes qui ne dépend que d'un nombre, si possible, limité de variables indépendantes.

TABLE 1 – Principales caractéristiques et métriques de paysage de fitness en optimisation combinatoire avec leur courte description. Pour plus de détails, consulter les articles originaux précisés ou par exemple [ME13].

Nom	Domaine	réf.
Longueur des marches adaptatives ℓ tel que $(x_0, x_1, \dots, x_\ell)$ avec $f(x_t) < f(x_{t+1})$	multi-modalité	[KL87]
Autocorrélation de fitness $\text{corr}(f(x_t), f(x_{t+k}))$ avec marche aléatoire (x_0, x_1, \dots)	épistasie	[Weig0]
Épistasie de signe Concordance de signe entre $f(x) - f(\bar{x}^i)$ et $f(\bar{x}^j) - f(\bar{x}^{j,i})$	épistasie	[WWC05]
Densité des états Distribution de fitness de solution aléatoire	recherche aléatoire	[REA96]
Fitness distance correlation Coefficient de corrélation de $(\text{dist}(x^*, x), f(x^*) - f(x))$	déceptivité	[JF95]
Paysage élémentaire $\Delta f - \lambda(f - k.\mathbf{1}) = 0$ où Δ Laplacien du graphe, λ et k des constantes.	évolvabilité	[Stag6]
Nuage de fitness Distribution bi-variée de $(f(x), f(\text{op}(x)))$	évolvabilité	[VCC03]
Nuage fitness-probabilité Distribution bi-variée de $(f(x), \Pr(f(\text{op}(x)) > f(x)))$	évolvabilité	[LLY11b]
Degré de neutralité $\text{deg}_n(x) = \#\{y \in \mathcal{N}(x) : f(y) = f(x)\}$	neutralité	[Sch+94]
Autocorrélation du degré de neutralité $\text{corr}(\text{deg}_n(x_t), \text{deg}_n(x_{t+k}))$ avec marche aléatoire neutre (x_0, x_1, \dots)	neutralité	[Bas+03]
Autocorrélation de l'évolvabilité $\text{corr}(\text{evol}(x_t), \text{evol}(x_{t+k}))$ avec marche aléatoire neutre (x_0, x_1, \dots)	neutralité	[VCC06]

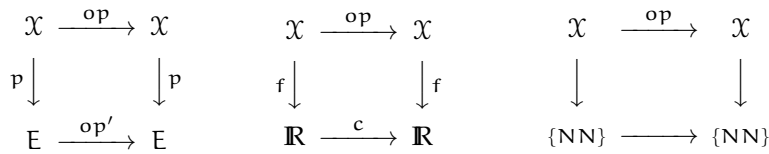


FIGURE 6 – Diagrammes illustrant l’agrégation de variables. À gauche, le cas général. \mathcal{X} est l’espace de recherche et op est l’opérateur stochastique associé à la relation \mathcal{N} de voisinage du paysage. E est l’espace agrégé de dimension plus petite que \mathcal{X} où se projette par p les solutions de l’espace de recherche. L’analyse de la dynamique du système et du paysage de fitness peut alors porter sur op' qui modélise la dynamique dans l’espace agrégé E . Au centre, un exemple où l’espace agrégé est \mathbb{R} celui des valeurs de fitness et la projection est directement la fonction de fitness. À droite, pour l’étude de la neutralité, l’espace agrégé est celui des réseaux de neutralité, c.-à-d. les plateaux structurés par la relation de voisinage.

les éléments qui appartient à la même «petite» échelle et d’exprimer la nouvelle dynamique à une échelle «large» avec un nombre plus restreint de variables agrégées.

En optimisation stochastique par métaheuristique, le système étudié est de grande taille, l’espace de recherche avec l’ensemble toutes les solutions possibles est souvent de taille exponentielle en fonction de la dimension. Il peut donc être approprié d’utiliser un principe d’agrégation de variables en partitionnant l’espace de recherche en un nouvel espace E de partitions (ou encore appelé états) et en associant à chaque solution un élément de E à l’aide d’une projection p . Ainsi, la dynamique de la métaheuristique s’exprime dans le nouvel espace agrégé E . En effet, la dynamique sur l’espace de recherche \mathcal{X} dépend du voisinage \mathcal{N} auquel est naturellement associé un opérateur stochastique op . La dynamique sur E peut se modéliser en exprimant en quelque sorte l’effet de op sur les projetés obtenus par p . La figure 6 donne une illustration par un diagramme de la technique d’agrégation de variables pour les paysages de fitness. Le choix de l’agrégation (espace E et projection p) est d’autant plus pertinent que la dimension de l’espace agrégé E est réduite et que le diagramme est «proche» de la commutativité, c’est-à-dire lorsque $\text{op}' \circ \text{p} = \text{p} \circ \text{op}$.

Les caractéristiques principales d’étude de paysage de fitness s’expriment dans ce cadre de l’agrégation de variables et du diagramme de la figure 6. Les mesures d’évolvabilité et d’épistasie s’étudient lorsque l’espace agrégé E est \mathbb{R} celui des valeurs de fitness. La neutralité s’étudie lorsque l’espace agrégé est celui des réseaux de neutralité, c’est-à-dire des plateaux du paysage. L’analyse du système agrégé porte sur les deux échelles : la dynamique inter à large échelle entre les nouveaux états définis par l’ensemble agrégé E et la dynamique intra à petite échelle au sein de chaque état pour l’ensemble des solutions qui ont le même représentant dans l’ensemble agrégé E . Pour l’étude large échelle de l’opérateur op' , on peut utiliser plusieurs techniques. Celle employée par les recherches sur le running time en évolution artificielle consiste à étudier la chaîne de Markov définie par op' à l’aide de martingale [LW13]. En cela, l’analyse des paysages de fitness s’approche de ces études théoriques. Une autre possibilité est de définir directement une statistique pertinente directement basée sur le diagramme (corrélation, autocorrélation, probabilité d’amélioration, etc.). Enfin un autre alternative riche est de définir un graphe à partir de E et de la relation de voisinage associée à op' . L’analyse de paysage de fitness repose alors sur l’étude d’un graphe de taille réduite et potentiellement porteur d’informations pertinentes sur la dynamique.

Le Réseau des Optima Locaux (ROL), défini dans la section 2.2.1, utilise le principe d’agrégation de variables. L’espace agrégé est celui des optima locaux et les arcs qui reflètent la dynamique entre les optima locaux sont définis de plusieurs manières qui seront discutées. L’autre source qui a mené à définir ce réseau des optima locaux est les travaux du chimiste J. Doye [Doy02] sur l’étude de la topologie des réseaux intrinsèques des surfaces d’énergie définis par Stillinger et Weber [SW84]. La surface d’énergie correspond au potentiel d’énergie des conformations d’une molécule. Les nœuds du réseau intrinsèque de cette surface sont les minima locaux de la surface qui correspondent aux conformations stables des molécules.

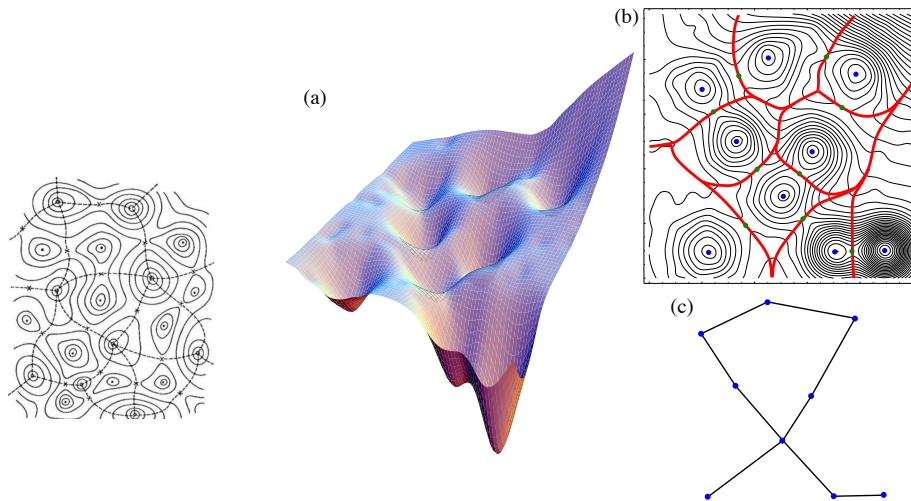


FIGURE 7 – Réseaux intrinsèques de surface d'énergie extraits des articles de Stiliinger et Weber [SW84] (à gauche) et de Doye [Doy02] (à droite). L'énergie est minimisée. Les lignes continues sont les lignes de niveaux de même énergie. Les cercles sont les minima, les bassins d'attraction sont délimités par les lignes pointillées ou rouges, et les possibilités de transitions entre les bassins sont marquées sur chaque ligne. Doye montre les différentes étapes (a-c) pour obtenir le réseau intrinsèque et en faire l'étude avec les outils des réseaux complexes.

La surface est alors découpée en bassins d'attraction qui entourent les minima, car le système passe la plus grande partie du temps dans ceux-ci et transite plus rarement d'un bassin à un autre. Ces transitions définissent les arcs du réseau intrinsèque. La figure 7 reprend les figures des articles cités précédemment pour illustrer le réseau intrinsèque. Clairement, le réseau intrinsèque utilise une technique agrégation de variables pertinente au vu de la dynamique du système étudié. L'originalité des travaux de Doye réside dans l'analyse du réseau intrinsèque de type «réseau complexe». Le ROL transpose au domaine de l'optimisation les définitions du réseau intrinsèque de Stiliinger et Weber et nous analysons la structure de ce réseau en nous inspirant des travaux de Doye.

2.1.3 Science des réseaux complexes

Au tournant du siècle dernier, les articles de D. J. Watts et S. H. Strogatz [WS98] et de A.-L. Barabasi et R. Albert [BA99] ont permis aux sciences des réseaux d'émerger, ou d'émerger de nouveau selon la date de commencement auquel on s'accorde. En effet depuis quelques années, l'intérêt pour l'étude de la structure des grands réseaux ne cesse d'augmenter. La définition d'un réseau repose sur une relation binaire qui permet de modéliser la relation directe entre deux entités et sur l'échange d'information entre elles. Les relations binaires étant très abstraites et usuelles, les réseaux sont l'un des universaux au même titre que les équations différentielles, les systèmes linéaires, les groupes d'invariants, etc. pouvant décrire des réalités diverses. Watts et Strogatz ont montré que de nombreux réseaux que l'on rencontre dans le monde réel ne sont ni complètement aléatoires au sens de Erdős ni complètement ordonnés comme dans un cristal. En définissant et étudiant deux propriétés, la densité locale du réseau et l'éloignement moyen qui peuvent être quantifiés par le coefficient d'agglomération³ et la distance moyenne entre deux nœuds⁴, ils montrent que beaucoup de réseaux sociaux, biologiques, ou artificiels ont une topologie de "petit-monde", c'est-à-dire ayant une densité très importante et une distance moyenne petite.

3. «clustering coefficient» en langue anglaise

4. «average path length» en langue anglaise

TABLE 2 – Principales métriques sur un réseau et leur définition succincte. Voir [Bar15] pour de plus amples détails. Le réseau est un réseau pondéré (V, E) dont les arcs de E sont pondérés par les poids positifs ou nuls $(w_{ij})_{(i,j) \in E}$.

Nom	Information
Distribution des degrés $k_i = \#\{j \neq i : w_{ij} > 0\}$ (deg. sortant) $h_i = \#\{j \neq i : w_{ji} > 0\}$ (deg. entrant)	Connectivité
Intensité d'un nœud $s_i = \sum_{j \in E(i) - \{i\}} w_{ij}$	Connectivité pondérée
Coefficient d'agglomération (pondéré) $c^w(i) = \frac{1}{s_i(k_i-1)} \sum_{j,h} \frac{w_{ij}+w_{ih}}{2} a_{ij}a_{jh}a_{hi}$ où $a_{\ell,m} = 1$ ssi $w_{\ell,m} > 0$ et $a_{\ell,m} = 0$ ssi $w_{\ell,m} = 0$	Densité locale
Distance moyenne $\text{avg}[\text{dist}(i,j) : (i,j) \in V^2]$	Proximité, transmission d'info.
Disparité $Y_2(i) = \sum_{j \neq i} \left(\frac{w_{ij}}{s_i}\right)^2$	Hétérogénéité locale
Appariement associatif $\text{knn} = \text{corr}(\text{deg}(i), \text{avg}_{j \in E(i)} \text{deg}(j))$	Similarité locale
Modularité $Q = \frac{1}{m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \frac{s_i s_j + 1}{2}$	Communauté de nœuds

Un autre aspect important de ces travaux sur les réseaux, en particulier de Barabási [BA99], est l'observation que la distribution du degré des nœuds est typiquement à «longue queue» où une petite fraction des nœuds, appelés «hubs», ont un très grand degré et où l'énorme majorité des nœuds ont une très faible connectivité en dessous de la moyenne. La distribution des degrés suit en première approximation une loi de puissance, autrement dit la probabilité qu'un nœud soit de degré k est proportionnelle à $k^{-\alpha}$, expérimentalement l'exposant α ayant souvent dans l'intervalle $[2, 3]$. Cette distribution se distingue d'une distribution modale comme la distribution gaussienne où la probabilité que le degré d'un nœud s'écarte beaucoup de la moyenne est quasiment nulle. Cette distribution typique donne le nom à ce type de réseau dit «sans échelle caractéristique»⁵. On trouve de tels réseaux sans échelle dans les réseaux de communication, le web, les réseaux de collaboration scientifique, les réseaux biochimiques, etc. [Wato4].

De nombreuses caractéristiques ont été proposées pour décrire ces grands réseaux. Il n'est pas possible de toutes les expliquer précisément dans ce document. Un tableau synthétique des principales mesures les plus utilisées et les plus robustes sont données dans le tableau 2, et l'on pourra se référer à [Bar15] pour plus de détails.

2.2 MODÉLISATION DES PAYSAGES DE FITNESS PAR LE RÉSEAU DES OPTIMA LOCAUX

Cette principale section contributive définit le réseau des optima locaux, analyse la structure de ce réseau pour des problèmes d'optimisation combinatoire, et montre plusieurs possibilités d'utilisation de ce réseau en particulier pour l'explication et la prédiction de performance de métaheuristiques.

5. «scale free» en langue anglaise

2.2.1 Réseau des optima locaux

Le réseau des optima locaux réduit l'espace de recherche aux optima locaux et définit des arcs de transition entre ceux-ci en prenant en compte la relation de voisinage du paysage de fitness et la dynamique de recherche des algorithmes d'optimisation. Initialement au début de ces recherches, nous avons beaucoup hésité à définir les nœuds du réseau comme étant soit les bassins d'attraction, soit les optima locaux. Finalement, comme la dynamique de nombreux algorithmes sur de nombreux problèmes combinatoires est principalement concentrée sur les optima locaux, nous avons choisi de définir les nœuds comme étant les optima locaux. Dans un premier temps jusqu'en 2010 [VOT08; Och+08; TVO08; VOT10], la définition des arcs a d'abord pris en compte les bassins d'attraction se rapprochant alors très fortement des travaux de Doye sur les réseaux intrinsèques. Ensuite depuis [Ver+11a], nous avons défini les arcs en lien direct avec les optima locaux ce qui permet de les estimer sans la connaissance des solutions des bassins d'attraction. Nous rappelons d'abord les définitions primordiales d'optimum local et nous précisons ensuite formellement les différentes définitions d'arc du réseau des optima locaux.

NŒUDS DU RÉSEAU. Un optimum local est une solution dont aucune solution voisine n'est meilleure. On peut distinguer un optimum local d'un optimum local strict où aucune solution n'est strictement meilleure. Un optimum global est un optimum local, mais pas nécessaire un optimum local strict :

Définition 4 *Optimum Local (maximisation).* Soit un paysage de fitness $(\mathcal{X}, \mathcal{N}, f)$. Un optimum local est une solution $x \in \mathcal{X}$ telle que $\forall y \in \mathcal{N}(x), f(y) \leq f(x)$.

Les optima locaux sont les points de convergence des marches adaptatives qui sont des séquences de solutions voisines dont la qualité augmente successivement. Les hill-climbers sont des recherches locales qui permettent de déterminer effectivement les optima locaux à travers une marche adaptative. Deux principales variantes d'hill-climbers existent qui diffèrent selon la règle pivot : un hill-climber *first-improvement* sélectionne l'une des solutions améliorantes selon une loi aléatoire uniforme et un hill-climber *best-improvement* sélectionne l'une des meilleures solutions voisines (selon une loi aléatoire uniforme en cas d'égalité). Les algorithmes 4 et 5 donne le détail des algorithmes. Lorsque plusieurs solutions voisines sont de même qualité (paysage avec neutralité), l'une des solutions est sélectionnée aléatoirement uniformément. Les algorithmes hill-climbers sont stochastiques puisque la solution voisine est sélectionnée aléatoirement même pour une variante best-improvement où le paysage peut être neutre avec des solutions de même qualité dans le voisinage. Par la suite, nous noterons par $hc(x)$ un optimum local obtenu à partir d'une solution x par un hill-climber.

Algorithm 4 Hill-climber best-improvement.

```

1: function HILLCLIMBING_BEST_IMPROVEMENT( $x$  : solution initiale)
2:   while  $x \neq$  Optimum Local do
3:     Sélectionner  $x' \in \mathcal{N}(x)$ , t.q.  $f(x') = \max_{y \in \mathcal{N}(x)} f(y)$ 
4:     if  $f(x) < f(x')$  then
5:        $x \leftarrow x'$ 
6:     end if
7:   end while
8:   return( $x$ )
9: end function

```

Les nœuds du réseau d'optima locaux sont généralement les optima locaux du paysage de fitness. Néanmoins, lorsque le paysage de paysage est neutre, nous avons défini dans [VOT10] les nœuds comme étant les réseaux de neutralité optimum local, c'est-à-dire les plateaux de solutions qui sont reliés pas la relation de voisinage et dont les solutions sont

Algorithm 5 Hill-climber first-improvement.

```

1: function HILLCLIMBING_FIRST_IMPROVEMENT( $x$  : solution initiale)
2:   while  $x \neq$  Optimum Local do
3:     Sélectionner  $x' \in \mathcal{N}(x)$  aléatoirement (sans remise)
4:     if  $f(x) < f(x')$  then
5:        $x \leftarrow x'$ 
6:     end if
7:   end while
8:   return( $x$ )
9: end function

```

tous optima locaux. Les espaces de recherche combinatoire étant fini, le nombre d'optima locaux également, les nœuds du réseau sont notés $\{lo_1, lo_2, \dots, lo_n\}$ que ce soit des optima locaux, ou par un abus de notation simplificateur, des réseaux de neutralité optima locaux.

ARCS DU RÉSEAU. Deux types d'arcs ont pu être définis, l'un basé sur la notion de bassin d'attraction, l'autre basé sur la possibilité d'échapper à un optimum local.

Un bassin d'attraction est un ensemble des solutions menant à un même optimum local. Notons $p_i(x)$ la probabilité qu'une solution x converge vers un nœud i : $p_i(x) = \Pr(\text{hc}(x) = lo_i)$.

Définition 5 *bassin d'attraction.* Le bassin d'attraction d'un nœud i est l'ensemble $b_i = \{x \in \mathcal{X} : p_i(x) > 0\}$. La taille du bassin d'attraction est alors $\sum_{x \in \mathcal{X}} p_i(x)$.

La figure 8 donne une représentation de ces définitions lorsque le hill-climber est la variante best-improvement, et que pour chaque solution, il n'y a qu'une solution voisine de fitness maximum. En effet dans ce cas classique, chaque solution n'appartient qu'à un seul bassin d'attraction, ceux-ci définissent alors une partition de l'espace de recherche comme pour le réseau intrinsèque de Doye.

La définition des arcs et des poids pour la version des arcs basés sur les bassins, *arc-bassin*, consiste à agréger la relation de voisinage \mathcal{N} initiale. Il existe un arc entre deux bassins lorsqu'ils existent deux solutions voisines une dans chaque bassin d'attraction. Les poids sont proportionnels à la probabilité qu'à partir d'une solution d'un bassin l'on atteigne une solution de l'autre bassin. Plus précisément [Och+08], il existe un arc (i, j) entre les nœuds i et j si et seulement s'il existe deux solutions $x_i \in b_i$ et $x_j \in b_j$ telles que $x_j \in \mathcal{N}(x_i)$. Pour définir les poids, notons par $p(x \rightarrow y)$ la probabilité qu'à partir d'une solution x d'atteindre la solution y . Par exemple, on pourra définir cette probabilité par une distribution uniforme lorsque le voisinage est de taille finie par $p(x \rightarrow y) = 1/\#\mathcal{N}(x)$ si $y \in \mathcal{N}(x)$, et $p(x \rightarrow y) = 0$ si $y \notin \mathcal{N}(x)$. La probabilité qu'une solution x ait un voisin dans le bassin b_j est alors $p(x \rightarrow b_j) = \sum_{y \in b_j} p(x \rightarrow y)p_j(y)$. Le poids w_{ij} de l'arc entre les nœuds i et j est défini comme la moyenne de toutes ces probabilités pondérées par la probabilité d'appartenir au bassin i : $w_{ij}^b = p(b_i \rightarrow b_j) = \frac{1}{\#b_i} \sum_{x \in b_i} p_i(x)p(x \rightarrow b_j)$.

Les arcs basés sur la possibilité d'échapper aux optima locaux [Ver+11a], *arc-échappement*, sont définis à partir d'une fonction distance d entre les solutions et d'un nombre entier strictement positif D . Par exemple, la fonction distance peut être le nombre minimal de hop entre deux solutions du graphe $(\mathcal{X}, \mathcal{N})$ défini par le paysage de fitness, ou encore la distance de Hamming dans le cas des espaces de recherche composés des chaînes binaires. Il existe un arc entre deux nœuds i et j , $(i, j) \in E^e$, lorsqu'il existe une solution y à une distance D de lo_i et que depuis cette solution le hill-climber permet d'atteindre l'optimum local lo_j : $d(lo_i, y) \leq D$ et que $p_j(y) = \Pr(\text{hc}(y) = lo_j) > 0$. Le poids d'un arc entre i et j est la moyenne sur toutes les solutions à une distance D de lo_i de la probabilité d'atteindre l'optimum local lo_j : $w_{ij}^e = 1/\#S(i) \sum_{y \in S(i)} p_j(y)$ avec $S(i) = \{y \in \mathcal{X} : d(lo_i, y) \leq D\}$.

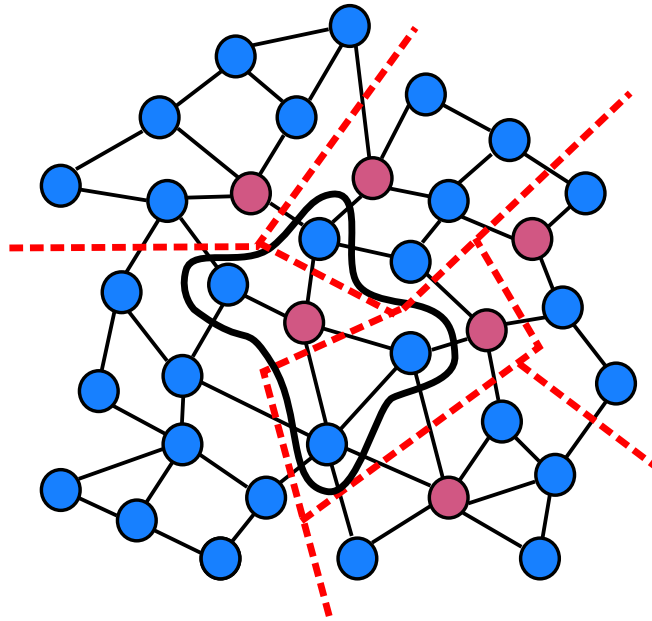


FIGURE 8 – Illustration des bassins d'attraction d'un paysage de fitness. Les solutions du problème combinatoire sont représentées par des disques. Ceux de couleur rouge sont des optima locaux. Les traits noirs les reliant représentent la relation de voisinage. Les traits en pointillés rouges délimitent les bassins d'attractions. Cette représentation est proche de celle de Doye [Doy02] couramment utilisée. Pourtant cette représentation est souvent trompeuse pour les problèmes combinatoires de grande dimension où les bassins ne sont pas des ensembles «compacts» avec une frontière en forme d'une ligne simple.

On peut maintenant définir le réseau des optima locaux. La figure 9 montre un exemple d'un tel réseau pour un problème combinatoire de paysage NK. La section suivante propose les résultats de l'analyse de tels réseaux.

Définition 6 Réseau des Optima Locaux (ROL). Le réseau des optima local $G_w = (N, E)$ est le graphe où les nœuds sont les solutions (ou les réseaux de neutralité) optima locaux $\{lo_1, lo_2, \dots, lo_n\}$. Les arcs avec leur pondération sont E^b et (w_{ij}^b) dans le cas des arcs-bassins, et E^e et (w_{ij}^e) dans le cas des arcs-échappements selon les définitions données précédemment.

2.2.2 Nouvelles propriétés des paysages de fitness

Les premières analyses ont cherché à montrer la structure générale des réseaux des optima locaux pour les problèmes combinatoires. Pour cela, nous avons employé le paysage NK [Kau93] qui est un problème indépendant d'un problème combinatoire particulier, mais qui possède des propriétés suffisamment pertinentes pour être très informatif sur de nombreux autres problèmes combinatoires. Un aspect très intéressant de ce problème est qu'il permet de régler le degré de non-linéarité du problème (épistasie) qui est l'un des facteurs majeurs de difficulté d'optimisation. De plus, ce problème NP-complet [Wei+96] est utilisé comme modèle en biologie évolutive [Orro5], en théorie des organisations [RS02c] et bien sûr en optimisation. En optimisation, les généralisations immédiates du problème peuvent coder les problèmes classiques (maxSAT, modèle Ising, etc.) et les problèmes dont la décomposition en fonction de Walsh est de degré inférieur à une constante [HW97].

L'espace de recherche d'un paysage NK est l'ensemble des chaînes binaires de longueur N . Originellement, chaque élément de la chaîne représentait la valeur d'un gène. La fonction de fitness Φ est la moyenne de fonctions de contribution Φ_i qui dépendent de $K + 1$ variables

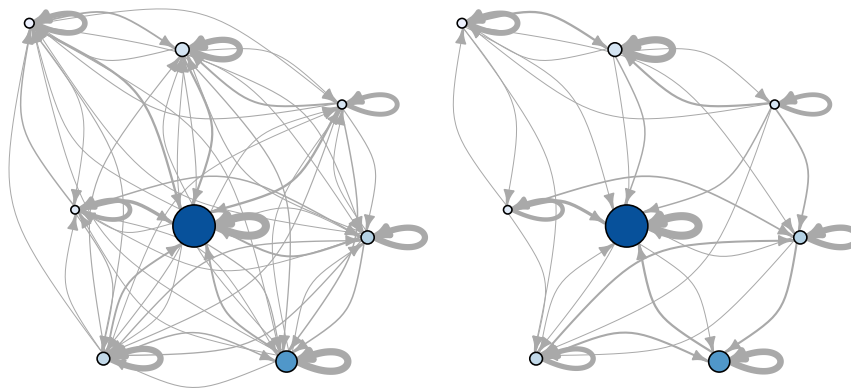


FIGURE 9 – Exemple de réseau des optima locaux avec des arcs basés sur les bassins (à gauche) et avec des arcs basés sur l'échappement (à droite) pour un paysage NK avec $N = 18$ et $K = 2$. La taille des nœuds est proportionnelle à la taille des bassins. La couleur représente la fitness : les valeurs de fitness les meilleures sont de couleurs plus sombres. L'épaisseur des arcs est proportionnelle au poids de chaque arc, c.-à-d. la probabilité de transition entre optima locaux.

distinctes : $\Phi(x) = \frac{1}{N} \sum_{i=1}^N \phi_i(x_i, x_{i_1}, \dots, x_{i_k})$. Les familles d'entier (i_k) définissent les dépendances entre les variables, appelé aussi liens épistatiques : pour chaque variable i , la fonction Φ_i dépend de la variable i et de K autres variables telles que $\{i, i_1, \dots, i_k\} \subset \{1, \dots, N\}$. En augmentant le paramètre K de 0 à $N - 1$, le paysage NK passe de linéaire à complètement aléatoire. Dans la suite, les liens épistatiques sont aléatoires choisis sans remise⁶. Les fonctions de contribution sont en généralement définies aléatoirement : chaque valeur de la fonction est définie par un nombre réel tiré aléatoirement entre 0 et 1.

Le premier pas de ces travaux consiste à analyser la topologie générale du réseau des optima locaux (ROL), par exemple en la comparant à celle d'un graphe aléatoire. Ensuite, la question est de savoir si la topologie du ROL reflète bien la difficulté d'optimisation du problème. Ici, le protocole expérimental consiste à étudier des petites instances du paysage NK où l'on peut extraire de façon exacte l'ensemble du ROL, puis à calculer les mesures classiques des réseaux complexes sur un grand nombre d'instances pour obtenir des statistiques robustes sur une classe de paysage NK. La critique récurrente du problème des paysages NK est de ne pas correspondre à un «vrai» problème combinatoire puisque par construction il semble très aléatoire. Au contraire, on pourrait plutôt le considérer comme un problème d'optimisation «moyen» représentatif. Néanmoins pour écarter cette critique, nous avons pu aussi analyser les ROL d'autres problèmes d'optimisation combinatoire plus classiques en recherche opérationnelle où les propriétés sont similaires aux paysages NK. De plus, nous avons pu montrer que la performance de métaheuristique peut être prédite à l'aide du ROL.

STRUCTURE GÉNÉRALE DU RÉSEAU DES OPTIMA LOCAUX. La structure du réseau des optima locaux n'est pas celle d'un graphe aléatoire. La figure 10 montre trois mesures de réseau complexe des instances de petite dimension de paysages NK [VOT08; TVO08; Och+14]. Tout d'abord, la distribution des poids sortant du réseau n'est pas géométrique comme

6. Une autre possibilité courante est d'utiliser des liens adjacents, mais dans ce cas un algorithme de programmation dynamique de complexité $\mathcal{O}(2^{K+1}N)$ existe.

celle attendue d'un graphe aléatoire pondéré [Garoga]. La distribution montre une queue longue, c'est-à-dire que de nombreux nœuds ont des connexions de poids importants. De même, le coefficient d'agglomération pondéré (cf. tableau 2) qui mesure la connectivité locale du réseau est bien plus important que celui d'un graphe aléatoire. Les nœuds du ROL sont plus agglomérés que dans un graphe aléatoire. Enfin, la disparité qui mesure l'hétérogénéité des distributions des poids sur un même nœud est là encore plus hétérogène que pour un graphe aléatoire. Pour donner une image synthétique et simple sur les paysages NK, les réseaux des optima locaux ont une distribution des poids hétérogènes au niveau local autour de chaque nœud et au niveau global où les poids importants existent. De plus, l'agglomération des connexions est forte comme dans un réseau complexe de type petit-monde.

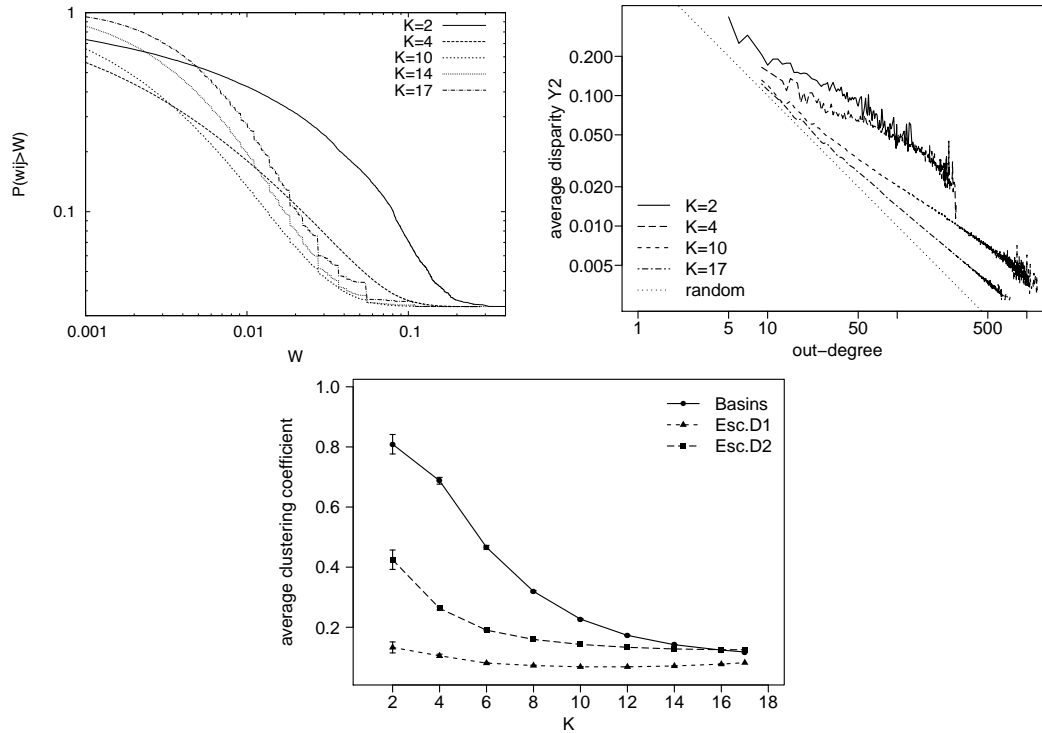


FIGURE 10 – Mesures de type réseau complexe sur le paysage NK avec $N = 18$ estimées sur 30 instances pour chacune d'elle. En haut à gauche, distribution de probabilité cumulative des poids sortants $w_{ij, j \neq i}$ du réseau dont les arcs sont des arcs-bassins. À noter, l'échelle logarithmique en abscisse qui montre ainsi une distribution à longue queue. En haut à droite, disparité pondérée en fonction du degré sortant des nœuds pour les mêmes réseaux des optima locaux. La ligne pointillée ("random") représente la valeur moyenne pour un graphe aléatoire : $Y_2 = 1/k$, où k est le degré sortant. Les valeurs sont donc plus hétérogènes que celle d'un graphe aléatoire. En bas, coefficient d'agglomération moyen en fonction du degré d'épistasie K (non-linéarité) pour différent type d'arc. (*basins* : basé sur les bassins d'attraction, *esc.D1* et *esc.D2* basés sur l'échappement à une distance $D = 1$ et $D = 2$).

LA STRUCTURE REFLÈTE LA DIFFICULTÉ. Un deuxième point important des travaux montre que la structure du réseau des optima locaux est corrélée avec la difficulté d'optimisation. Pour les paysages NK, le degré d'épistasie K , c'est-à-dire la non-linéarité du problème, qui constitue le facteur principal de difficulté, influence la structure du ROL. Comme attendu, le nombre d'optima locaux augmente avec ce degré de non-linéarité K , mais aussi la densité du ROL diminue avec celui-ci (voir tableau 3). Les trois mesures de poids moyen des arcs, de

TABLE 3 – Moyenne et écart-type du nombre de nœuds et de la densité d’arcs du réseau pour les deux types d’arcs basés sur les bassins d’attraction («Bassin») et sur l’échappement à distance $D = 1$ et $D = 2$ (respectivement Esc.D1 et Esc.D2). Les réseaux basés sur l’échappement offrent une plus faible densité et ces arcs ne nécessitent pas l’énumération complète des bassins pour être calculés.

κ	nb. nœuds	densité des arcs (%)		
		Bassin	Esc.D1	Esc.D2
2	43.0 _{27.7}	74.182 _{13.128}	8.298 _{4.716}	22.750 _{9.301}
4	220.6 _{39.1}	54.061 _{4.413}	1.463 _{0.231}	7.066 _{0.810}
6	748.4 _{70.2}	26.343 _{1.963}	0.469 _{0.047}	3.466 _{0.279}
8	1668.8 _{73.5}	12.709 _{0.512}	0.228 _{0.009}	2.201 _{0.066}
10	3147.6 _{109.9}	6.269 _{0.244}	0.132 _{0.004}	1.531 _{0.036}
12	5270.3 _{103.9}	3.240 _{0.079}	0.088 _{0.001}	1.115 _{0.015}
14	8099.6 _{121.1}	1.774 _{0.035}	0.064 _{0.001}	0.838 _{0.009}
16	11688.1 _{101.3}	1.030 _{0.013}	0.051 _{0.000}	0.647 _{0.004}
17	13801.0 _{74.1}	0.801 _{0.007}	0.047 _{0.000}	0.574 _{0.002}

coefficient d’agglomération et de disparité varient avec le degré de non-linéarité (voir figures 10 et 11). La structuration du réseau est plus importante lorsque le degré d’épistasie est faible. Au fur et à mesure que la non-linéarité du problème augmente, le réseau a tendance à devenir moins structuré et plus proche d’un graphe aléatoire. Les deux types de définitions d’arc, arcs-bassin et arcs-échappement, ont les mêmes tendances.

Les poids des arcs peuvent être interprétés comme la probabilité de passer d’un optimum local à un autre. L’inverse du poids $1/w_{ij}$ peut donc s’interpréter comme une distance, le nombre moyen de tentatives pour pouvoir passer d’un optimum local à un autre. La distance moyenne à l’optimum global sur le ROL (voir figure 11) donne une indication importante quant à la difficulté d’optimisation. En effet, cette distance augmente avec le degré de non-linéarité justifiant que les définitions d’arcs reflètent bien la difficulté d’optimisation.

La définition du réseau des optima locaux a été étendue aux paysages de fitness neutre dont beaucoup de solutions voisines sont de même performance [VOT10]. Également dans ce cas, la structure du réseau reflète la difficulté. Parmi d’autres résultats, la figure 12 montre la disparité et la distance moyenne à l’optimum global sur le réseau. Pour les variantes neutres des paysages NK (NKq et NKp), la neutralité réduit l’effet de la multimodalité due à la non-linéarité du problème : plus le problème est neutre, plus la structure du réseau des optima locaux devient plus favorable et le problème d’optimisation plus facile à résoudre, comme avec une multimodalité moindre.

BASSINS D’ATTRACTION. Les solutions d’un même bassin d’attraction se projettent sur le même nœud du réseau des optima locaux. Les bassins d’attraction sont à la «petite» échelle du système mise en œuvre par l’agrégation de variables du réseau des optima locaux. La forme des bassins eux-mêmes est une caractéristique importante des paysages de fitness.

La figure 13 montre un exemple de distribution des tailles des bassins. Celle-ci n’est pas une distribution uniforme. Pour les paysages NK, la distribution est log-normale pour les degrés de non-linéarité plus petits que $N/2$ [VOT10]. La variance des tailles est d’autant plus petite que le problème devient non-linéaire. Lorsque le problème devient de plus en plus aléatoire, la taille des bassins s’écarte peu d’une moyenne faible; et au contraire, lorsque le problème est de plus en plus linéaire, il existe de plus en plus de grands bassins d’attraction dans la distribution. Contrairement peut-être à une intuition première, les plus grands bassins d’attraction sont ceux relatifs aux meilleurs optima locaux. En effet, la taille des bassins est corrélée positivement avec la fitness des optima locaux (voir figure 13). Néanmoins, les optima locaux de bonne qualité sont peu fréquents et la taille de leur bassin d’attrac-

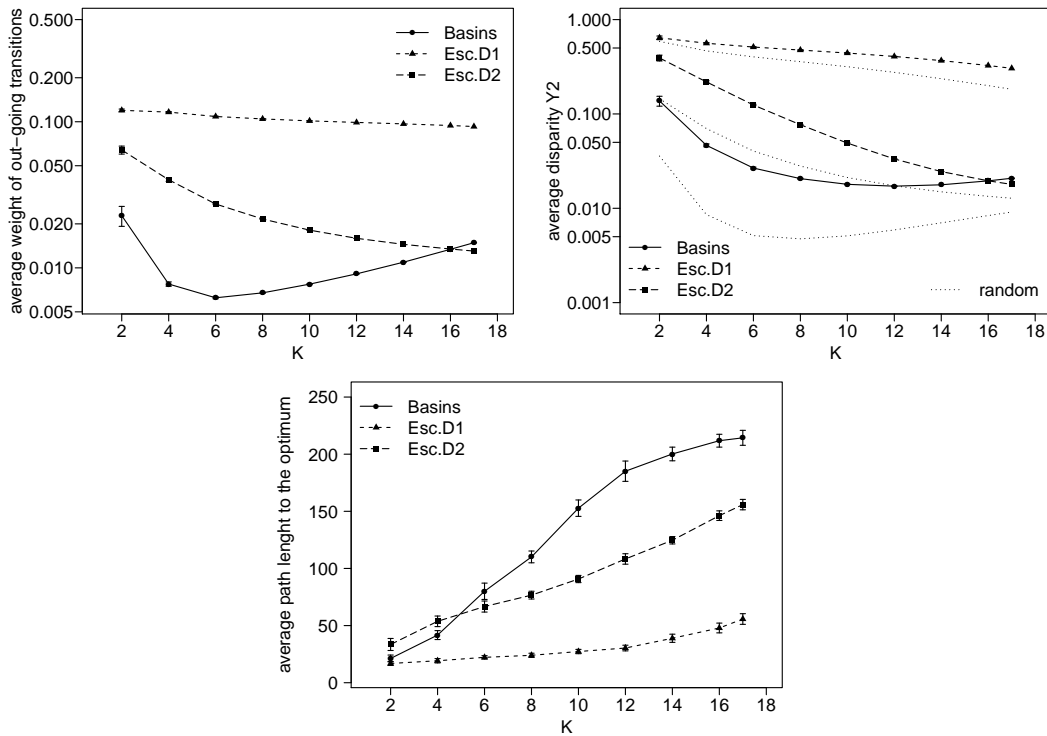


FIGURE 11 – Paysages NK pour $N = 18$: de haut en bas, poids des degrés sortant (strength), coefficient d'agglomération moyen et distance moyenne à l'optimum global pour les différents types d'arcs (bassin et échappement à une distance 1 et 2). Estimations sur 30 instances.

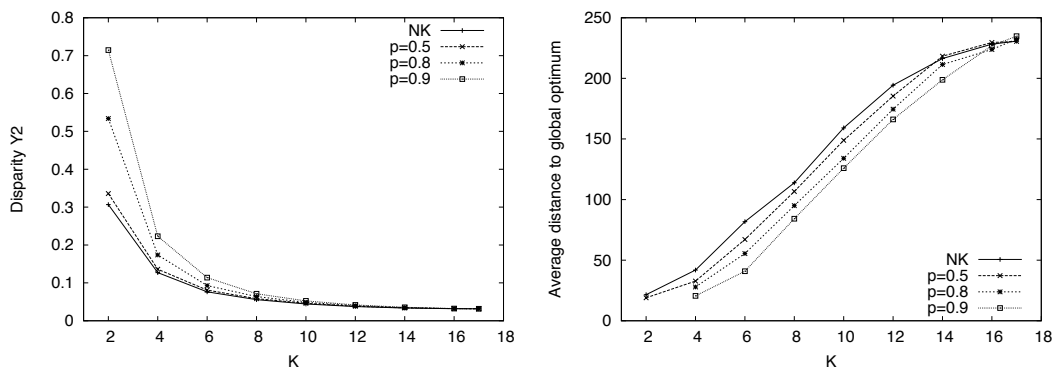


FIGURE 12 – Variante neutre NK_p des paysages NK pour $N = 18$, plus le paramètre p est proche de 1, plus le problème est neutre. À gauche, la disparité moyenne et à droite la distance moyenne à l'optimum global. Estimations sur 30 instances. Un plus grand degré de neutralité augmente la disparité (certains nœuds ont une plus grande probabilité d'être obtenus) et diminue la distance à l'optimum global.

tion est exponentiellement petite par rapport à la taille de l'espace de recherche et diminue exponentiellement avec la non-linéarité du problème.

La forme des bassins d'attraction n'est pas celle reproduite schématiquement dans la figure 8. La figure 14 donne la proportion de solutions dont toutes les solutions voisines sont également dans le même bassin. Ainsi ces solutions peuvent être considérées pleinement à l'intérieur du bassin. Les solutions sont ultra-majoritairement à la frontière des bassins d'attraction, même si pour chacune des solutions une proportion significative de solutions

voisines appartiennent au même bassin. Les études menées ici montrent qu'il vaut mieux imaginer une représentation comme celle de la figure 15 qui donne un exemple concret de paysage de fitness pour $N = 6$.

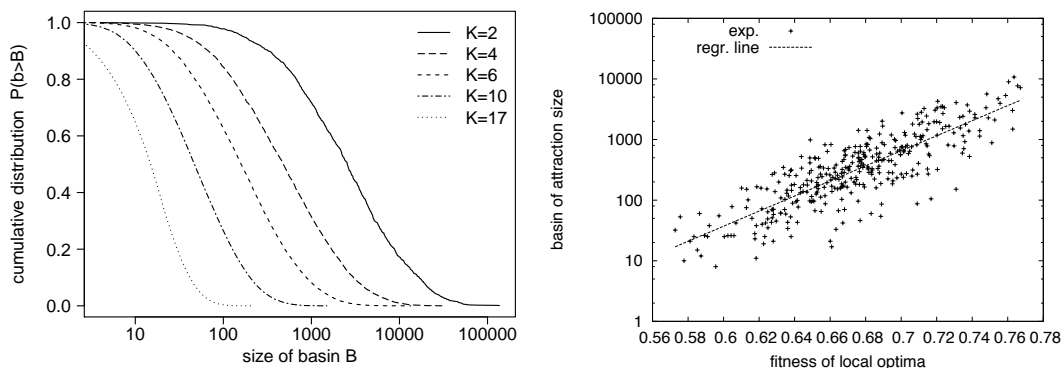


FIGURE 13 – Paysages NK avec $N = 18$: à gauche distribution cumulative des tailles des bassins d'attraction pour différents degrés de non-linéarité K (abscisse en échelle log) et à droite, exemple de corrélation entre la taille des bassins et la fitness de maxima locaux pour un paysage particulier où $K = 4$.

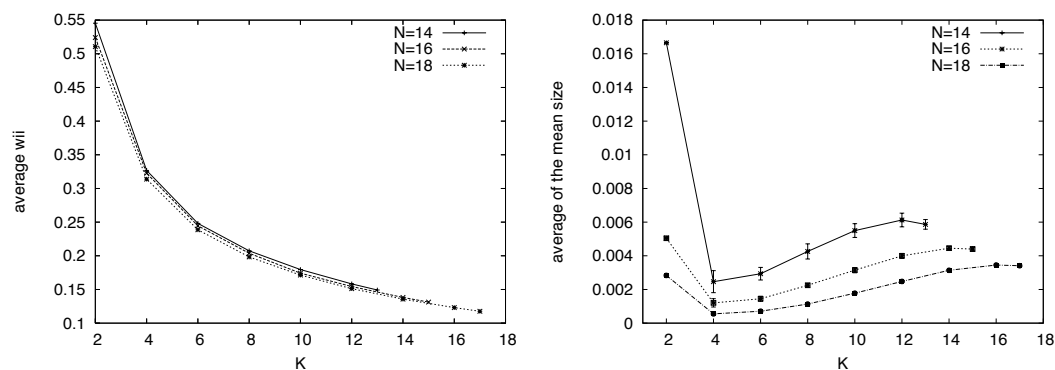


FIGURE 14 – Paysages NK pour $N = 14, 16$ et 18 et différentes valeurs de K : à gauche moyenne des poids w_{ii} des boucles donnant une estimation de la probabilité de rester dans le même bassin ; à droite, proportion moyenne des nœuds «intérieurs» dont tous les voisins sont dans le même bassin d'attraction.

2.2.3 Analyses basées sur le réseau des optima locaux

Dans les finalités identifiées des paysages de fitness à la section 1.2.2 et synthétisées par la figure 5, l'une des situations classiques de l'utilisation des paysages de fitness consiste à comparer de manière off-line soit différents types d'instance de problème ou différentes options de conception d'algorithme afin de sélectionner celle qui est la plus pertinente à la structure du problème. Nous mettons en œuvre cette méthodologie à l'aide du réseau des optima locaux pour comparer différentes instances du problème d'affectation quadratique, différent opérateur de voisinage pour le problème du flow-shop et enfin différentes règles pivot d'un hill-climber.

COMPARAISON D'INSTANCES DU PROBLÈME D'AFFECTATION QUADRATIQUE Le réseau des optima locaux permet aussi d'étudier d'autres problèmes combinatoires qui sont sans doute

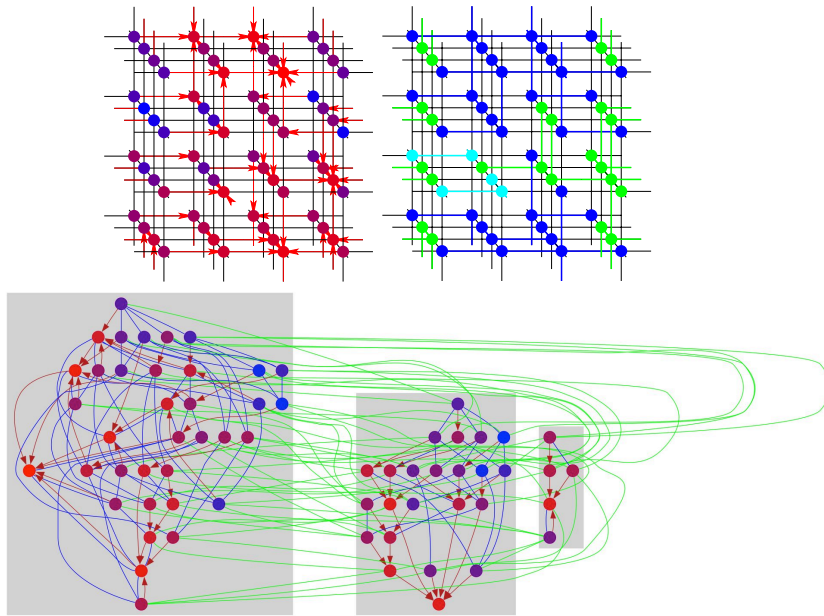


FIGURE 15 – Exemple de paysage NK pour $N = 6$. Les solutions sont les nœuds. En haut, les 64 solutions sont reliées par un trait représentant la relation de voisinage, les extrémités étant reliées aussi entre elles. À gauche, les couleurs sont la fitness (maximum en rouge et minimum en bleu) et les flèches indiquent le meilleur voisin. À droite, les couleurs indiquent les 3 bassins d’attraction qui sont entremêlés. En bas, le paysage en représentation planaire avec les 3 bassins d’attraction, les liens montrent les nombreuses relations de voisinage entre les solutions où quasiment toutes les solutions sont à la frontière. Ici, le bassin le plus grand est celui de l’optimum global.

plus courants dans la littérature en recherche opérationnelle. Le problème d’affectation quadratique (QAP) [KB57] est un problème de minimisation NP-complet dont l’espace de recherche est l’ensemble des permutations de longueur n et la fonction de fitness est donnée par : $f_{\text{QAP}}(\pi) = \sum_{ij} a_{ij} b_{\pi_i \pi_j}$. Selon les matrices (a_{ij}) et (b_{ij}) , les instances sont de difficultés différentes. Le réseau des optima locaux permet de comprendre et de quantifier la différence de difficulté entre deux types d’instances classiques l’une «uniforme» dont les coefficients de matrices sont uniformément répartis dans un intervalle, et l’autre «real-like» dont la distribution des coefficients imite celle d’instances de problèmes réels [Tai95].

La figure 16 montre quelques résultats des mesures sur le ROL [Dao+10] calculés sur les deux types de petites instances de QAP ($n = 10$) décrites ci-dessus. La structure globale de ces réseaux ne diffère pas des observations faites sur les paysages NK : distribution des poids est à longue queue, disparité des poids est importante, agglomération est importante, corrélation significative existe entre fitness et taille des bassins d’attraction, etc. Les instances real-like sont connues pour être plus faciles à résoudre que les instances uniformes. On constate aussi ce résultat à travers les métriques du ROL. Pour les instances real-like, l’agglomération est plus importante, la distance sur le ROL à l’optimum global est plus courte, la disparité est plus marquée, etc. Comme supposé, les instances real-like sont plus structurées. Nous avons pu montrer que les optima locaux des instances real-like peuvent être regroupés en communauté d’optima locaux [Dao+11]. La figure 17 montre visuellement les communautés d’optima locaux locaux d’instances real-like et uniforme, et la figure 16 donne la modularité des partitions correspondantes. La stratégie de recherche sur les instances real-like semble plus facile, car les optima locaux proches au sens de la distance sur le réseau sont aussi des optima locaux de qualité proche.

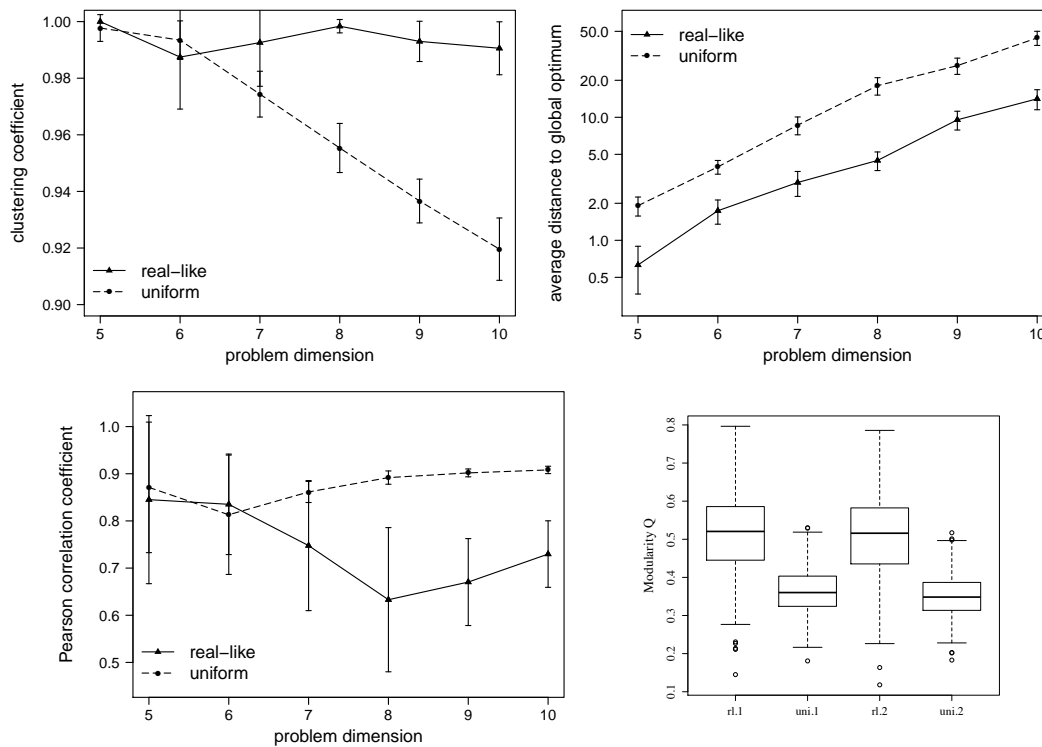


FIGURE 16 – Mesures de type réseaux complexes sur le problème QAP estimées sur 30 instances. Coefficient d’agglomération, distance moyenne à l’optimum global sur le ROL, coefficient de corrélation (Pearson) entre le logarithme de la taille des bassins d’attraction et la fitness, modularité des réseaux après détection de communauté.

COMPARAISON D’OPÉRATEUR POUR LE PROBLÈME FLOW-SHOP. L’analyse de paysage de fitness à l’aide du réseau des optima locaux a été aussi appliquée au problème d’optimisation combinatoire de flow-shop. Brièvement, le problème flow-shop consiste à ordonnancer n jobs qui doivent passer par le traitement de m machines successives. Deux jobs ne peuvent pas être traités en même temps sur une même machine, les machines et les traitements sont «à la chaîne» et le temps de traitement de la tâche du job i sur la machine j est donné par un paramètre t_{ij} . Dans sa version mono-objectif, le problème consiste à trouver l’ordonnement des tâches pour minimiser le temps de traitement total (*makespan*) [GJS76].

La figure 18 montre quelques métriques relatives au réseau des optima locaux pour les arcs-échappement [Dao+13] sur des instances uniformes de Taillard. Comme pour les autres problèmes d’optimisation combinatoire, le réseau est fortement aggloméré, la disparité des connexions est forte, la distribution des poids sortants montre une large queue. Plus le nombre de machines est important, plus le problème d’optimisation devient difficile à résoudre pour les métaheuristiques. La structure du ROL reflète cette difficulté, les principales mesures montrent que le réseau perd de sa structure au fur à mesure de l’augmentation du nombre de machines. L’analyse du ROL permet aussi de décider d’un choix d’opérateur de recherche locale. En effet, deux opérateurs classiques sont possibles : l’opérateur d’échange (*exchange*) permute l’ordre de deux jobs et l’opérateur d’insertion (*insertion*) intercale un job parmi d’autres. La structure du ROL est plus favorable pour l’opérateur d’insertion que pour l’opérateur d’échange ce qui permet de faire une sélection parmi ces deux opérateurs et de justifier également la pertinence du ROL dans l’analyse de paysage de fitness.

COMPARAISON DE RÈGLES PIVOTS. Le réseau des optima locaux permet aussi de mieux comprendre la différence entre deux règles pivots first-improvement et best-improvement

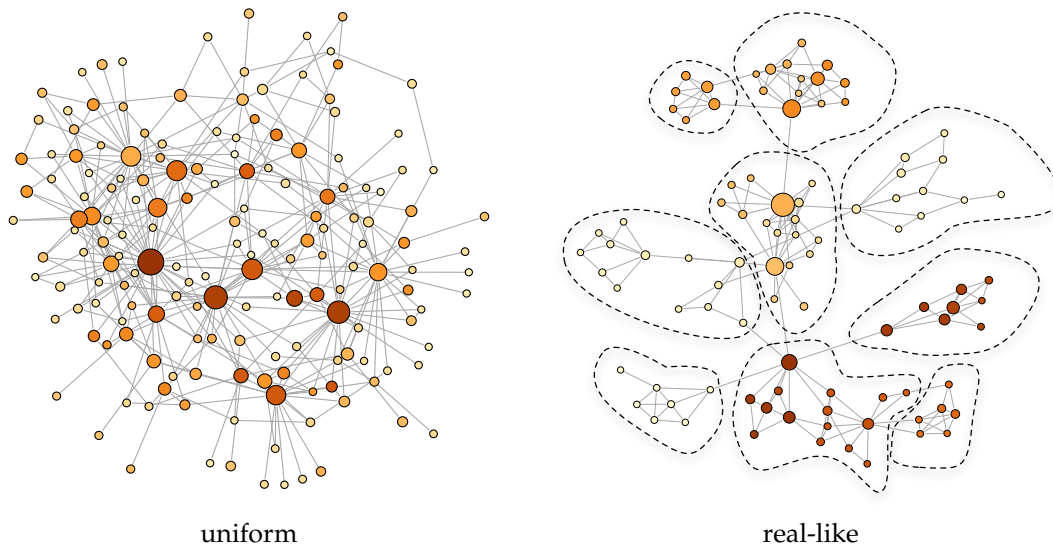


FIGURE 17 – Représentation des communautés sur le ROL pour deux instances de QAP. Les nœuds sont les optima locaux, les couleurs correspondent aux valeurs de fitness (les plus rouges sont les meilleures valeurs), et la taille correspond à la taille des bassins d’attraction correspondants. Les communautés sont détectées par l’algorithme de Clauset *et al.* [CNM04] et regroupent des nœuds de manière à maximiser la modularité, c’est-à-dire à trouver une partition où les connections intra-communautés sont importantes et celles inter-communautés faibles. Les instances real-like font apparaître des communautés d’optima locaux homogènes en qualité d’optima locaux.

couramment utilisés dans les hill-climbers (voir section 2.2.1). Deux types de ROL avec des arcs-échappements ont été obtenus. Les uns utilisent un hill-climber best-improvement pour obtenir les nœuds et surtout les arcs (b-LON) et les autres utilisent un first-improvement (f-LON) [OVT10]. Le tableau 4 donne les principales caractéristiques ROL dans ces deux cas sur les paysages NK.

Les résultats peuvent être résumés ainsi. L’impact de la rugosité (paramètre K d’épistasie) sur les caractéristiques du ROL est similaire pour les deux types de ROL. Le first-improvement induit un réseau dense (en fait complètement connecté) alors que ce n’est pas le cas pour le best-improvement. Toutefois, de nombreux arcs dans le cas first-improvement ont des poids très faibles. En particulier, le poids des boucles (probabilité de revenir sur le même optimum local) est bien plus faible pour le first-improvement que pour le best-improvement (voir figure 19). La longueur du chemin vers l’optimum global sur le réseau est plus grande pour le first-improvement, mais il faut considérer qu’un hill-climber first improvement demande N fois moins d’évaluations à chaque pas. Ces propriétés des ROLs suggèrent que le first-improvement est une meilleure heuristique pour explorer le paysage de fitness du paysage NK.

2.2.4 Explication de performance des métaheuristiques

Dans les buts de l’analyse de paysage de fitness (voir section 1.2.2 et figure 5), l’un des défis est la prédiction de performance à partir des propriétés du paysage de fitness. Dans cette section, nous examinons la possibilité d’explication des performances à partir des caractéristiques du réseau des optima locaux par un modèle de régression dont les prédicteurs sont sur les métriques du ROL.

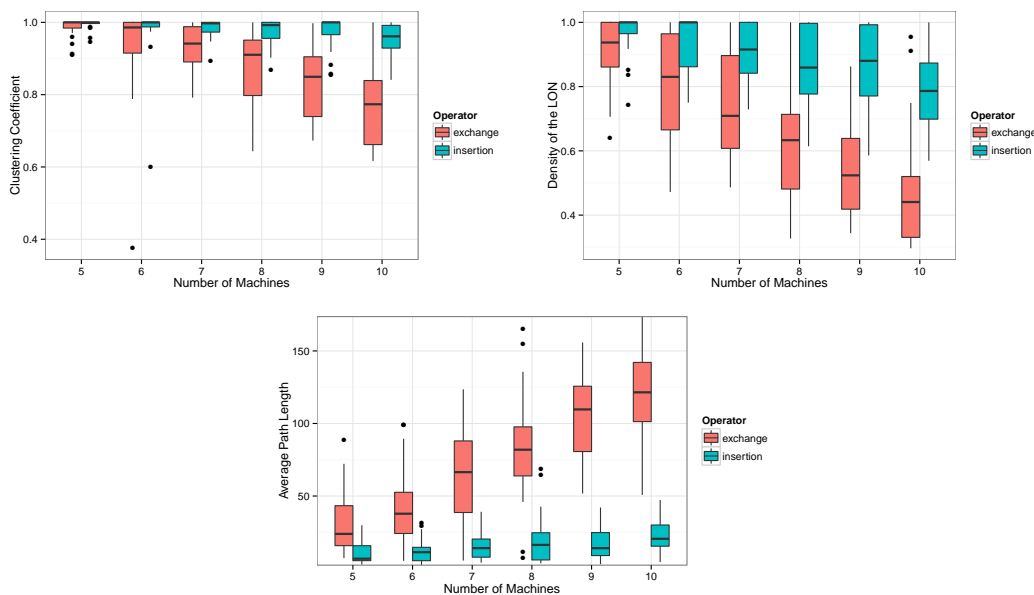


FIGURE 18 – Métriques extraites du réseau des optima locaux sur le problème flow-shop estimées sur 30 instances avec $n = 10$ jobs : coefficient d’agglomération, disparité moyenne et distance moyenne à l’optimum global sur le ROL. Deux opérateurs classiques sont comparés : insertion et échange de tâches. Le nombre de machines m est un paramètre dont il est connu qu’un plus grand nombre de machines augmente la difficulté du problème d’optimisation. Les arcs du ROL sont des arcs-échappement calculés avec une distance $D = 2$.

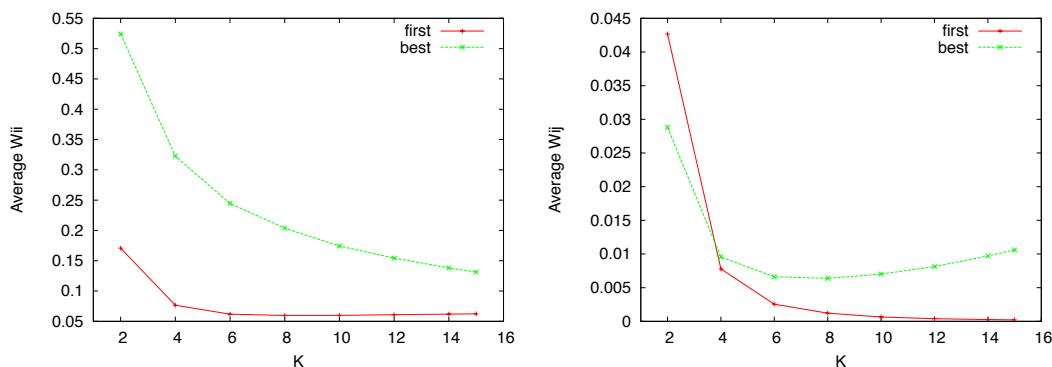


FIGURE 19 – Moyenne des poids des boucles w_{ii} (gauche) et des poids sortants w_{ij} (droite) sur le ROL pour les paysages NK avec les variantes first-improvement et best-improvement.

Le premier travail s’inscrit le cadre des paysages NK [Dao+12]. La métaheuristique dont les performances sont à expliquer est une recherche locale itérée, ou Iterated Local Search (ILS), qui consiste en l’itération d’une perturbation qui modifie $k = 2$ bits aléatoirement suivi d’un hill-climbing best-improvement. Comme mesure de performance, nous adoptons le nombre d’évaluations moyen (*expected run-length*) pour atteindre l’optimum global en simulant des redémarrages indépendants [AH05] : $erl = \mathbb{E}(T_s) + \left(\frac{1-p_s}{p_s}\right) T_{max}$ où T_s est le nombre d’évaluation lorsque l’optimum global est atteint, p_s le taux de succès de l’algorithme et T_{max} le nombre d’évaluations maximums du critère d’arrêt de l’algorithme. Tous les détails expérimentaux sont donnés dans l’article [Dao+12].

La figure 20 montre quelques nuages de points et les coefficients de corrélation entre les métriques du ROL et le logarithme de la performance. Excepté pour la distance moyenne

TABLE 4 – Propriétés du réseau des optima locaux pour les paysages NK avec $N = 16$. Les valeurs sont les moyennes sur 30 instances. n_v et n_e représentent le nombre de nœuds et d’arcs, \bar{C}^w , la moyenne du coefficient d’agglomération pondéré, \bar{Y} la disparité moyenne, \bar{d} la distance moyenne entre les nœuds, et \bar{d}_{best} la distance moyenne à l’optimum global (pour mémoire, la distance entre i et j est $1/w_{ij}$).

K	\bar{n}_v	\bar{n}_e/\bar{n}_v^2		\bar{C}^w		\bar{Y}		\bar{d}		\bar{d}_{best}	
		both	b-LON	f-LON	b-LON	b-LON	f-LON	b-LON	f-LON	b-LON	f-LON
2	33	0.81	1.00	0.96	0.326	0.110	56	39	16	12	
4	178	0.60	1.00	0.92	0.137	0.033	126	127	35	32	
6	460	0.32	1.00	0.79	0.084	0.016	170	215	60	70	
8	890	0.17	1.00	0.65	0.062	0.011	194	282	83	118	
10	1,470	0.09	1.00	0.53	0.050	0.009	206	340	112	183	
12	2,254	0.05	1.00	0.44	0.043	0.008	207	380	143	271	
14	3,264	0.03	1.00	0.38	0.040	0.008	203	411	158	351	
15	3,868	0.02	1.00	0.35	0.039	0.008	200	423	162	391	

entre nœuds, les corrélations entre les métriques et la performance sont fortes. Cela suggère que les métriques contiennent de l’information en rapport avec la mesure de performance. Nous avons ajouté deux autres métriques : la corrélation entre la fitness d’un nœud et la moyenne de fitness de ses nœuds voisins (fnn), et une métrique classique dans les réseaux complexes, le degré d’assortivité (knn) qui est la corrélation entre l’intensité du nœud (somme des poids sortants) et la moyenne des intensités des nœuds voisins. Ces deux métriques locales qui mesurent l’homogénéité de structure et la régularité de fitness dans le réseau sont également très fortement corrélées avec la performance de l’ILS.

Pour expliquer les performances à partir des métriques, nous proposons tout d’abord un modèle simple de régression multilinéaire de la forme :

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p} + \epsilon_i \quad (1)$$

où la variable y de réponse du modèle est le logarithme d’expected run-length, les p différents prédicteurs sont les métriques du ROL et ϵ est le terme usuel de bruit aléatoire. Un tel modèle donne de bons résultats de régression avec un R^2 ⁷ proche de 0.85. Seulement, certaines métriques étant linéairement corrélées. En conséquence, les coefficients β_j de la régression sont souvent non significatifs statistiquement. Aussi, nous avons procédé à une élimination des prédicteurs les moins significatifs (backward elimination). Itérativement, le prédicteur minimisant le score AIC est éliminé. Nous obtenons alors le modèle suivant dont les coefficients sont reportés dans le tableau 5 :

$$\log(\text{erl}) = \beta_0 + \beta_1 \bar{d}_{best} + \beta_2 \text{zout} + \beta_3 \bar{Y} + \beta_4 \text{knn} + \epsilon \quad (2)$$

Par cette technique, les métriques qui expliquent le plus la performance sont donc la distance moyenne à l’optimum global sur le ROL, le degré sortant moyen des nœuds, la disparité pondérée et le degré d’assortivité. Le R^2 de la régression est important, 0.8471, montrant que l’on peut expliquer pour les paysages NK et un ILS les performances mesurées par l’expected run-length avec les caractéristiques du ROL.

Nous avons entrepris une démarche similaire sur le problème combinatoire QAP en essayant d’expliquer cette fois le taux de succès de deux métaheuristiques, un recuit simulé (SA) et un algorithme génétique simple avec croisement, mutation et sélection par tournoi (GA) [Chi+12]. Le but est aussi de comparer les métriques du ROL avec la métrique de la longueur d’autocorrélation de la fitness [Weigo] bien établie dans la littérature des paysages de fitness.

La figure 21 montre des nuages de la corrélation entre les métriques du ROL, la longueur d’autocorrélation et le taux de succès des métaheuristiques à trouver l’optimum global pour

⁷ Le R^2 est le coefficient de détermination qui mesure la qualité d’une régression : $R^2 = \sum_i (\hat{y}_i - \bar{y})^2 / \sum_i (y_i - \bar{y})^2$ où les \hat{y}_i sont les valeurs prédites.

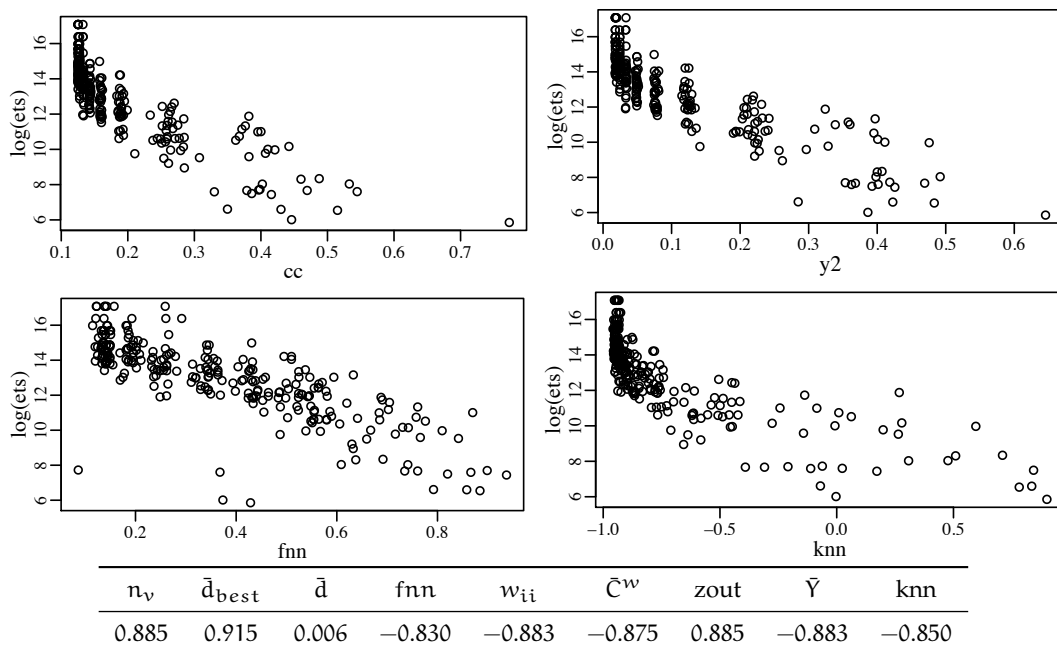


FIGURE 20 – Corrélation entre les métriques du réseau des optima locaux et le logarithme de la performance mesurée par l’expected run-length pour l’ensemble (est sur les graphiques) des instances des paysages NK étudiés avec $N = 18$. Les coefficients de corrélation non paramétrique de Spearman’s sont calculés pour toutes les métriques : n_v nombre de nœuds (optima locaux), \bar{d}_{best} la distance moyenne à l’optimum global, \bar{d} la distance moyenne entre les nœuds, fnn la corrélation entre la fitness du nœud et la moyenne des fitness des nœuds voisins sur le ROL, w_{ii} poids moyen du self-loop, \bar{C}^w , la moyenne du coefficient d’agglomération pondéré, zout moyen du degré sortant, \bar{Y} la disparité moyenne, et knn degré d’assortivité (corrélation de poids sortants entre nœuds voisins).

des instances real-like du QAP de dimension $n = 11$. Contrairement aux résultats précédents, les coefficients de corrélation sont plutôt faibles hormis pour le recuit simulé (SA) avec la distance moyenne à l’optimum sur le ROL et la longueur d’autocorrélation de fitness. Il est difficile de conclure globalement et nous pouvons formuler trois hypothèses. Il se peut que le taux de succès ait une mesure de performance plus difficile à expliquer. Les performances d’un AG ne sont peut-être pas linéairement directement corrélées aux métriques du ROL. Ou encore, l’ensemble des instances (real-like pour une seule dimension $n = 11$) peut ne pas fournir suffisamment de variance pour établir un modèle de régression de qualité.

Les résultats sur le problème combinatoire du flow-shop (instances de dimension $n = 10$ et $m \in \{5, 6, 7, 8, 9, 10\}$ machines) sont plus positifs. Les performances de recherches locales itérées mesurées par l’expected run-length peuvent être expliquées par les métriques extraites du réseau des optima locaux [Dao+13]. Comme pour la section 2.2.3, deux types d’opérateurs de recherche locale sont considérés pour le hill-climber best-improvement utilisés dans l’ILS : l’insertion et l’échange. La perturbation de l’ILS est basée sur l’opérateur d’échange appliqué k fois aléatoirement. Le tableau 6 donne les coefficients de corrélation de Spearman entre chaque métrique du ROL et les différentes variantes de l’ILS. De manière attendue, plus le nombre d’optima locaux augmente plus le temps de résolution (expected run-length) augmente également. La distance moyenne sur le réseau à l’optimum global est aussi la métrique la plus corrélée (et positivement) au temps moyen de résolution des

TABLE 5 – Résumé du modèle multilinéaire de régression de la performance de l'ILS dont les prédicteurs sont les métriques et après la technique d'élimination des prédicteurs pour les paysages NK avec $N = 18$. R^2 ajusté : 0.8471. F-statistique : 368.1 sur 4 et 261 DF, p-value : $< 2.2e - 16$.

	Estimé	Err. Std.	valeur t	Pr(> t)
(Intercept)	10.3838	0.58512	17.75	$9.24e - 47$
\bar{d}_{best}	0.0439	0.00434	10.11	$1.67e - 20$
zout	-0.0306	0.00831	-3.68	$2.81e - 04$
\bar{Y}	-7.2831	1.63038	-4.47	$1.18e - 05$
knn	-0.7457	0.40501	-1.84	$6.67e - 02$

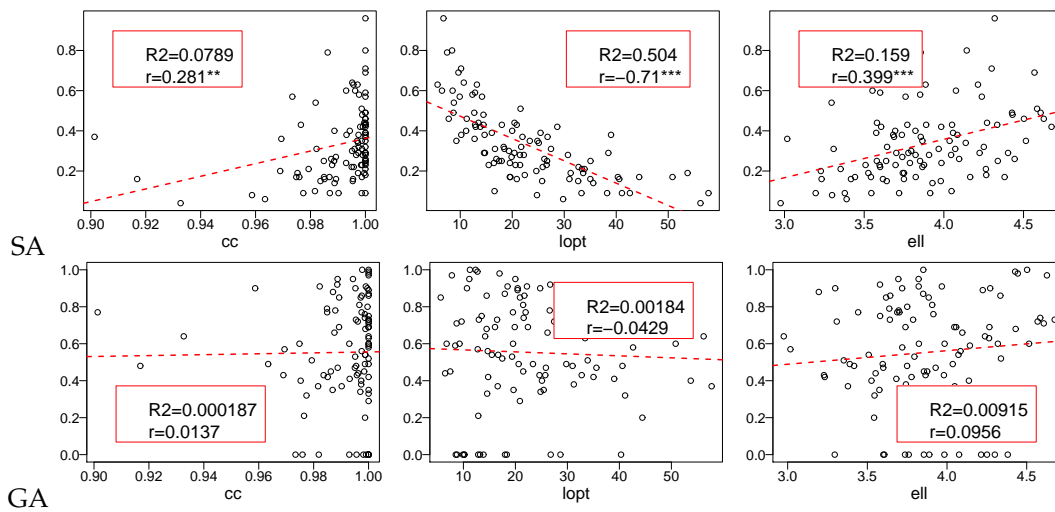


FIGURE 21 – Corrélation entre, de gauche à droite, le coefficient d'agglomération pondéré, la distance moyenne à l'optimum global sur le ROL et la longueur d'autocorrélation de fitness et, de haut en bas, le taux de succès des métaheuristiques recuit simulé (SA) et algorithme génétique (GA) sur les instances QAP real-like de dimension $n = 11$.

algorithmes. Les corrélations avec les autres métriques sont également relativement significatives.

Afin d'approfondir comment les métriques peuvent expliquer les performances des algorithmes et tout en évitant la colinéarité des variables de prédiction, nous avons étudié l'ensemble des modèles de régression multilinéaires dont le nombre de prédicteurs est donné. Parmi tous les modèles possibles avec p prédicteurs, nous avons sélectionné le meilleur selon la statistique C_p de Mallow [Gil96] qui est un indicateur de la qualité de l'ajustement du modèle. Les résultats sont donnés dans le tableau 7. Le prédicteur principal n'est pas le nombre d'optima locaux, mais la distance moyenne à l'optimum global qui explique 57% de la variance de la performance log-transformée. À partir de 3 prédicteurs, plus de 70% de la variance est expliquée par les métriques du ROL avec un simple modèle multilinéaire.

Ces résultats indiquent que le réseau des optima locaux et ses métriques peuvent expliquer les performances, au moins lorsqu'elles sont mesurées par le nombre moyen d'évaluations et pour des métaheuristiques de recherche locale comme l'ILS sur de petites instances de problèmes combinatoires. L'étape suivante est de montrer qu'il est possible de prédire les performances de différentes métaheuristiques sur de larges instances à partir d'estimation des métriques du ROL.

TABLE 6 – Corrélation entre les métriques du réseau des optima locaux et le logarithme de la performance mesurée par la run-length pour l'ensemble des instances de flow-shop étudiées avec $n = 10$ jobs et $m \in \{5, 6, 7, 8, 9, 10\}$ machines. Les coefficients de corrélation de Spearman's sont calculés pour toutes les métriques et les différents types d'ILS utilisant soit l'opérateur d'insertion ou d'échange et dont l'intensité de la perturbation est soit $k = 1$ ou $k = 2$ échanges aléatoires.

ILS	N_V	CC^w	f_{nn}	k_{nn}	r	L_{opt}	L_V	w_{ii}	Y_2	k_{out}
insertion, $k = 1$	0.46	-0.221	0.199	0.078	0.238	0.634	0.40	-0.31	-0.41	0.479
insertion, $k = 2$	0.54	-0.209	0.316	-0.165	0.117	0.691	0.45	-0.476	-0.46	0.55
échange, $k = 1$	0.535	-0.506	-0.004	0.142	0.353	0.624	0.536	-0.235	-0.473	0.448
échange, $k = 2$	0.408	-0.255	0.22	-0.111	0.165	0.527	0.353	-0.272	-0.434	0.409

TABLE 7 – Modèles de régression multilinéaires pour prédire la performance (logarithme du run-length de l'ILS avec insertion et perturbation $k = 2$) à partir des métriques du ROL pour le problème flow-shop. Pour chaque nombre de prédicteurs $\#P$, le meilleur modèle selon la statistique C_p de Mallows est sélectionné, les coefficients de la régression et le R^2 ajusté correspondant sont indiqués.

$\#P$	$\log(N_V)$	CC^w	F_{nn}	k_{nn}	r	$\log(L_{opt})$	$\log(L_V)$	w_{ii}	Y_2	k_{out}	C_p	adjR ²
1						2.13					265.54	0.574
2		-5.18				1.43					64.06	0.675
3						1.481	0.895			-0.042	16.48	0.700
4		-2.079				1.473	0.540			-0.032	8.75	0.704
5		-2.388			-1.633	1.470	0.528			-0.030	5.97	0.706

2.2.5 Prédiction de performance pour les problèmes de grande dimension

Naturellement, le réseau des optima locaux ne peut pas être énuméré exhaustivement pour les grands espaces de recherche. Il s'agit alors d'échantillonner un certain nombre de nœuds et d'arcs pour estimer les métriques les plus pertinentes. Là encore, l'ensemble des métriques du réseau ne peuvent pas être estimées. Par exemple, la métrique souvent la plus porteuse d'information comme la distance moyenne à l'optimum global ne pourra pas être calculée faute de connaître l'optimum global, mais aussi l'ensemble des arcs du réseau.

L'échantillonnage du réseau des optima locaux présente plusieurs difficultés. Le nombre de nœuds (optima locaux) est très important, bien plus important que dans les grands réseaux complexes usuels considérés dans la littérature. De plus, l'échantillonnage de chaque nœud nécessite l'exécution d'un hill-climber. Le nombre d'arcs est lui aussi potentiellement très grand et aussi coûteux à énumérer. De plus, la réalisation d'un échantillonnage uniforme avec une faible complexité des nœuds n'est pas claire voir impossible. Depuis l'avènement des «grands» réseaux du web et des réseaux sociaux «on-line», beaucoup de techniques d'échantillonnage ont été proposés et leur biais étudié dans certains cas [KMT10]. La plupart de ces techniques nécessitent la connaissance au minimum de centaines de milliers de nœuds, et souvent la connaissance du degré des nœuds pour corriger les biais introduits par l'échantillonnage ce qui n'est a priori pas possible pour le réseau des optima locaux. Les techniques d'échantillonnage sont principalement basées sur des marches aléatoires (avec biais ou non) et sur des algorithmes de parcours en largeur. Nous avons adopté ce type d'échantillonnage en réalisant une marche aléatoire couplée avec un échantillonnage dit «boule de neige» (snowball) qui est un parcours en largeur où le nombre d'arcs échantillonné est un paramètre. La procédure snowball vise principalement à estimer les métriques qui nécessitent de connaître les voisins des voisins comme le coefficient d'agglomération.

La figure 22 donne l'idée générale de la technique d'échantillonnage et l'algorithme 6 en donne une description précise. Les arcs échantillonnés sont ceux basés sur la notion d'échappement à une distance D ; ceux basés sur l'énumération du bassin d'attraction étant

impossible à échantillonner efficacement. La marche aléatoire de longueur l commence par un optimum local aléatoire. Pour chaque pas de la marche une courte procédure d'échantillonnage snowball est effectuée : m arcs sont échantillonnés aléatoirement et la procédure snowball s'exécute récursivement jusqu'à une courte profondeur d , en pratique fixée à $d = 2$. Les arcs sont les arcs-échappement et sont donc échantillonnés en produisant une solution aléatoire à une distance maximale de D du nœud puis en effectuant un hill-climber à partir de cette solution afin de trouver un autre optimum local. À chaque pas de la marche aléatoire d'échantillonnage, un nouvel optimum local est sélectionné aléatoire uniformément (sans retour en arrière) parmi ceux déjà échantillonnés lors de l'échantillonnage des arcs.

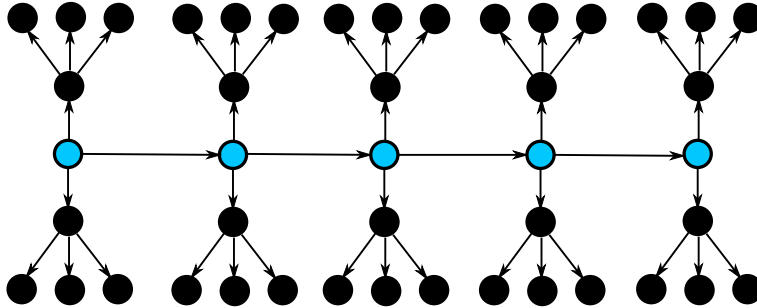


FIGURE 22 – Illustration de l'échantillonnage du réseau des optima locaux où la longueur de la marche aléatoire $l = 5$ et le nombre d'arcs échantillonnés par la procédure snowball est $m = 3$ avec une profondeur de $d = 2$. Les cercles en bleu clair sont les solutions x_i de la marche aléatoire et les cercles en noir sont ceux obtenus par la procédure snowball.

ESTIMATION DES MÉTRIQUES DU RÉSEAU DES OPTIMA LOCAUX. Toutes les métriques sur les réseaux ne peuvent pas être estimées à l'aide de la technique d'échantillonnage, et nous avons essayé d'en retenir le plus grand nombre. De plus, pendant l'échantillonnage, il est possible de mesurer de manière opportuniste sans coût supplémentaire certaines métriques qui ne sont pas issues du ROL comme la longueur des marches adaptatives pour atteindre un optimum local. Le tableau 8 résume l'ensemble des métriques estimées. Comme attendu, l'échantillonnage introduit des biais d'estimation que nous avons étudiés sur de petites instances de problèmes mais qui ne sont pas présentés ici. Sur les larges instances, nous pouvons observer l'effet des paramètres de l'échantillonnage à savoir la longueur de la marche aléatoire (l) et le nombre d'arcs échantillonnés (m). La figure 23 reporte les valeurs de certaines métriques pour le problème QAP. Les observations sont identiques pour les paysages NK. À taille d'échantillon identique, il est préférable pour améliorer la précision des mesures d'augmenter le nombre d'arcs échantillonnés plutôt que la longueur de la marche aléatoire. Par la suite, nous adopterons les paramètres minimaux d'échantillonnage c'est-à-dire une longueur de $l = 100$ et un nombre d'arcs $m = 30$. Afin d'estimer le coefficient d'agglomération, la profondeur d'échantillonnage est $d = 2$. Au maximum, $9 \cdot 10^4$ optima locaux différents sont échantillonnés, certains optima pouvant être échantillonnés plusieurs fois.

PRÉDICTION BASÉE SUR LES MÉTRIQUES ESTIMÉES. Au-delà de la mesure du biais dans l'estimation des métriques, nous pouvons montrer l'utilité des métriques estimées en essayant de prédire la performance de métaheuristiques à partir de celles-ci. Pour montrer cela, nous considérons les paysages NK et le problème QAP. Pour les paysages NK, différents degrés d'épistasie $K \in \{2, 4, 8, 12, 16\}$ et des chaînes binaires de dimension $N \in \{64, 128, 256\}$ sont utilisés, et pour les problèmes QAP, trois types d'instances (une uniforme et deux real-likes avec différents paramètres) et les dimensions $N \in \{30, 50, 70, 100, 150\}$ sont testées. Pour chaque jeu de paramètres, 30 instances sont générées pour former un ensemble de 450 ins-

Algorithm 6 Méthode d'échantillonnage du réseau des optima locaux.

```

1: procedure LONSAMPLING( $d, m, l$ )
2:    $x_0 \leftarrow \text{hc}(x)$  ▷ où  $x$  est une solution aléatoire
3:    $\hat{N} \leftarrow \{x_0\}$ 
4:    $\hat{E} \leftarrow \emptyset$ 
5:   for  $t \leftarrow 0, \dots, l-1$  do
6:     Snowball( $d, m, x_t$ )
7:      $x_{t+1} \leftarrow \text{RandomWalkStep}(x_t)$ 
8:   end for
9: end procedure

1: procedure SNOWBALL( $d, m, x$ )
2:   if  $d > 0$  then
3:     for  $j \leftarrow 1, \dots, m$  do
4:        $x' \leftarrow \text{hc}(\text{op}(x))$ 
5:        $\hat{N} \leftarrow \hat{N} \cup \{x'\}$  ▷ Ajoute un nœud à l'échantillon
6:       if  $(x, x') \in \hat{E}$  then
7:          $\hat{w}_{x,x'} \leftarrow \hat{w}_{x,x'} + 1$ 
8:       else
9:          $\hat{E} \leftarrow \hat{E} \cup \{(x, x')\}$  ▷ Ajoute un arc à l'échantillon
10:         $\hat{w}_{x,x'} \leftarrow 1$ 
11:        Snowball( $d-1, m, x'$ )
12:      end if
13:    end for
14:  end if
15: end procedure

1: procedure RANDOMWALKSTEP( $x_t$ )
2:   neighborSet  $\leftarrow \{x : (x_t, x) \in \hat{E} \wedge x \notin \{x_0, \dots, x_t\}\}$ 
3:   if neighborSet  $\neq \emptyset$  then ▷ Sélectionne aléatoirement un voisin
4:     Select randomly  $x_{t+1} \in \text{neighborSet}$ 
5:   else ▷ Redémarrage à partir d'une solution aléatoire  $x$ 
6:      $x_{t+1} \leftarrow \text{hc}(x)$ 
7:      $\hat{N} \leftarrow \hat{N} \cup \{x_{t+1}\}$ 
8:   end if
9:   return  $x_{t+1}$ 
10: end procedure

```

tances. Nous avons choisi 4 métaheuristiques représentatives : une recherche locale itérée (ILS), un recuit simulé (SA), une recherche taboue (TS) et un algorithme génétique simple (GA). Les détails des algorithmes et les paramètres ne sont pas donnés ici, ils reprennent les implémentations standards de la plateforme paradisEO-MO [Hum+13a] pour imiter un scénario où un utilisateur emploie des versions «off the shell». Pour chaque instance, les métaheuristiques sont exécutées 100 fois indépendamment. Le critère d'arrêt est de 10^6 évaluations et la performance des algorithmes est mesurée de deux manières. L'une (fit) est simplement la valeur moyenne de la fitness de la solution trouvée, et l'autre (rang) est le rang moyen de la solution trouvée par la métaheuristique parmi les 400 solutions trouvées par l'ensemble des 4 métaheuristiques sur l'instance considérée. La seconde mesure de performance est indépendante de l'échelle des valeurs de fitness qui peut varier grandement d'un type d'instance à un autre et permet d'avoir la performance relative aux autres métaheuristiques. Pour les modèles de régression, nous avons utilisé un modèle multilinéaire, mais aussi une régression forêt aléatoire [Bre01].

Les coefficients de détermination R^2 issus d'une validation croisée sont donnés dans le tableau 9 où l'on peut comparer différents ensembles de métrique : les métriques basiques (basic) sont celles que l'on peut trouver dans la littérature des paysages de fitness

TABLE 8 – Métriques estimées par l'échantillonnage du ROL et métriques classiques des paysages de fitness retenues à titre de comparaison.

<i>Métriques du réseau des optima locaux</i>	
<i>fit</i>	Moyenne des fitness des optima locaux.
<i>wii</i>	Poids moyen des self-loops.
<i>zout</i>	Degré sortant moyen (nombre d'arcs sortants).
<i>y2</i>	Disparité pondérée moyenne.
<i>knn</i>	Assortativité pondérée.
<i>wcc</i>	Coefficient d'agglomération pondéré moyen.
<i>fnn</i>	Corrélation fitness / moyenne des fitness voisines.
<i>Métriques de l'échantillonnage</i>	
<i>lhc</i>	Long. moyenne des marches adaptatives aux nœuds.
<i>mlhc</i>	Long. maximale des marches adaptatives aux nœuds.
<i>Métriques classiques de paysage de fitness</i>	
<i>r1</i>	Premier coefficient d'autocorrélation de fitness (marche aléatoire de 10^4 pas).
<i>flo</i>	Moyenne des fitness des optima échantillonnés par 10^3 marches adaptatives.
<i>llo</i>	Longueur moyenne de 10^3 marches adaptatives.

($N, r1, flo, llo$), celles (rol) déduites du réseau des optima locaux ($N, fit, wii, zout, y2, knn, wcc, fnn, nhc, lhc, mlhc$) et enfin celles de la réunion des deux précédentes (all). La validation croisée consiste à partitionner l'ensemble des 450 instances en 80% d'instances d'apprentissage et 20% d'instances de test et à calculer la moyenne des R^2 en répétant 400 fois cette opération (random sub-sampling validation).

Pour tous les ensembles de métriques et les modèles de régression, les R^2 sont élevés sauf pour le modèle linéaire et la prédiction de la fitness et du rang du problème QAP en utilisant les métriques basiques. Sur les 32 combinaisons d'ensemble de métriques, d'algorithmes, de mesures de performance et de modèles de prédiction, l'ensemble des prédicteurs qui contient toutes les métriques est le meilleur ensemble 27 fois, l'ensemble avec les métriques du ROL 12 fois et moitié moins pour les métriques basiques. Les prédictions avec toutes les métriques obtiennent la meilleure qualité et indiquent que d'ajouter les métriques du ROL apporte des informations supplémentaires pertinentes.

Pour les modèles de régression linéaire, l'ensemble des métriques basiques ne donne jamais les meilleurs R^2 , alors que dans la moitié des cas, les métriques ROL font partie des meilleurs prédicteurs avec l'ensemble qui regroupe toutes les métriques. Quand la performance est mesurée par le rang, 7 fois sur 8 les métriques du ROL sont les meilleures en qualité de prédiction. Cela suggère que les métriques du ROL sont plus linéairement corrélées avec la performance que les métriques basiques, et que de considérer l'ensemble de toutes les métriques améliore la qualité de prédiction.

Dans tous les scénarios, les modèles de régression forêt aléatoire donnent de meilleures prédictions que les modèles multilinéaires en rendant compte les interactions non linéaires entre les métriques, et cela même dans le cas des métriques basiques seules. En effet, pour la prédiction de la fitness, 2 fois sur 8 les métriques basiques donnent les meilleurs scores alors que les métriques ROL ne le donnent jamais. Toutefois, lorsque l'ensemble des métriques est utilisé, la qualité de la prédiction est meilleure. Ces résultats montrent que d'ajouter les métriques ROL apporte une information supplémentaire pertinente pour améliorer la qualité de la prédiction même lorsque des modèles de régression efficaces tels que les forêts aléatoires sont employés. Pour illustrer les corrélations, la figure 24 montre des exemples de

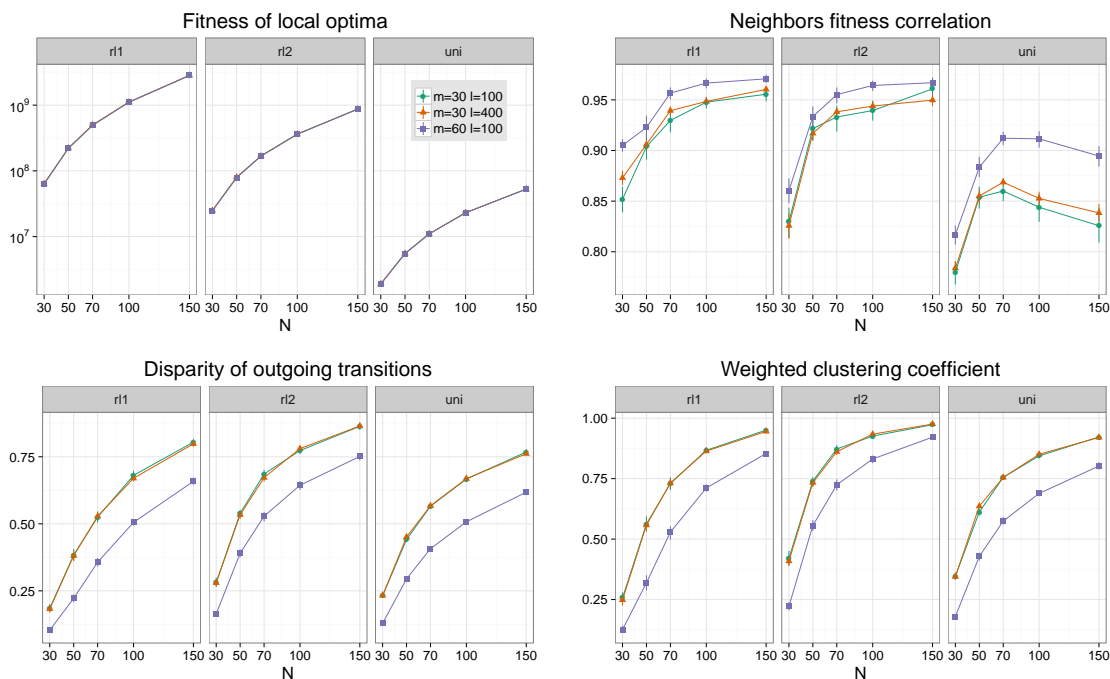


FIGURE 23 – Grandes instances du problème QAP : trois types d’instances, uniforme, real-like1 et real-like2 (*uni*, *rl1*, *rl2*) pour les dimensions $N \in \{30, 50, 70, 100, 150\}$. Valeurs estimées des métriques pour le réseau des optima locaux pour différentes valeurs des paramètres d’échantillonnage, longueur de la marche aléatoire l et nombre d’arcs échantillonnés m .

nuages de points de la performance prédite en fonction de celle observée pour l’ILS sur les problèmes QAP. Les nuages de points sont globalement similaires pour les autres scénarios. Les métriques basiques ont tendance à regrouper les observations sur de petits segments, alors que cet effet est nettement moins marqué avec les métriques ROL.

SÉLECTION POUR PORTFOLIO DE MÉTAHEURISTIQUES. Dans un scénario de portfolio de métaheuristiques, il s’agit de sélectionner une des métaheuristiques qui obtient la meilleure performance, mesurée par la qualité moyenne de solution, en fonction des caractéristiques du paysage de fitness, c’est-à-dire dans notre contexte en fonction du réseau des optima locaux. Nous considérons toujours les mêmes 4 métaheuristiques qu’auparavant. D’un point de vue de l’apprentissage automatique, la sélection de métaheuristique correspond à un problème de classification. Il s’agit d’associer une métaheuristique en fonction de prédicteurs. Ici, nous utilisons 3 méthodes populaires de classification : le modèle multinomial logistique (*logit*), la forêt aléatoire (*rf*) et la machine à vecteurs de support (*svm*).

Le tableau 10 rapporte les erreurs moyennes sur la validation croisée qui répète 400 fois la partition 80/20 aléatoire entre ensembles d’apprentissage et de test. Une erreur de classification est comptée lorsque la métaheuristique n’est pas l’une des meilleures selon un test statistique de Mann-Whitney au seuil de confiance de 5%. Nous utilisons aussi le même test statistique pour indiquer le meilleur ensemble de métrique. À titre de comparaison, nous avons aussi testé deux méthodes naïves de sélection à savoir une sélection aléatoire de l’une des 4 métaheuristiques et la sélection constante qui sélectionne pour l’ensemble des instances toujours la même métaheuristique de meilleure performance sur l’ensemble des instances, c.-à-d. la métaheuristique qui obtient globalement le taux d’erreur le plus faible comme choisirait idéalement une méthode de *tuning* off-line.

Toutes les sélections basées sur les métriques obtiennent des taux d’erreur très faibles par rapport aux deux méthodes naïves confirmant les capacités de cette approche de port-

TABLE 9 – Coefficients de détermination R^2 des régressions en utilisant une validation croisée d'échantillonnages répétés aléatoires. Les deux modèles de régression sont les modèles multilinéaire (lm) et forêt aléatoire (rf). L'ensemble des métriques sont décrites dans le tableau 8 et les mesures de performances sont la moyenne de la fitness de la meilleure solution trouvée (fit) et le rang moyen de la métaheuristique (rang). avg est la moyenne des coefficients R^2 pour les 4 métaheuristicques. Les nombres en gras sont les meilleurs ensembles de métriques selon un test de Mann-Whitney à 5% de confiance.

Mod.	Métr.	Perf.	NK					QAP				
			TS	SA	EA	ILS	avg	TS	SA	EA	ILS	avg
lm	basic	fit	0.8573	0.8739	0.8763	0.8874	0.8737	-38.42	-42.83	-41.63	-39.06	-40.48
lm	rol	fit	0.8996	0.9015	0.9061	0.8954	0.9007	0.9995	1.0000	1.0000	0.9997	0.9998
lm	all	fit	0.9356	0.9455	0.9442	0.9501	0.9439	0.9996	0.9997	0.9999	0.9997	0.9997
lm	basic	rang	0.8591	0.9147	0.6571	0.6401	0.7678	0.2123	0.8324	-0.0123	0.4517	0.3710
lm	rol	rang	0.9517	0.9332	0.7783	0.7166	0.8449	0.7893	0.9673	0.8794	0.9015	0.8844
lm	all	rang	0.9534	0.9355	0.7809	0.7177	0.8469	0.6199	0.9340	0.8577	0.9029	0.8286
rf	basic	fit	0.9043	0.9104	0.9074	0.8871	0.9023	0.8811	0.8820	0.8806	0.8801	0.8809
rf	rol	fit	0.8323	0.8767	0.8567	0.8116	0.8443	0.9009	0.9025	0.9027	0.9019	0.9020
rf	all	fit	0.8886	0.9334	0.9196	0.8778	0.9048	0.9431	0.9445	0.9437	0.9429	0.9436
rf	basic	rang	0.9513	0.9433	0.7729	0.8075	0.8687	0.9375	0.9653	0.8710	0.9569	0.9327
rf	rol	rang	0.9198	0.9291	0.7979	0.7798	0.8566	0.9308	0.9630	0.8820	0.9601	0.9340
rf	all	rang	0.9554	0.9465	0.8153	0.8151	0.8831	0.9381	0.9668	0.8779	0.9643	0.9368

TABLE 10 – Taux d'erreur moyen sur la validation croisée qui répète 400 tirages aléatoires d'échantillon d'apprentissage et de test pour les paysages NK et le problème QAP. Les nombres en gras repèrent les meilleurs ensembles de métrique selon le test de Mann-Whitney à 5% de confiance.

Probl.	Métr.	Taux d'erreur moyen				
		logit	rf	svm	cst	rnd
NK	basic	0.0379	0.0278	0.0158	0.4711	0.6749
	rol	0.0203	0.0249	0.0168		
	all	0.0244	0.0269	0.0165		
QAP	basic	0.0142	0.0107	0.0771	0.4222	0.6706
	rol	0.0156	0.0086	0.0456		
	all	0.0161	0.0106	0.0431		

folio basée sur les caractéristiques du paysage de fitness. Les taux d'erreur des différents ensembles de métriques sont proches. Sur les paysages NK, les modèles svm obtiennent de bons résultats, quel que soit l'ensemble de métriques (moins de 1,7% d'erreur), mais ces modèles svm obtiennent des résultats clairement moins bons pour les problèmes QAP où les métriques du ROL sont les meilleures. Avec le simple modèle logistique (logit), le meilleur ensemble de métriques est celui du ROL sur les paysages NK et il n'y a pas de différence sur les problèmes QAP. Avec le modèle forêt aléatoire qui obtient les taux d'erreur les plus faibles globalement, les métriques du ROL sont nécessaires pour sélectionner la meilleure métaheuristique (moins de 1% d'erreur).

Ces premiers résultats sur les grandes instances montrent que les métriques du réseau des optima locaux peuvent être échantillonnées dans le but de prédire des performances ou de sélectionner la métaheuristique adéquate au sein d'une méthode portfolio de métaheuristicques.

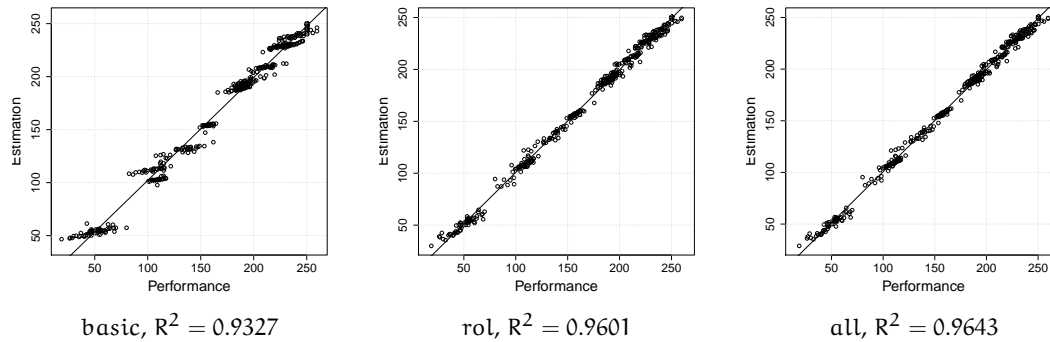


FIGURE 24 – Nuage de points pour le problème QAP entre la performance observée (le rang) et la performance prédite par le modèle de régression forêt aléatoire pour la métaheuristique ILS. Les R^2 issus de la validation croisée sont donnés en légende.

2.3 CONCLUSION ET PERSPECTIVES

Le domaine des paysages de fitness s’inscrit dans les sciences de la complexité : un paysage de fitness est un modèle qui permet d’étudier les systèmes dynamiques induits par les algorithmes d’optimisation tels que les métaheuristiques. Le principe est simple. L’espace de recherche et la relation de voisinage définissent un graphe dont les nœuds sont valués par la valeur de la fonction de fitness à optimiser. De nombreux outils d’analyse des paysages de fitness ont été proposés reposant sur l’analyse du graphe lui-même.

Dans cette partie, nous avons exposé les travaux sur un nouvel outil d’analyse des paysages de fitness, le *réseau des optima locaux*, qui repose sur deux notions utilisées dans les systèmes complexes : le principe d’agrégation de variables et les réseaux complexes. Dans le réseau des optima locaux, l’espace de recherche est condensé sur les optima locaux qui sont les solutions saillantes du paysage de fitness pour un algorithme de recherche local et les arcs sont construits en rapport à la possibilité de «passer» d’un optimum local à un autre.

Dans un premier temps, afin de montrer l’intérêt et l’apport de ce nouvel outil. Nous avons étudié le réseau des optima locaux sur une classe de problèmes génériques de difficulté réglable, les paysages NK, avec un espace de recherche de petite dimension pour permettre l’énumération complète du réseau. À travers cette étude, nous avons pu mettre en évidence que la structure générale des réseaux est très éloignée d’un réseau aléatoire, elle est au contraire dense et très agglomérée. De plus, la structure du réseau définie pour les deux définitions d’arcs proposées reflète la difficulté liée à la non-linéarité des problèmes. Les bassins d’attraction ont une structure qui n’est pas toujours intuitive et qui pourtant peut bien être prise en compte par les définitions d’arcs.

Dans un deuxième temps, nous avons étudié la robustesse du modèle du réseau des optima locaux en étudiant d’autres classes de problèmes dans des situations classiques d’analyse des paysages de fitness. De nouveau sur des petites instances, sur le problème d’affectation quadratique (QAP), le réseau des optima locaux permet clairement de différencier la difficulté entre deux classes d’instances. De plus, nous avons montré en utilisant des outils d’analyse de communauté des réseaux complexes que les optima locaux forment des clusters sur le réseau dans le cas des instances structurées. Sur le problème de type flow-shop, nous avons montré que la structure du réseau des optima locaux met en évidence la différence entre deux opérateurs de recherche locale et par la même, permet de choisir de manière raisonnée un opérateur pertinent. De manière originale dans l’étude des paysages de fitness, le réseau des optima locaux nous a permis d’étudier différents types de

hill-climbers selon la règle pivot, first ou best improvement, à travers la structure de réseau qu'ils impliquent.

Dans un troisième temps, toujours sur de petites instances de problèmes combinatoires, nous avons pu expliquer les performances de métaheuristiques à l'aide de modèles de régression linéaire dont les variables explicatives sont les métriques extraites du réseau des optima locaux. Les performances de métaheuristique de type recherche locale itérée (ILS) mesurées par le nombre moyen d'évaluations ont pu être corrélées avec les métriques du ROL en particulier pour les problèmes NK et flow shop. Cette étape confirme les potentialités du réseau des optima locaux à rendre compte de manière effective de la difficulté d'optimisation.

Enfin, dans l'étape qui présente le plus grand défi, nous avons étudié le réseau des optima locaux pour les grands espaces de recherche des instances de paysages NK et de problèmes QAP. Nous avons défini une procédure d'échantillonnage du réseau en nous inspirant des techniques existantes pour les grands réseaux complexes. Les métriques estimées du réseau ont servi de variables de prédiction pour prédire la performance de quatre types différents de métaheuristiques usuelles. La qualité de régression en utilisant des modèles de régression de l'apprentissage automatique est très satisfaisante. À la suite, nous avons pu mettre en œuvre un scénario de portfolio de métaheuristiques où il s'agit de sélectionner la métaheuristique la plus adaptée à l'instance à résoudre. Les taux d'erreur de sélection d'une des meilleures métaheuristiques sont très faibles et montrent là encore la pertinence et les perspectives d'utilisation du réseau des optima locaux.

Le réseau des optima locaux vient donc compléter la liste des outils d'analyse de paysage de fitness. Naturellement, ce réseau ne se substitue pas aux autres outils existants. Il est évidemment que lorsque la principale difficulté consiste à converger vers un optimum local, ou lorsque le paysage est dominé par de larges plateaux, le réseau ne sera pas judicieux. Mais au-delà de la simple métrique classique dans les paysages de fitness, le réseau des optima locaux propose un modèle dont il est possible d'extraire de nombreuses métriques potentiellement pertinentes pour de nombreux problèmes et métaheuristiques.

PERSPECTIVES. Ce modèle de réseau des optima locaux présente maintenant de nombreuses perspectives à plusieurs niveaux.

Une des améliorations possibles que tout chercheur dans le domaine des paysages de fitness a toujours en tête est de proposer une nouvelle métrique qui reflète la difficulté d'un problème ou d'une classe de problèmes pour une classe d'algorithmes. Avec le réseau des optima locaux, il est sans doute possible de construire de nouvelles métriques en profitant du degré de liberté qu'offre un graphe. Par exemple, nous avons utilisé le coefficient d'agglomération qui est une mesure classique dans les réseaux sociaux pour caractériser le fait fréquent que «les amis de mes amis sont souvent aussi mes amis». En reprenant le sens de cette caractéristique dans le contexte de l'optimisation, on peut surement imaginer une nouvelle métrique qui donne la probabilité que «les optima locaux voisins de mes voisins ne soient pas des optima locaux voisins».

Plus généralement, dans le but d'améliorer l'analyse, on peut changer d'échelle d'observation en considérant finalement que le réseau des optima locaux et lui-même un paysage de fitness à une échelle supérieure «méta». En effet, les nœuds du réseau (optima locaux) peuvent être considérés comme des solutions, les arcs du réseau peuvent définir une nouvelle relation de voisinage et enfin la fonction de fitness originale peut toujours être utilisée comme fonction de fitness de ce «méta» paysage de fitness. Dans cette perspective, on pourrait utiliser tous les outils d'analyse des paysages de fitness et surement observer de nouvelles structures et propriétés à ce niveau «méta» au niveau des optima locaux.

D'un point de vue fondamental et qui n'a pas été abordé dans ce travail, il doit être possible d'obtenir de manière théorique les réseaux des optima locaux pour certains problèmes simples de la littérature. Cette approche est l'une des voies d'analyse dans les travaux théoriques concernant le temps d'exécution moyen (expected running time). Dans ces travaux,

le nombre moyen d'évaluations est calculé pour des problèmes linéaires ou proches du linéaire comme le problème one-max, les problèmes trompeurs, etc. Nous pouvons sûrement calculer le réseau des optima locaux pour ces problèmes ou des problèmes particuliers qu'il reste à définir. Ainsi, il pourrait être possible de comprendre plus fondamentalement pourquoi certaines propriétés comme la densité du graphe ou la distance moyenne sur le graphe semblent être la signature de nombreux problèmes d'optimisation.

Au-delà de la compréhension fondamentale des propriétés des paysages de fitness par le réseau des optima locaux, l'utilisation effective du réseau des optima locaux pour des problèmes de grande dimension dépend du coût de l'estimation des métriques par rapport au coût de l'algorithme lui-même. Des travaux sont donc à développer pour réduire le plus possible le coût de l'échantillonnage du réseau et de l'estimation des métriques. Maintenant que les propriétés génériques des réseaux sont connues, il est sans doute possible d'analyser les biais introduits par des techniques échantillonnages et de choisir un échantillonnage plus efficace.

Dans le portfolio de métaheuristiques proposé, qui est encore à un stade de principe, nous avons suivi le schéma générique de Rice qui sépare clairement les phases d'estimation des métriques du problème, de sélection d'un algorithme adapté par une méthode d'apprentissage automatique et d'exécution cet algorithme. En fait, dans le cas de l'optimisation où l'algorithme est une métaheuristique, il est possible d'imaginer de lier la partie estimation des métriques avec le processus d'optimisation. Il est en effet malheureux d'oublier les optima locaux trouver dans l'estimation des métriques qui pourraient servir de base pour l'algorithme d'optimisation. De nombreuses possibilités de combinaison originales sont à développer dans cette perspective.

À un niveau plus général encore, l'analyse des problèmes l'optimisation par le réseau des optima locaux réside sur un principe d'agrégation de variables. On étudie les propriétés du problème par la structure du graphe résultant plutôt qu'à l'aide de chaînes de Markov. Ici dans ce travail, les nœuds sont les optima locaux. On peut imaginer de changer cette définition de nœud pour d'autres solutions saillantes du paysage. Le principe resterait le même. À partir d'une heuristique simple et peu coûteuse, les nœuds peuvent être définis comme les points de convergence de cette heuristique. La définition des arcs peut être aussi modifiée en fonction de l'heuristique simple choisie, comme nous avons pu le faire avec les variantes first et best improvement des hill-climbers, et de la technique qui permet d'échapper aux optima locaux. De cette manière, on obtient un nouveau réseau dont il faudrait, de la même manière à ce travail, étudier la pertinence par rapport aux performances des métaheuristiques visées.

3

ESPACES DE RECHERCHE DE L'OPTIMISATION MULTIOBJECTIVE

CONTENU

3.1	Introduction	46
3.1.1	Source de difficulté en optimisation multiobjective	46
3.1.2	Paysage de fitness et optimisation multiobjective	47
3.2	Structure des espaces de recherche de l'optimisation multiobjective	48
3.2.1	Portrait des problèmes d'optimisation multiobjective	48
3.2.2	Explication des performances des heuristiques	53
3.2.3	Prédiction de performance pour les problèmes de grande dimension	58
3.2.4	Paysage de fitness multiobjectif basé sur les ensembles	63
3.2.5	Approches scalaires par décomposition	68
3.3	Conclusion et perspectives	77

RÉSUMÉ

Ce chapitre comprend les travaux suivants relatifs aux publications principales indiquées :

- Une brève présentation des difficultés d'optimisation potentielles relatives à chacune des différentes approches de résolution en optimisation multiobjective.
- Une revue des travaux existants sur les paysages de fitness en optimisation multiobjective.
- La présentation de la méthodologie pour obtenir un portrait des problèmes d'optimisation multiobjective à partir des métriques de paysage de fitness. [Ver+11d; Ver+11c; Ver+11b; Ver+13]
- L'explication de performance par des modèles de régression linéaire dont les prédicteurs sont les métriques identifiées et définies précédemment. [Lie+13; Dao+15; Lie+15a; Dao+16]
- La prédiction de performance d'algorithmes pour les problèmes multiobjectifs de grande dimension.
- La proposition d'une nouvelle définition de paysage de fitness multiobjectif qui se base sur un espace de recherche composé d'ensembles de solutions. [VLD11; Bas+13; LLV14]
- Une étude des approches scalaires, la proposition d'un algorithme distribué dont les directions de recherche sont adaptatives et une proposition d'un nouvel algorithme hybride entre une approche scalaire et une approche Pareto. [Der+14a; Der+14c; LVH14; Lie+15b]

Principaux collaborateurs : Arnaud Liefooghe (Université Lille 1, inria Lille-Nord Europe), Bilel Derbel (Université Lille 1, inria Lille-Nord Europe), Clarisse Dahanens (Université Lille 1, inria Lille-Nord Europe), Laetitia Jourdan (Université Lille 1, inria Lille-Nord Europe), Kioshi Tanaka (Shinshu university), Hernán Aguire (Shinshu university), Dima Brockhoff (inria Lille-Nord Europe), A. Goëffon (Université d'Angers), M. Basseur (Université d'Angers).

3.1 INTRODUCTION

Dans cette section, nous introduisons brièvement les principales classes d'algorithmes d'optimisation multiobjective afin d'intuitivement comprendre les difficultés d'optimisation spécifiques de ces algorithmes. En répertoriant les travaux précédents sur les espaces de recherche multiobjectif, nous pourrions proposer, dans la section suivante, de nouvelles métriques qui relatent cette difficulté.

3.1.1 Source de difficulté en optimisation multiobjective

Les problèmes d'optimisation multiobjective présentent des sources de difficultés spécifiques. Il s'agit non plus de trouver une solution comme dans le cas mono-objectif, mais de trouver un ensemble de solutions et donc potentiellement augmenter le nombre d'évaluations pour les calculer toutes. Le nombre de solutions Pareto optimales peut être exponentiel en fonction de la dimension du problème [Ehr06]. On dit dans ce cas que les problèmes d'optimisation qui ont pour but de rechercher toutes les solutions Pareto optimales sont insolubles¹. Par ailleurs, une autre source de difficulté vient de la définition d'optimalité des solutions Pareto optimales. En effet, décider si une solution est une solution Pareto optimale est un problème NP-complet [Ser87] et cela même si aucun des objectifs sous-jacents n'est un problème NP-difficile. Les trois types d'approches décrites précédemment (section 1.1.4) contournent ces difficultés en recherchant des approximations. Toutefois, d'autres sources de difficultés apparaissent spécifiques soit liées à la notion de localité utilisée par ces approches, soit liées à la coopération et l'enchevêtrement des multiples recherches nécessaires pour explorer différentes zones de l'espace objectif.

Dans les approches Pareto, on espère trouver une nouvelle solution non-dominée dans le voisinage des solutions courantes et cela jusqu'à trouver une première solution Pareto optimale. La structure de l'ensemble Pareto optimal influence directement les performances de ces algorithmes. Si celle-ci permet de trouver une nouvelle solution Pareto optimale à partir d'une autre, la difficulté va être moindre que dans le cas contraire. De la même manière que les optima locaux sont les points de convergence des hill-climbers, les approches Pareto vont avoir des solutions au rôle équivalent sur lequel l'algorithme va être stopper et qui peuvent là aussi fortement influencer les performances de ces algorithmes.

Les approches scalaires vont-elles être influencées naturellement par la difficulté d'optimisation des fonctions mono-objectives choisies (non-linéarité, etc.). Mais en plus, la forme du front de Pareto, comme la convexité ou la discontinuité du front, peut aussi influencer les performances. En effet a priori, l'efficacité de la coopération des optimisations scalaires n'est pas la même selon la forme de ce front.

Pour tous les types d'approches, comme dans le cas de l'optimisation mono-objective, la qualité du voisinage dans sa capacité à avoir de meilleures solutions est essentielle dans les performances de tous les types d'approches basés sur un voisinage. La qualité peut être en termes de relation de dominance, de fonction scalaire, ou encore d'indicateur de qualité. Enfin, la corrélation entre les objectifs influence la difficulté générale du problème. En effet, lorsque les objectifs sont parfaitement corrélés, le problème est équivalent à un problème mono-objectif. Intuitivement, l'image de l'espace de recherche dans l'espace objectif et en particulier du front de Pareto influence la coordination globale de la recherche multiobjective. Toutes ces différentes notions intuitives sont les bases pour définir un paysage de fitness multiobjectif et les métriques le caractérisent.

1. «Intractable» en langue anglaise

3.1.2 Paysage de fitness et optimisation multiobjective

Il existe finalement peu de travaux sur l'étude des paysages de fitness multiobjectifs. Par exemple, on peut trouver des études sur la distribution des optima locaux mono-objectifs dans l'espace objectif pour le problème de voyageur de commerce multiobjectif [BH98]. Ce type de travail permet de comprendre les performances des approches scalaires. En fait, le premier obstacle est de définir clairement le paysage de fitness multiobjectif. La définition sous-entendue dans de nombreux travaux initiés par [KC02] consiste simplement à changer la fonction de fitness par celle utilisée pour le problème multiobjectif : un paysage de fitness multiobjectif est un triplet $(\mathcal{X}, \mathcal{N}, f)$ où $f : \mathcal{X} \rightarrow \mathbb{R}^M$ représente la fonction multiobjective. Avec cette perspective et dans le but de caractériser la difficulté des approches basées sur la relation de dominance Pareto, il est possible de définir des notions d'optimum local pour l'optimisation multiobjective [PCSo4]. Une solution *optimum locale Pareto* ne peut être dominée par aucune de ces solutions voisines :

Définition 7 *Optimum Local Pareto (PLO) [PCSo4]. Une solution $x \in \mathcal{X}$ est optimum local Pareto (PLO) si et seulement s'il n'existe pas de solution voisine $y \in \mathcal{N}(x)$ qui domine x .*

L'ensemble de tous les optima locaux Pareto est appelé *ensemble des optima locaux Pareto* (\mathcal{X}_{PLO}). On peut aussi définir une notion d'optimalité pour les ensembles de solutions :

Définition 8 *Ensemble localement non-dominé. Un ensemble A localement non-dominé est un ensemble de solutions mutuellement non-dominées telles que $\forall x \in A, \forall y \in (\mathcal{N}(x) \setminus A)$, il existe $z \in A$ tel que $y \prec z$.*

Autrement dit, un ensemble localement non-dominé ne peut pas être amélioré en termes de dominance par une solution voisine d'une des solutions de l'ensemble. Naturellement, ces définitions ont été proposées dans le but d'étudier les points de convergence d'algorithmes tels que PLS (algorithme 1). De nouveau pour étudier les capacités d'un algorithme de PLS à trouver des solutions Pareto optimales, des travaux sur le problème TSP et QAP multiobjectif ont étudié l'ensemble Pareto Optimal du point de vue de la relation de voisinage [PS09]. Plusieurs travaux [KC02; GD06; GD07] basés sur la même définition de paysage de fitness ont étendu la corrélation fitness-distance mono-objective (voir tableau 1) au cas multiobjectif. Ils ont étudié sur différents types d'échantillon de solutions la corrélation entre la distance dans l'espace objectif (distance euclidienne) et la distance dans l'espace de décision \mathcal{X} à l'ensemble Pareto optimal qu'il faut connaître *a priori*. Les résultats montrent une certaine corrélation, mais dont il est difficile de donner une relation précise avec les performances des algorithmes.

Un autre type d'étude de paysage de fitness se focalise sur la dynamique d'algorithmes Pareto comme NSGA-II. La sélection dans NSGA-II associe un rang de dominance à chaque solution sur lequel portent les études de paysage. Sur de petites instances de problèmes, les auteurs partitionnent l'espace de recherche en fronts de solution de même rang selon la dominance Pareto. On peut remarquer que l'étude du paysage de fitness repose encore ici sur une agrégation de variable (voir section 2.1.2) qui consiste à regrouper des solutions de même rang. Dans [Gar09b] en faisant un parallèle entre la non-comparabilité et la neutralité, le graphe réduit aux fronts et leur relation par voisinage de solution est représenté pour quelques instances de problèmes en montrant visuellement les différences. Les travaux de [AT07] vont plus loin en fournissant plusieurs métriques très pertinentes relativement aux fronts comme leur nombre, leur taille respective, la probabilité de passer d'un front à un autre par l'opérateur de recherche local, l'hypervolume des fronts, etc. De plus, contrairement aux travaux précédents qui portent sur des instances assez spécifiques (TSP, QAP), l'étude expérimentale porte sur les paysages NK multiobjectifs introduits par l'article dont les résultats sont potentiellement plus génériques. Ceci est le point de départ de nombre de nos travaux que nous allons développer par la suite.

3.2 STRUCTURE DES ESPACES DE RECHERCHE DE L'OPTIMISATION MULTIOBJECTIVE

3.2.1 Portrait des problèmes d'optimisation multiobjective

La méthodologie pour établir un portrait des paysages de fitness multiobjectifs consiste à définir une classe de problèmes génériques dont les paramètres sont les principales caractéristiques d'un problème d'optimisation multiobjective et de mesurer ensuite les métriques du paysage de fitness afin de comprendre la structure générale et d'en tirer des recommandations pour la conception de nouveaux algorithmes. Les principales métriques étudiées portent principalement sur l'ensemble Pareto optimal, dans un premier temps. Sauf exception, des instances de petites dimensions sont analysées afin de pouvoir obtenir l'ensemble Pareto optimal et d'éviter les biais possibles d'un échantillonnage.

PROBLÈME GÉNÉRIQUE DES PAYSAGES ρMNK . Comme dans le cas mono-objectif, le paysage NK est un bon candidat pour avoir un modèle de problème générique. La dimension et la non-linéarité sont réglables et une version multiobjective de ce problème existe [AT07], mais dont la corrélation entre objectifs n'est pas ajustable. La corrélation entre objectifs est pourtant l'une des caractéristiques majeures sur la forme du front de Pareto comme nous allons l'expliquer. Et bien qu'il existe une version multiobjective du problème d'affectation quadratique (mQAP) [KC02] dont la corrélation est partiellement réglable, nous proposons ici une version augmentée du paysage NK dont la corrélation est finement contrôlable entre toutes les paires d'objectifs ce qui n'est pas le cas pour le problème mQAP.

Les paysages NK multiobjectifs [AT07], paysages MNK, sont naturellement définis par une fonction NK pour chaque objectif Φ^m :

$$\forall m \in \{1, \dots, M\}, \quad \forall x \in \mathcal{X}, \quad \Phi^m(x) = \frac{1}{N} \sum_{i=1}^N \Phi_i^m(x_i, x_{i_1^m}, \dots, x_{i_{K_m}^m})$$

Le degré d'épistasies K_m peut être différent pour chaque objectif Φ^m , mais en pratique les études ont toujours choisi un degré identique K et nous faisons également ce choix ici. Les fonctions de contribution Φ_i^m sont définies en extension en associant à chaque vecteur de longueur $K + 1$ les nombres $y_{x_i, x_{i_1^m}, \dots, x_{i_{K_m}^m}}^{m,i}$. Originellement dans les paysages MNK, ces nombres sont choisis aléatoirement uniformément et indépendamment d'un objectif à l'autre dans l'intervalle $[0, 1[$.

Les paysages NK multiobjectifs avec corrélation ajustable entre objectifs [Ver+11c], appelés *paysages ρMNK* , changent la façon dont les fonctions de contribution sont générées. D'abord, définissons les CMNK-landscapes où les liens épistatiques sont identiques pour toutes les fonctions : $\forall m \in \{1, \dots, M\}, K_m = K$ et $\forall m \in \{1, \dots, M\}, \forall j \in \{1, \dots, K\}, i_j^m = i_j$. Les fonctions de contributions ne sont plus indépendantes. Les M nombres $(y_{x_i, x_{i_1}, \dots, x_{i_K}}^{1,i}, \dots, y_{x_i, x_{i_1}, \dots, x_{i_K}}^{M,i})$ associés aux vecteurs $(x_i, x_{i_1}, \dots, x_{i_K})$ suivent une distribution uniforme multivariée de dimension M définie par la matrice de corrélation C . Autrement dit, les nombres y suivent une distribution multivariée dont la loi marginale est uniforme et dont la corrélation entre contributions est définie par la matrice symétrique semi-définie positive C avec $M(M - 1)/2$ coefficients. Pour limiter le nombre de paramètres qui définit la matrice de corrélation, les paysages ρMNK ont la même corrélation ρ entre toutes les paires d'objectifs avec $c_{nn} = 1$ pour tout n et $c_{np} = \rho$ pour tous les $n \neq p$.

C est une matrice de corrélation ce qui impose qu'elle soit symétrique semi-définie positive. Par conséquent, ρ doit être strictement plus grand que $\frac{-1}{M-1}$. La corrélation ne peut pas être excessivement négative entre toutes les paires d'objectifs. Pour 2 objectifs, ρ appartient à l'intervalle, $] -1, 1[$ mais à partir de 3 objectifs, ρ est strictement plus grand que -0.5 . Bien sûr, il est possible de définir des matrices de corrélation plus spécifiques pour l'étude de certaines situations comme dans le cas many-objectif. Pour générer les distributions uniformes

multivariées, nous utilisons des nombres aléatoires générés par la loi normale multivariée de corrélation C et la densité de la loi normale pour obtenir une distribution uniforme. Les générateurs ainsi que les instances ont été mis en ligne sur la plateforme mocoBench².

Il est possible de prouver que les coefficients de la matrice C définissent la corrélation moyenne entre les objectifs : pour tous les objectifs i et j , $E[\text{corr}(\{\Phi^i(x), \Phi^j(x)\} : x \in \mathcal{X})] = c_{ij}$ où $C = (c_{ij})_{(i,j) \in [1,m]^2}$. Il est donc possible de précisément régler la corrélation entre les objectifs. La figure 25 donne l'image de l'espace de recherche dans l'espace objectif pour différentes valeurs de corrélation.

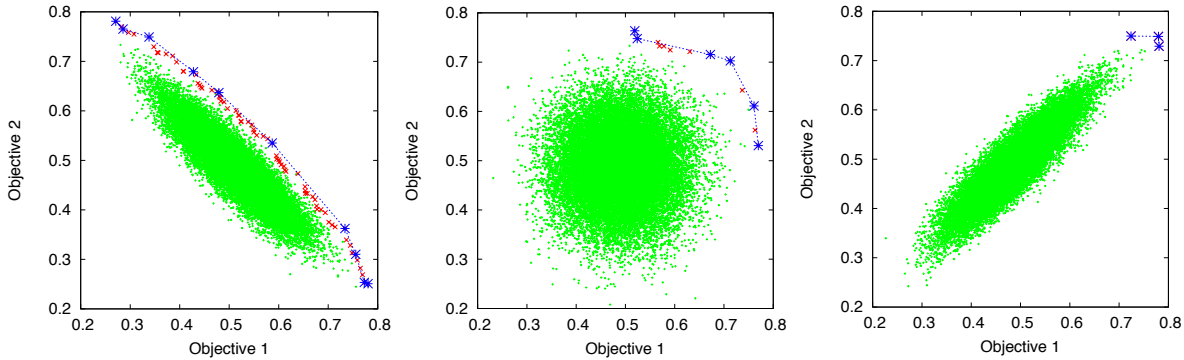


FIGURE 25 – Espace objectif pour 3 instances de paysages ρMNK . Le nombre d'objectifs est $M = 2$, la longueur des chaînes binaires est $N = 18$, et le degré d'épistasies est $K = 4$. De gauche à droite, la corrélation entre les objectifs passe de négative à positive ($\rho = -0.9, 0.0$ et 0.9). Les points verts représentent 10% des solutions (aléatoires) de l'espace de recherche, les points rouges sont les solutions Pareto optimales et les points bleus sont les solutions supportées.

ENSEMBLE PARETO OPTIMAL. Pour dresser un portrait d'un paysage de fitness multiobjectif, nous commençons par étudier l'ensemble des solutions Pareto optimales comme on peut le trouver dans divers travaux de la littérature. Dans cette étape, les paysages ρMNK de petites dimensions ($N = 18$ le plus souvent) sont considérés afin de pouvoir extraire l'ensemble Pareto optimal et d'analyser les métriques relatives. Le tableau 12 donne la liste de ces métriques. Le nombre de solutions Pareto optimales est un facteur important de difficulté de problème, plus le nombre est important plus la qualité d'un ensemble d'approximation est affectée. Les solutions supportées sont les solutions dont l'image dans l'espace objectif est sur l'enveloppe convexe du front de Pareto. Ces solutions sont les seules solutions Pareto optimales que peut atteindre une fonction scalaire de somme pondérée couramment utilisée dans les approches scalaires. En ajoutant la relation de voisinage entre les solutions à l'ensemble Pareto, il est possible de définir un graphe, le *graphe Pareto*. Lorsque ce graphe est connexe, il est possible à partir d'une solution Pareto optimale et de la relation de voisinage d'obtenir l'ensemble Pareto optimal. Cette technique est à la base des approches deux-phases tel que [PS06].

L'étude expérimentale sur de nombreux jeux de paramètres et de nombreuses instances (30 instances par jeu de paramètres) permet d'obtenir une étude quantitative qui met en relation les métriques avec le nombre d'objectifs, le degré de non-linéarité et la corrélation entre les objectifs. La figure 26 donne un aperçu du nombre de solution Pareto optimale et de la proportion du nombre de solutions supportées. Le nombre de solutions Pareto optimales diminue exponentiellement avec la corrélation et augmente exponentiellement avec la dimension, par contre la non-linéarité a une influence très limitée. Dans le même temps, la proportion de solutions supportée diminue à l'inverse du nombre de solutions Pareto optimales.

2. Projet sourceforge mocoBench disponible à <http://mocobench.sourceforge.net/>

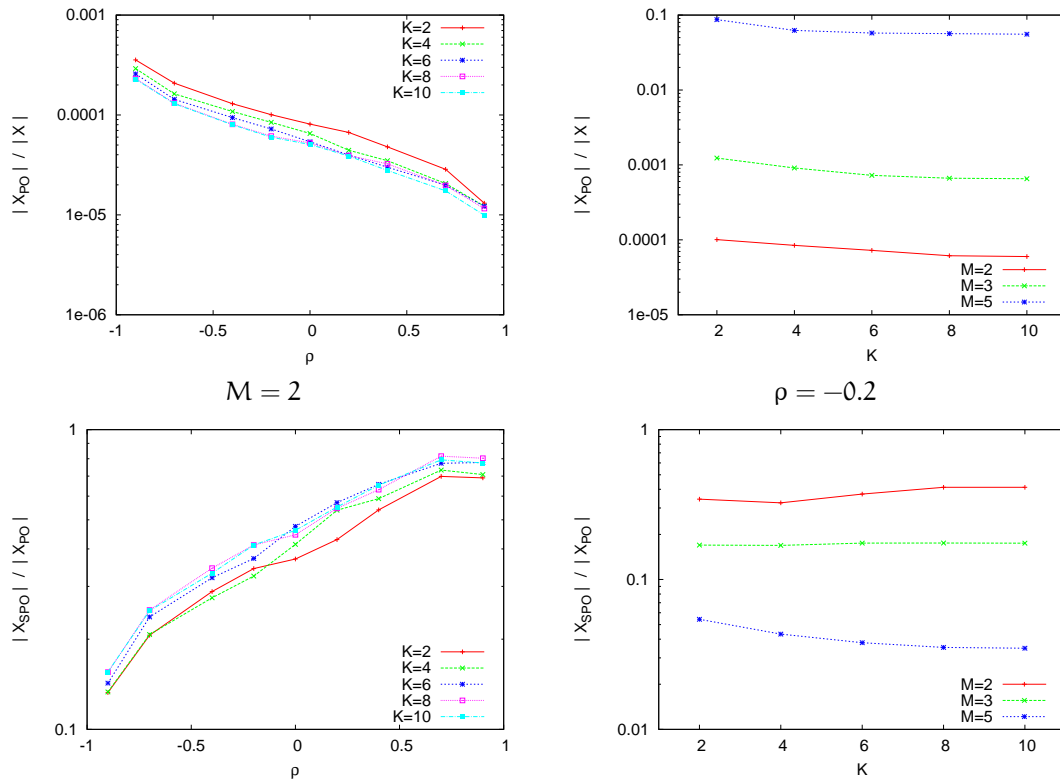


FIGURE 26 – Paysages ρ MNK : en fonction des paramètres ρ , M et K , proportion dans l'espace de recherche (X) de solution Pareto optimale (X_{PO}) et proportion dans l'ensemble Pareto optimal de solutions supportées (X_{SPO}), c.-à-d. de solutions dont l'image est sur l'enveloppe convexe du front de Pareto.

La figure 27 donne les principales caractéristiques du graphe Pareto d'un point de vue de l'optimisation à savoir la proportion de la taille de la plus composante connexe et de la distance minimale entre les solutions pour rendre complètement connexe le graphe. Lorsque le nombre d'objectifs est important ($M = 5$) et la corrélation négative, quasiment la totalité des solutions Pareto optimales sont connectées par la relation de voisinage. La proportion décroît avec le degré de non-linéarité du problème. Par contre, le lien n'est pas direct avec le nombre d'objectifs et la corrélation. Par exemple, la taille relative de la plus grande composante augmente avec le nombre d'objectif lorsque la corrélation est négative et diminue avec une corrélation positive.

Ces premiers résultats montrent l'interdépendance entre les paramètres globaux, en particulier le nombre d'objectifs et la corrélation entre objectifs. De plus, la non-linéarité du problème n'influence pas seulement la difficulté intrinsèque de chaque objectif, mais également la structure globale de l'ensemble Pareto optimal dont il faut tenir compte dans la conception d'heuristique.

OPTIMA LOCAUX PARETO. L'ensemble Pareto optimal est le point de convergence que vise généralement une heuristique. Les caractéristiques de cet ensemble ne peuvent expliquer à elles seules toute la difficulté d'optimisation puisque les heuristiques rencontrent aussi un certain nombre de solutions non Pareto optimales. D'ailleurs dans l'article [Ver+11b], en nous inspirant de la construction du problème long-path mono-objectif [HGD94], nous avons pu construire un problème dont le graphe Pareto est connexe et dont le temps de résolution pour trouver une approximation est exponentiel avec la dimension du problème. Et a contrario, nous avons aussi proposé un problème dont le graphe Pareto n'a pas a priori

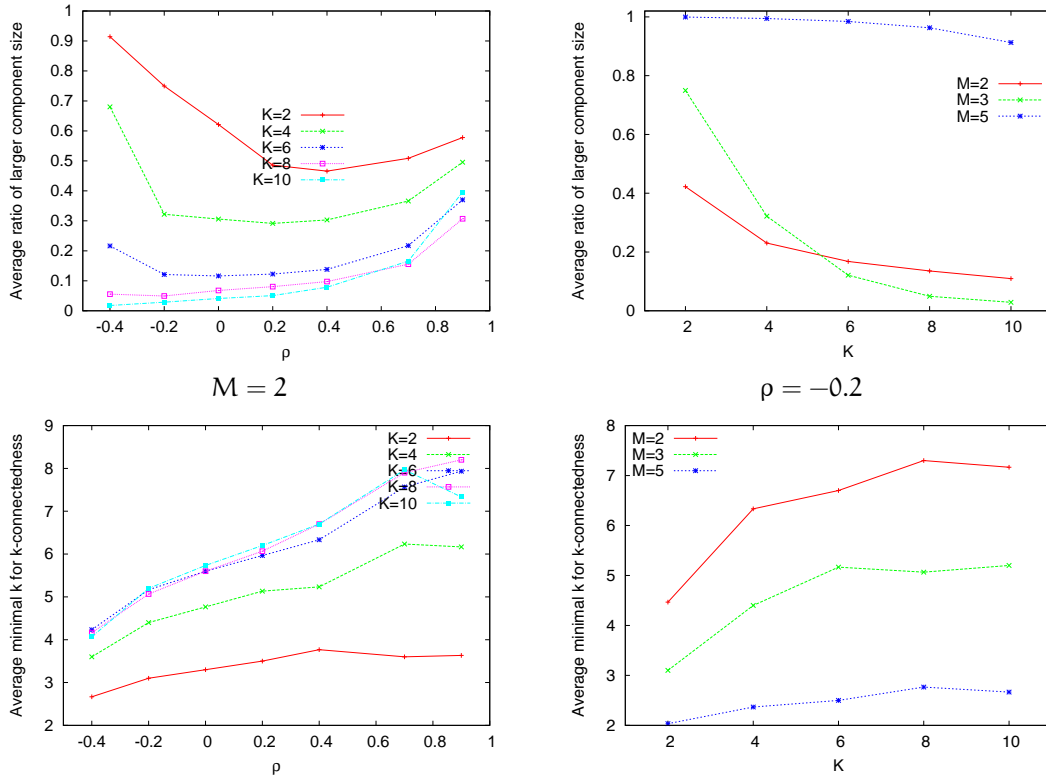


FIGURE 27 – Paysages ρMNK : ratio (sur le nombre de solutions Pareto optimales) de la taille de la plus grande composante connexe du graphe Pareto et distance minimale pour obtenir la connexité de toutes les solutions Pareto optimales.

de bonnes propriétés (non connexe), mais dont le temps de résolution est expérimentalement polynomial avec la dimension. Les propriétés de l'ensemble Pareto optimal, comme la connexité, ne sont pas suffisantes pour expliquer les performances.

Une caractéristique importante des paysages de fitness multiobjectifs est également le nombre d'optima locaux Pareto (définition 7), noté PLO pour Pareto Local Optimum, qui sont des solutions rencontrées par les algorithmes pendant la recherche. Sur les mêmes petites instances, nous avons énuméré le nombre d'optima locaux Pareto [Ver+11d; Ver+13] (voir la figure 28). Comme pour le nombre de solutions Pareto optimales, le nombre d'optima locaux Pareto décroît exponentiellement avec la corrélation entre objectifs et augmente avec le nombre d'objectif. Par contre, comme dans le cas mono-objectif avec les optima locaux, le nombre d'optima locaux Pareto augmente largement avec le degré de non-linéarité du problème. Sur une figure non présentée ici, nous avons aussi montré que le nombre de PLO n'est pas directement corrélé avec le nombre de solutions Pareto optimales (noté POS), ainsi le nombre de PLO donne bien une information supplémentaire sur la difficulté du problème multiobjectif.

Nous avons pu proposer une méthode d'estimation du nombre d'optima locaux Pareto [Ver+11d] qui nous permet d'estimer ce nombre pour les grandes instances de problème d'optimisation multiobjectif. Comme dans le cas mono-objectif, l'estimation est basée sur la longueur de marches adaptatives. Une *marche adaptative Pareto* est définie comme une suite de solutions telle que la solution du pas suivant domine la solution du pas précédent. L'algorithme 7 donne un exemple hill-climber Pareto. La figure 28 montre la corrélation entre la longueur de la marche adaptative Pareto et le logarithme de la proportion du nombre d'optima locaux Pareto pour des instances de dimension $N \leq 18$. La corrélation est forte (0.997) et l'équation de la régression est par donnée par $|\text{PLO}|/|\text{POS}| = be^{aL}$ où L est la

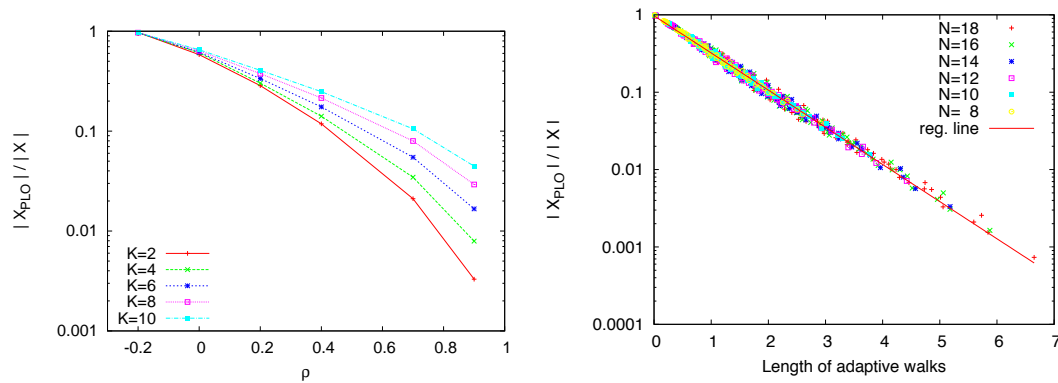


FIGURE 28 – Paysages ρ MNK. À gauche, proportion d’optima locaux Pareto pour un nombre d’objectifs $M = 5$ et différentes valeurs de corrélation ρ entre objectif et degré K de non-linéarité. À droite, corrélation entre la longueur des marches adaptatives Pareto et la proportion d’optima locaux Pareto (logarithme) pour toutes les instances (différents ρ , M , N et K). Le coefficient de corrélation de Pearson est de 0.997.

longueur moyenne de la marche adaptative et $a = -1.60$ et $b = 0.97$ sont les coefficients de la régression.

Algorithm 7 Pareto hill-climber définissant une marche adaptative Pareto.

```

1: Commencer par une solution aléatoire  $x \in X$ 
2:  $\text{step} \leftarrow 0$ 
3: while  $x$  is not a PLO do
4:   Choisir aléatoirement  $x'$  dans  $\{y \in \mathcal{N}(x) : x \prec y\}$ 
5:    $x \leftarrow x'$ 
6:    $\text{step} \leftarrow \text{step} + 1$ 
7: end while

```

En supposant que le modèle de régression sur les petites instances est aussi valide pour les grandes instances, on peut avoir une estimation du nombre d’optima locaux Pareto pour les instances de grandes tailles. La figure 29 montre les longueurs moyennes des marches adaptatives Pareto pour de grandes instances non énumérables. Pour un même nombre d’objectif, de corrélation entre objectifs et de degré de non-linéarité, la longueur augmente linéairement avec la dimension N du problème : $L \approx \alpha N$ avec $\alpha < 0.33$ pour toutes les régressions linéaires. Les longueurs de marche adaptative augmentent moins vite que la dimension du problème. En suivant l’hypothèse initiale, la proportion d’optima locaux Pareto dans l’espace de recherche diminue exponentiellement avec la dimension : $|\text{PLO}|/|\text{POS}| = be^{a \alpha N}$ avec $-0.6 < a \alpha < 0$ et le nombre absolu de PLO augmente exponentiellement puisque $|\text{PLO}| = be^{(1+a \alpha) N}$ avec $0.4 < 1 + a \alpha$.

La longueur des marches adaptatives Pareto permet d’estimer le nombre d’optima locaux Pareto pour les espaces de recherche de grande taille. On peut soit alors étudier ce nombre en fonction des caractéristiques du problème, ou soit, nous le verrons par la suite, utiliser cette estimation comme variable de prédiction de la difficulté.

RECOMMANDATION POUR LA CONCEPTION D’ALGORITHMES MULTIOBJECTIFS. En se référant aux différents buts de l’analyse de paysage de fitness (voir tableau 5), l’étude de paysage de fitness multiobjectif précédente permet de donner un aperçu et une compréhension de la structure générale du paysage de fitness afin de choisir de manière raisonnée une conception d’heuristiques multiobjectives. Par exemple, souvent le nombre d’objectif sert à déterminer si le problème est un problème multiobjectif ou many-objectif (4 ou plus objectifs) et à choisir l’algorithme en conséquence. Clairement, cette étude contraste cette simplification en insistant sur le fait que la corrélation entre objectifs est tout aussi importante. Plus largement,

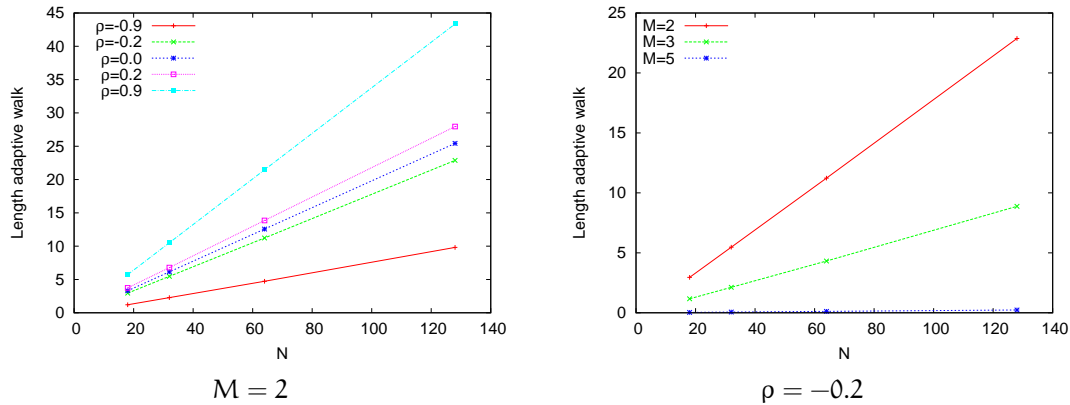


FIGURE 29 – Longueur moyenne des marches adaptatives Pareto sur les paysages ρMNK de grande dimension en fonction de N pour différentes valeurs du nombre d'objectifs M et de corrélation entre objectifs ρ .

les propriétés du paysage de fitness doivent être prises en compte pour la conception d'un algorithme efficace.

En esquisant une vision globale à grand trait, nous avons essayé de résumer qualitativement dans le tableau 11 les résultats et les recommandations de conception pour les grandes approches de résolution [Ver+13] :

- (i) **Taille de l'archive.** La qualité de l'ensemble manipulé par l'algorithme, c.-à-d. l'archive, est directement liée à la cardinalité de l'ensemble Pareto optimal. Au vu des résultats, la taille de l'archive doit être limitée lorsque la dimension du problème est grande et que les objectifs sont contradictoires (anti-corrélés), sinon une archive de taille non bornée est envisageable.
- (ii) **Approche scalaire.** Les approches scalaires qui utilisent une somme pondérée trouvent un sous-ensemble des solutions Pareto optimales qui sont supportées. Ces approches doivent être efficaces lorsqu'il y a peu d'objectifs ou qu'ils sont corrélés. Pour les problèmes many-objectif, il semble peu approprié d'utiliser un tel type d'approche.
- (iii) **Approche Pareto.** Le nombre d'optima locaux Pareto augmente avec la non-linéarité du problème, le nombre d'objectifs, ou encore diminue avec la corrélation des objectifs. Les approches Pareto devraient être plus efficaces lorsque ce nombre est faible. Toutefois, il faut aussi prendre en compte la complexité de l'archivage pour borner la taille de l'ensemble de recherche lorsque le nombre d'optima locaux semble favorable à ce type d'approche.
- (iv) **Approche Pareto deux-phases.** Quand toutes les solutions Pareto optimales sont connectées ou presque connectées par la relation de voisinage, une approche Pareto deux-phases peut parcourir un grand nombre de solutions Pareto optimales par voisinage et semble donc pertinente. La connexité décroît avec la non-linéarité et l'anti-corrélation entre les objectifs. Lorsque la corrélation est positive, elle décroît également avec le nombre d'objectifs. Ceci donne des indications sur la capacité de la méthode deux-phases et peut fournir une information utile pour le nombre de solutions initiales ou la taille du voisinage à considérer.

Il est à noter que peu de conclusion peut être donnée pour les approches basées sur un indicateur de qualité. Nous reviendrons sur ces approches dans la section 3.2.4.

3.2.2 Explication des performances des heuristiques

Le but ici est de montrer que l'on peut expliquer les performances d'heuristiques multiobjectives de manière effective à partir des métriques du paysage de fitness multiobjectif.

TABLE 11 – Synthèse qualitative et recommandations de conception pour les grandes approches de résolution. (N) dimension de l'espace de recherche, (K) degré de non-linéarité, (M) nombre d'objectifs et (ρ) corrélation entre objectifs.

Métriques	Propriété du problème				Suggestion pour la conception de recherche locale
	N	K	M	ρ	
Cardinalité de l'ensemble Pareto optimal					+ archive de taille limitée – archive non bornée
Proportion de solutions supportées					+ scalaire efficace – scalaire non efficace
Connexité du graphe Pareto	$\rho \geq 0$				+ deux-phases efficace
	$\rho < 0$				– deux-phases non efficace
Nombre de optima locaux Pareto					+ Pareto non efficace
					– Pareto efficace

De manière semblable aux travaux sur le réseau des optima locaux de la section 2.2.4, nous essayons d'atteindre un objectif difficile de l'analyse des paysages de fitness, le premier de ce type pour un problème d'optimisation multiobjectif.

CORRÉLATION ENTRE MÉTRIQUES. Dans le souci d'avoir des métriques aussi utilisables dans un contexte d'optimisation boîte noire, et en suivant les notions des paysages de fitness mono-objectif comme nous avons pu le faire avec les marches adaptatives, nous avons introduit deux nouvelles métriques basées sur des marches aléatoires sur le paysage [Lie+13]. Nous proposons une métrique qui s'inspire de l'autocorrélation de fitness (voir tableau 1). Une marche aléatoire sur le paysage multiobjectif est effectuée, mais au lieu de reporter la fitness de la solution x à chaque pas de la marche aléatoire, nous associons soit l'hypervolume de la solution unique $hv(\{x\})$ dans la métrique appelée l'autocorrélation de l'hypervolume ($corhv$), soit l'hypervolume local, c'est-à-dire celui couvert par la solution et ses solutions voisines $hv(\mathcal{N}(x) \cup \{x\})$ dans l'autocorrélation de l'hypervolume local ($corlhv$). Ces deux nouvelles métriques visent à mesurer la rugosité du paysage de fitness multiobjectif.

Le tableau 12 résume l'ensemble des métriques que nous avons maintenant à disposition qui peuvent être utilisées comme prédicteurs. Ces métriques ne sont pas indépendantes. Sur les mêmes instances de petite taille des paysages ρMNK avec $N = 18$, nous avons calculé dans la figure 30 la matrice de corrélation entre toutes les paires de métriques.

Le nombre d'objectifs M ou la corrélation entre objectifs sont modérément corrélés avec la cardinalité de l'ensemble Pareto optimal $\log(np_0)$. Aucune de ces caractéristiques seules ne peut expliquer le nombre de solutions Pareto optimales comme nous l'avons déjà souligné auparavant. En fait, un modèle de régression multilinéaire pour prédire $\log(np_0)$ basé sur M et ρ peut expliquer 70% de la variance du logarithme du nombre de solutions Pareto optimales avec un coefficient de corrélation élevé de 0.84. Aussi, nous pensons qu'il n'est pas possible d'analyser un problème d'optimisation many-objectif sans tenir compte de la corrélation entre objectifs.

En plus des corrélations déjà commentées dans les sections précédentes avec les paramètres globaux des paysages ρMNK , des groupes de métriques sont corrélés entre elles. Par exemple, et de manière non surprenante, la distance moyenne et maximale entre les solutions Pareto optimales, le nombre de composantes connexes dans le graphe Pareto et la distance minimale pour qu'il soit complètement connecté sont très corrélés entre eux. Le degré de non-linéarité K est le paramètre du paysage ρMNK le plus corrélé avec les métriques sur la structure du graphe Pareto. Par contre, l'hypervolume du front de Pareto n'est que faiblement corrélé avec ces métriques du graphe Pareto. De manière plus intéressante encore, les nouvelles métriques sur la rugosité du paysage multiobjectif sont fortement corrélées avec le degré d'épistasie K comme dans le cas mono-objectif : le coefficient d'autocorréla-

TABLE 12 – Résumé des métriques de paysage de fitness multiobjectif et des paramètres des paysages ρ MNK.

Métriques paysage multiobjectif (12)		
npo	Nombre de solutions Pareto optimales	[KC03]
hv	Hypervolume du front de Pareto	[AT07]
avgd	Distance moyenne entre les solutions Pareto optimales	[Lie+13]
maxd	Distance maximale entre les solutions Pareto optimales	[KC03]
supp	Proportion de solutions supportées	[KC03]
nplo	Nombre d'optima locaux Pareto	[PS06]
ladapt	Longueur des marches adaptatives Pareto	[Ver+11d]
ncomp	Nombre relatif de composantes connexes du graphe Pareto	[PS09]
lcomp	Taille relative de la plus grande composante connexe	[Ver+11c]
dconn	Distance minimale pour connecter le graphe Pareto	[PS09]
corhv	Premier coefficient d'autocorrélation de l'hypervolume	[Lie+13]
corlhv	Premier coefficient d'autocorrélation de l'hypervolume local	[Lie+13]

Paramètres des paysages ρMNK (3)		
ρ	Corrélation entre objectifs	
M	Nombre d'objectifs	
K	Degré de non-linéarité (épistasie)	

tion est plus grand quand le degré d'épistasie est faible. Même si ces métriques ne sont pas directement corrélées avec le nombre de solutions Pareto optimales ou encore avec le nombre d'optima locaux Pareto, les autocorrélations de l'hypervolume peuvent nous servir à estimer le degré de non-linéarité dans un contexte d'optimisation boîte noire. La matrice de corrélation donne un aperçu général des prédicteurs pour la prédiction de performance d'un algorithme. Nous allons maintenant donner un exemple d'heuristique multiobjective et mettre en relation la performance avec les prédicteurs.

EXPLICATION PAR UN MODÈLE LINÉAIRE. L'heuristique Global Simple EMO (G-SEMO) [LTZ04] est une approche Pareto très proche d'une Pareto Local Search - PLS - (voir algorithme 1). La différence principale est dans la manière de créer de nouvelles solutions. Alors que PLS génère l'ensemble des solutions voisines, par exemple pour les chaînes binaires l'ensemble des solutions avec un bit est modifié, G-SEMO ne génère qu'une seule solution (méthode steady-state) par un opérateur stochastique qui potentiellement peut générer toutes les solutions de l'espace de recherche (opérateur ergodique). Typiquement pour une chaîne binaire de longueur N , chaque bit est modifié avec un taux de $1/N$ par l'opérateur stochastique. Le pseudo-code est donné par l'algorithme 8. Le critère d'arrêt n'est généralement pas précisé. Ici, sur ces petites instances où l'ensemble Pareto optimal est connu, l'algorithme s'arrête lorsque l'algorithme approxime le front de Pareto à un facteur de $(1 + \epsilon)$ (cf. indicateur ϵ de la section 1.1.4), ou encore l'algorithme s'arrête lorsque le nombre d'évaluations T_{\max} est atteint. Comme dans la section 2.2.4, la performance est mesurée par le nombre d'évaluations moyen (*expected run-length*) pour atteindre l'approximation en simulant des redémarrages indépendants : $ert = \mathbb{E}(T_s) + \left(\frac{1-p_s}{p_s}\right) T_{\max}$ où T_s est

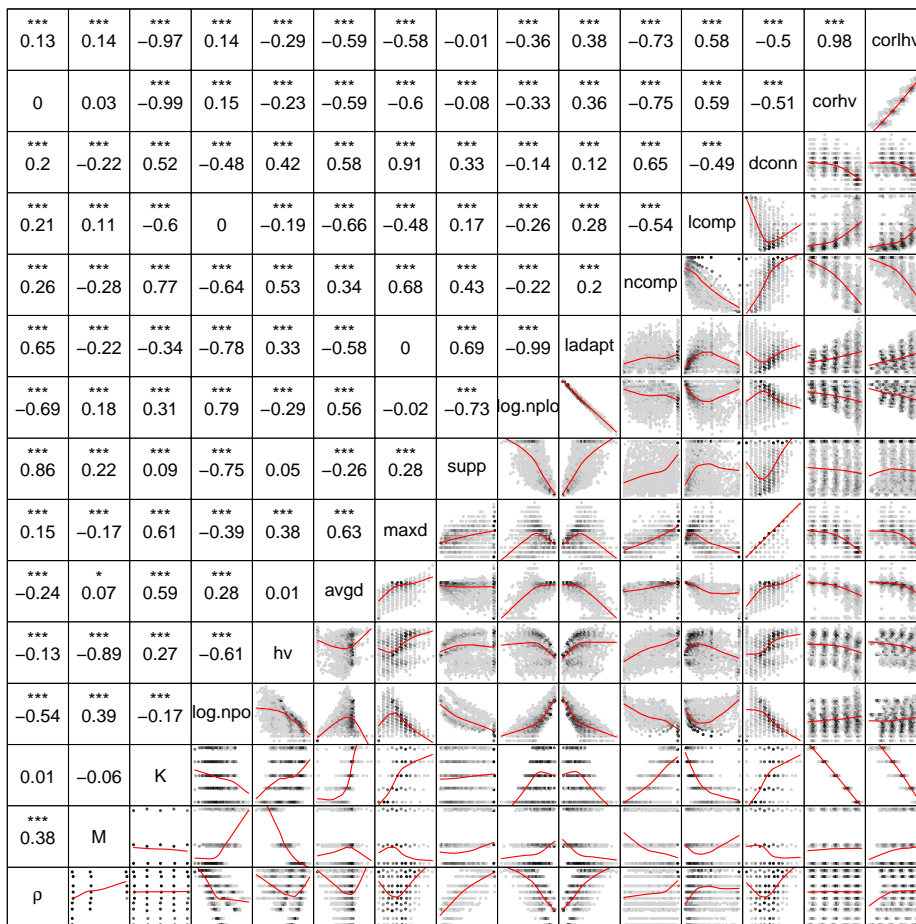


FIGURE 30 – Matrice de corrélation entre toutes les paires de métriques. La diagonale reporte le nom de la métrique. Pour chaque paire de métriques, le nuage de points avec une régression de type spline est représenté et le coefficient de corrélation (Pearson) correspondant est indiqué à l'intersection opposée. Les coefficients sont testés avec comme hypothèse nulle la corrélation nulle. Les p-valeur avec correction de Bonferroni sont encodées par niveau par : 0.05 (*), 0.01 (**), and 0.001 (***)

le nombre d'évaluations lorsque la $(1 + \epsilon)$ -approximation est atteinte et p_s le taux de succès de l'algorithme.

La figure 31 reporte les corrélations entre chaque métrique et les performances de G-SEMO. Parmi les trois paramètres ρ , M et K , le degré de non-linéarité K est le plus fortement corrélé avec $\log(\text{ert})$ (0.85) alors qu'il y a peu de corrélation respectivement avec les paramètres M et ρ . Aussi, les plus fortes corrélations avec les métriques sont avec les métriques d'autocorrélation de l'hypervolume. De manière surprenante, ni le nombre de solutions Pareto optimales ni l'hypervolume du front de Pareto ne sont corrélés avec les performances. En revanche, lorsque la distance entre ces solutions augmente, le temps de résolution augmente également. De même, chaque métrique du graphe Pareto est modérément corrélée avec la performance de l'algorithme. La structure d'ensemble Pareto optimal a tout de même plus d'effet sur les performances que la taille de cet ensemble. Enfin, le nombre d'optima locaux Pareto est modérément corrélé avec la performance, le temps de résolution étant plus grand lorsque ce nombre augmente. Individuellement, certaines métriques sont linéairement corrélées avec les performances ce qui amène à mesurer leur effet conjoint.

Algorithm 8 Pseudo-code du Global Simple EMO (G-SEMO)

```

1:  $A \leftarrow \{x^0\}$ 
2: loop
3:   sélectionner aléatoirement (uniformément)  $x$  dans  $A$ 
4:   créer  $x'$  en modifiant chaque bit de  $x$  avec le taux  $1/N$ 
5:    $A \leftarrow$  solutions non dominées de  $A \cup \{x'\}$ 
6: end loop
7: return  $A$ 

```

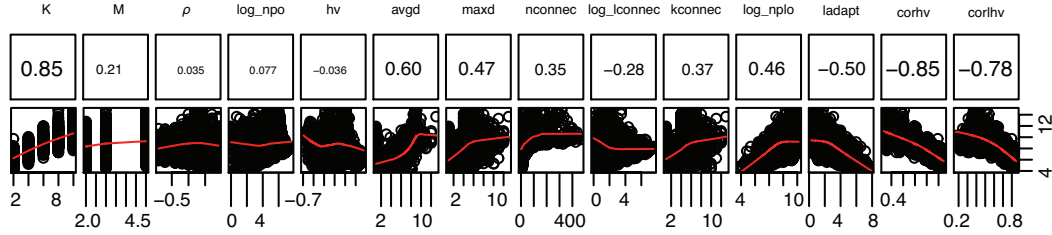


FIGURE 31 – Corrélation entre la performance mesurée par le logarithme de l'expected running-time ($\log(\text{ert})$) et chaque métrique du paysage de fitness multiobjectif. Le nom de la métrique est indiqué dans la première ligne, les coefficients de corrélation sont reportés dans la deuxième ligne et sur la troisième les nuages de points avec une spline sont dessinés.

L'effet conjoint de toutes les métriques sur les performances peut être modélisé par une régression linéaire du type [Lie+15a] : $y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{k,i} + \epsilon_i$ où pour une instance i , y_i sont les observations de la performance $\log(\text{ert})$, $x_{k,i}$ sont les p métriques du paysage de fitness et ϵ_i sont les résidus $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. Les statistiques sur la qualité du modèle se trouvent à la première ligne du tableau 13. Pour ce modèle avec toutes les métriques, le coefficient de corrélation est très élevé supérieur à 0.9. La validation croisée du modèle indique aussi une bonne capacité de généralisation avec une erreur moyenne quadratique faible. Afin de déterminer les métriques les plus importantes dans le modèle linéaire, nous procédons à une élimination itérative des prédicteurs les moins significatifs dans la prédiction (backward elimination). En commençant avec un modèle avec l'ensemble de tous les prédicteurs, nous retirons un à un les prédicteurs qui réduisent le moins la qualité du modèle. Les derniers prédicteurs restants sont ceux les plus significatifs pour le modèle selon cette technique. Le même tableau 13 reporte les éliminations successives. La qualité du modèle se dégrade de façon notable lorsque le paramètre ρ est retiré. Les métriques les plus importantes sont donc par ordre décroissant K , m , $\log(\text{nplo})$, hv et ρ . Bien que leur effet individuel puisse être faible, leur effet conjoint est important. Le degré de non-linéarité K est la métrique la plus importante. Ce paramètre K n'est pas connu dans un scénario boîte noire, mais l'autocorrélation de l'hypervolume pourra servir à l'estimer si besoin.

EXPLICATION PAR UN MODÈLE LINÉAIRE MIXTE. Il est possible d'aller encore plus loin dans l'analyse de l'effet conjoint des métriques sur les performances. En effet, au-delà des paramètres principaux ρ , M et K , il est possible de connaître l'effet des métriques du paysage de fitness à l'aide d'un modèle de régression linéaire mixte [Dao+15; Dao+16]. Dans le modèle linéaire mixte, les instances ayant le même triplet (ρ, M, K) sont mises ensemble dans un même groupe. Le modèle permet d'analyser l'effet (linéaire) des autres métriques sur la variance de la variable observée au sein de chaque groupe. La variance est due, ici, à la seule variabilité des fonctions de contribution des paysages ρMNK . Plus formellement

TABLE 13 – Résumé des statistiques des régressions multilinéaires avec sélection des métriques selon une backward-élimination. r est le coefficient de corrélation linéaire de Pearson, MAE l'erreur moyenne absolue, RMSE la racine carrée de l'erreur quadratique moyenne et %RRSE est le rapport entre RMSE de la régression et RMSE du modèle basique qui donne la valeur moyenne.

	ensemble d'entraînement				validation croisée 10-folds			
	r	MAE	RMSE	%RRSE	r	MAE	RMSE	%RRSE
toutes les métriques	0.91	0.58	0.76	40.34	0.91	0.58	0.76	40.48
\ maxd	0.91	0.58	0.76	40.38	0.91	0.58	0.76	40.50
\ log(ncomp)	0.91	0.58	0.76	40.49	0.91	0.58	0.77	40.60
\ log(np0)	0.91	0.58	0.77	40.53	0.91	0.58	0.77	40.62
\ lcomp	0.91	0.58	0.77	40.54	0.91	0.58	0.77	40.61
\ avgd	0.91	0.58	0.77	40.58	0.91	0.58	0.77	40.64
\ dconn	0.91	0.58	0.77	40.66	0.91	0.58	0.77	40.71
\ ρ	0.91	0.61	0.80	42.20	0.91	0.61	0.80	42.26
\ hv	0.89	0.67	0.87	46.29	0.89	0.67	0.87	46.33
\ log(np10)	0.88	0.70	0.91	48.19	0.88	0.70	0.91	48.21
\ m	0.85	0.77	1.00	52.95	0.85	0.78	1.00	52.96
\ k	0.00	1.53	1.89	100.00	0.04	1.53	1.89	100.00

dans un modèle mixte, la performance sur l'instance i appartenant au groupe d'instances j est modélisée par :

$$y_{ij} = \beta_0 + \sum_{k=1}^p \beta_k x_{kij} + \alpha_j + \epsilon_{ij} \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

où les α_j sont des variables aléatoires i.i.d tel que $\alpha_j \sim \mathcal{N}(0, \sigma_\alpha^2)$. Les α_j caractérisent l'effet du groupe j . Dans le vocabulaire des modèles mixtes, les coefficients linéaires β_k sont les effets *fixes* communs à tous les groupes et les coefficients α_j sont les effets aléatoires associés à chaque groupe. Pour plus de précision, par exemple sur les méthodes d'estimation des paramètres du modèle, on pourra se reporter à [CG10].

La figure 32 reporte les valeurs des coefficients du modèle multilinéaire mixte. Lorsque la métrique augmente d'une unité, le coefficient correspondant indique la variation de la performance ($\log(\text{ert})$) à travers l'ensemble des groupes et donc indépendamment des paramètres d'instances (ρ, M, K). Cette interprétation simple du modèle mixte est l'une des motivations du choix de ce type de modèle. En plus dans cette étude, nous avons comparé deux heuristiques, G-SEMO et aussi la Pareto Local Search (PLS) avec le même critère d'arrêt. Pour les deux algorithmes, les métriques qui ont le plus d'effet sur la performance sont l'hypervolume du front de Pareto et les métriques, que nous avons proposées, de l'autocorrélation de l'hypervolume. Ce qui prouve qu'au-delà du lien avec le degré de non-linéarité K , la rugosité du paysage a un effet notable sur les performances. Parmi d'autres observations, remarquons, de même, que soit le nombre d'optima locaux Pareto pour G-SEMO, soit la longueur des marches adaptatives pour PLS ont un effet notable sur les performances. Ceci montre d'une autre manière l'importance des optima locaux Pareto sur les performances de ces algorithmes : plus il y a d'optima locaux Pareto, plus le temps d'approximation est long et cela indépendamment de la valeur de ρ, M ou K .

Le modèle mixte montre de nouveau sur de petites instances de problèmes que les métriques identifiées des paysages de fitness multiobjectifs permettent d'expliquer les performances d'heuristiques multiobjectives en particulier avec les métriques liées à l'autocorrélation de l'hypervolume lors de marches aléatoires et la longueur de marches adaptatives Pareto.

3.2.3 Prédiction de performance pour les problèmes de grande dimension

Dans cette section, il s'agit de montrer comment l'on peut prédire les performances d'heuristiques multiobjectives à partir de métriques estimées du paysage de fitness pour des problèmes de grande taille.

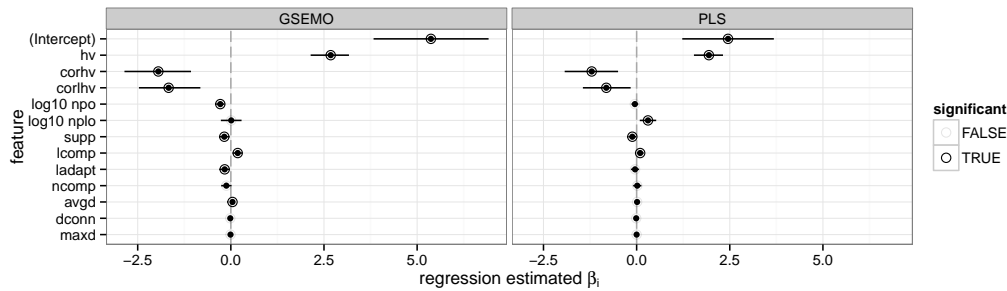


FIGURE 32 – Valeurs estimées des coefficients du modèle linéaire mixte qui explique la performance mesurée par le logarithme de l'expected running-time ($\log(\text{ert})$) de G-SEMO et PLS en fonction des différentes métriques. Le point indique la valeur estimée du paramètre β_k et la barre l'intervalle de confiance à 95%. Les points encadrés indiquent les coefficients significativement non nuls ($H_0 : \beta_k = 0$).

NOUVELLES MÉTRIQUES POUR LES GRANDS ESPACES. Les métriques étudiées jusqu'à présent pour les espaces de recherche de petite taille portent principalement sur l'ensemble Pareto optimal et sa structure avec deux notables exceptions, la longueur des marches adaptatives Pareto et les autocorrélations de l'hypervolume lors d'une marche aléatoire. Pour les grands espaces de recherche, il n'est évidemment pas possible d'avoir l'ensemble Pareto optimal pour prédire les performances d'une heuristique dont le but est de le trouver. C'est pourquoi nous proposons de nouvelles métriques basées sur des marches aléatoires et adaptatives. Pour chacune des métriques, il s'agit d'extraire des informations sur le voisinage des solutions, à savoir : l'hypervolume de la solution $\{x\}$ (lhv), l'hypervolume de la solution avec ses solutions voisines $\{x\} \cup \mathcal{N}(x)$ (nhv), la différence entre l'hypervolume de tout le voisinage et celui de la solution ($dlhv$), la proportion de solutions non-dominées dans $\{x\} \cup \mathcal{N}(x)$, la proportion de solutions supportées dans $\{x\} \cup \mathcal{N}(x)$, le ratio du nombre de solutions voisines de $\mathcal{N}(x)$ qui domine la solution x ($rsup$), le ratio du nombre de solutions voisines de $\mathcal{N}(x)$ incomparables avec x (req) et enfin le ratio du nombre de solutions voisines de $\mathcal{N}(x)$ qui est dominé par la solution x ($rinf$). Ensuite, au cours des marches aléatoires ou adaptatives, nous mesurons la valeur moyenne de ces informations sur l'échantillon disponible de solutions. En plus pour les marches aléatoires, nous pouvons aussi calculer l'autocorrélation de chacune de ces métriques pour avoir la régularité de ces informations dans le voisinage des solutions et ainsi potentiellement avoir des indices sur la qualité du voisinage. Par ailleurs, à partir de l'échantillon de solution obtenu par une marche aléatoire, il est aussi possible d'estimer la corrélation entre les objectifs. Le tableau 14 donne la liste complète des 26 métriques qui ne sont bien sûr pas toutes indépendantes.

PORTRAIT DES MÉTRIQUES. Nous pouvons maintenant donner un portrait des paysages de fitness pour les problèmes de grande dimension à travers les nouvelles métriques proposées. Pour cela, nous générons un ensemble varié d'instances de paysage ρMKNK avec un nombre d'objectifs M entre 1 et 5, une dimension du problème N entre 64 et 256], un degré d'épistasie K entre 0 et 8 et une corrélation entre objectifs ρ comprise dans l'intervalle $]-1, 1]$. 1000 instances sont produites par un échantillonnage des quadruplets de paramètres en hypercube latin (tout en respectant la contrainte $-1/(M-1) < \rho$).

La figure 33 montre la matrice de corrélation entre toutes les métriques obtenues sur les 1000 instances à partir d'une marche aléatoire de longueur 10^4 et de 10^3 marches adaptatives Pareto. Au vu des distributions des métriques, nous avons choisi le coefficient de corrélation non paramétrique de Kendall qui donne le nombre de couples de points concordants dans la variation des couples de métriques. En plus, de la matrice de corrélation, nous avons réalisé un clustering hiérarchique en prenant la distance $1 - |\tau|$ où τ est le coefficient de corrélation

TABLE 14 – Synthèse de l'ensemble des métriques du paysage de fitness multiobjectif dédié aux problèmes de grande taille.

Abrév.	Métriques (17) à partir d'une marche aléatoire de longueur 10^4
lhv_mean.rw	hypervolume moy. de la solution
nhv_mean.rw	hypervolume moy. de tout le voisinage
dlhv_mean.rw	différence moy. entre "lhv" et "nhv"
rlpos_mean.rw	proportion moy. de solutions PO localement
rlspos_mean.rw	proportion moy. des solutions supportées localement
rinf_mean.rw	proportion moy. de solutions dominées dans le voisinage
req_mean.rw	proportion moy. de solutions incomparables dans le voisinage
rsup_mean.rw	proportion moy. de solutions dominantes dans le voisinage
corf.rw	corrélation estimée entre objectifs
lhv_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "lhv"
dlhv_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "dlhv"
nhv_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "nhv"
rlpos_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "rlpos"
rlspos_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "rlspos"
rinf_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "rinf"
req_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "req"
rsup_autocor1.rw	1 ^{er} coefficient d'autocorrélation de "rsup"
Métriques (9) à partir des 10^3 marches adaptatives Pareto	
lhv_mean.phc	hypervolume moy. de la solution
nhv_mean.phc	hypervolume moy. de tout le voisinage
dlhv_mean.phc	différence moy. entre "lhv" et "nhv"
rlpos_mean.phc	proportion moy. de solutions PO localement
rlspos_mean.phc	proportion moy. des solutions supportées localement
rinf_mean.phc	proportion moy. de solutions dominées dans le voisinage
req_mean.phc	proportion moy. de solutions incomparables dans le voisinage
rsup_mean.phc	proportion moy. de solutions dominantes dans el voisinage
phclength_mean.phc	longueur moy. des marches adaptatives Pareto

de Kendall entre les deux métriques. De cette façon, nous pouvons faire apparaître des groupes de métriques similaires, très corrélées (ou anti-corrélées) entre elles.

Par le clustering hiérarchique, trois groupes de métriques se distinguent donc. Chacun d'eux contient un paramètre des paysages ρ MNK : l'un est associé au degré d'épistasie relatif K/N (et aussi faiblement associé avec la dimension N), l'un avec la corrélation entre objectifs ρ et l'un plus petit avec le nombre de fonctions objectif M . Dans le groupe avec K/N , on retrouve à une exception près toutes les métriques d'autocorrélation ce qui fait sens puisqu'elles sont a priori destinées à mesurer la rugosité du paysage multiobjectif. On note aussi la présence des métriques sur la différence d'hypervolume local qui sont aussi des métriques relatives à la variation d'hypervolume dans le voisinage des solutions et sans doute liée également à la rugosité locale. Dans le groupe associé au paramètre de corrélation entre objectifs, on retrouve naturellement l'estimation de cette corrélation sur l'échantillon de solutions. On trouve également les métriques sur la moyenne du nombre de solutions supportées dans le voisinage comme si le voisinage donnait localement une indication sur la forme globale du front de Pareto. Enfin, on trouve toutes les métriques relatives à la dominance (nombre de solutions non-dominées localement, etc.). Le dernier groupe réunit le nombre d'objectifs avec les métriques relatives à l'hypervolume moyen.

PRÉDICTION DES PERFORMANCES. Les performances prédites sont celles de deux algorithmes qui servent de référence dans cette partie consacrée à l'optimisation multiobjective, Global Simple EMO, G-SEMO (algorithme 8) et Pareto Local Search, PLS (algorithme 1). Les ensembles Pareto optimaux n'étant pas connus pour les instances des paysages ρ MNK consi-

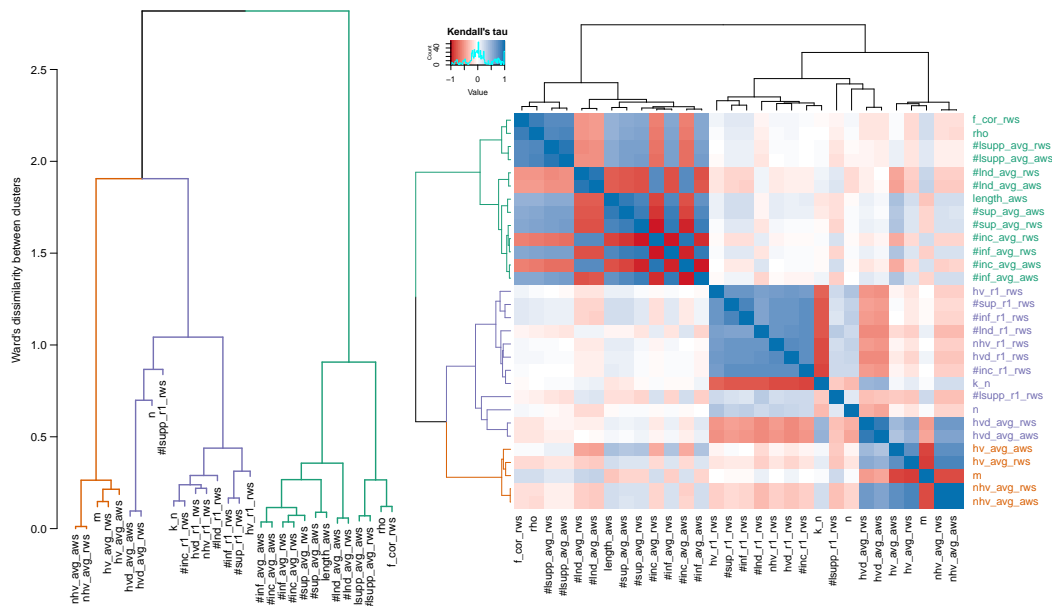


FIGURE 33 – Clustering des métriques (gauche) et matrice de corrélation (droite) pour les paysages ρ MNK de grande taille. Le coefficient de corrélation de Kendall τ qui est basé sur le rang et est non-linéaire, est reporté par une couleur pour chaque paire de métriques. La distance entre deux métriques est définie par $1 - |\tau|$. Le clustering est un clustering hiérarchique, au plus bas est la séparation, au plus est la corrélation. Ce résultat du clustering est aussi utilisé pour ordonner les lignes et colonnes de la matrice de corrélation. On peut voir que 3 groupes auxquels on peut associer les paramètres du paysage se distinguent.

dérés, un front de référence est calculé à partir de l'ensemble des solutions non-dominées trouvées à travers toutes les exécutions des algorithmes. Le critère d'arrêt des algorithmes est un nombre d'évaluations et la performance est mesurée par l'indicateur ϵ par rapport au front de référence. Dans le but prédictif et moins explicatif, nous utilisons une variante des forêts aléatoires comme modèle de régression, les «Extremely randomized trees» [GEW06] qui peuvent détecter les relations non-linéaires entre les prédicteurs.

La table 15 donne la qualité, selon la validation croisée, des régressions de la performance pour différents ensembles de métriques. Les qualités sont quasiment identiques que cela soit pour l'algorithme G-SEMO ou PLS. Pour les régressions basées seulement sur les paramètres des paysages ρ MNK, la qualité de la régression n'est bonne que lorsque le degré de non-linéarité relatif K/N est introduit dans l'ensemble des prédicteurs. Dans ce cas, le coefficient de détermination R^2 indique que plus de 89% de la variance peut être expliquée par le modèle. En fait, avec seulement M et N comme prédicteurs, la qualité de la prédiction est moins bonne que la prédiction par la moyenne des performances. Considérer la corrélation entre objectifs, mais aussi la non-linéarité du problème multiobjectif est indispensable pour comprendre et prédire les performances d'un algorithme.

Dans un scénario boîte noire d'optimisation, la corrélation ρ et surtout le degré de non-linéarité K ne sont pas connus. Dans ce contexte, la prédiction ne peut se baser que sur l'ensemble des métriques estimées à l'aide des marches aléatoires (sampling rw) et adaptatives Pareto (sampling phc). La qualité de la régression est très bonne lorsque le modèle utilise l'ensemble de ces métriques (85% de la variance expliquée). Les modèles basés sur les paramètres ou sur les métriques estimés ont la même qualité de prédiction (excepté sur l'erreur absolue). En effet, une seule instance est générée par quadruplet de paramètres et la performance s'explique surtout par la valeur de ces paramètres. Les métriques estimées n'ajoutent pas de nouvelles informations qui pourraient traduire la variance dans les fonctions de contribution du paysage ρ MNK. Enfin, nous pouvons observer que les modèles

TABLE 15 – Qualité des régressions de la validation croisée random subsampling, avec 50 itérations utilisant une proportion 90/10 entre les ensembles d'apprentissage et de test, sur les instances de grande taille de paysage ρ MNK en fonction des ensembles de métriques. (sampling phc) ensemble de prédicteurs obtenu à l'aide des marches adaptatives Pareto, (sampling rw) obtenu par la marche aléatoire, (sampling all) réunion de sampling phc et rw, (all) ensemble de tous les prédicteurs. (MAE) erreur moyenne absolue, (RMSE) erreur quadratique moyenne et (R^2) coefficient de détermination.

Algo.	Ens. de métri.	MAE		MSE		R^2		Rang
		avg	std	avg	std	avg	std	
GSEMO	all	0.003049	0.000285	0.000017	0.000004	0.891227	0.024584	1
	sampling all	0.003152	0.000295	0.000018	0.000004	0.883909	0.026863	1.3
	sampling rw	0.003220	0.000314	0.000019	0.000004	0.878212	0.028956	2
	sampling phc	0.003525	0.000329	0.000023	0.000006	0.854199	0.032339	3
	{ ρ , M, N, K/N}	0.003084	0.000270	0.000017	0.000003	0.892947	0.020658	1
	{ ρ , M, N}	0.009062	0.000681	0.000148	0.000021	0.065258	0.111207	4
	{M, N}	0.010813	0.000830	0.000206	0.000030	-0.303336	0.188046	5
IPLS	all	0.004290	0.000430	0.000034	0.000008	0.886568	0.026980	1
	sampling all	0.004359	0.000423	0.000035	0.000008	0.883323	0.027274	1
	sampling rw	0.004449	0.000394	0.000036	0.000008	0.879936	0.026335	1.3
	sampling phc	0.004663	0.000403	0.000039	0.000008	0.871011	0.025903	2
	{ ρ , M, N, K/N}	0.004353	0.000320	0.000033	0.000006	0.889872	0.024505	1
	{ ρ , M, N}	0.008415	0.000711	0.000119	0.000018	0.600965	0.074428	3
	{M, N}	0.016959	0.001473	0.000472	0.000077	-0.568495	0.228629	4

basés sur les métriques extraites des marches aléatoires sont légèrement de meilleure qualité que ceux basés sur les marches adaptatives, et finalement très légèrement de moins bonnes qualités qu'en considérant toutes les métriques. Dans un contexte où l'on chercherait à réduire le coût de l'échantillonnage, on pourrait utiliser seulement les métriques basées sur les marches aléatoires qui sont moins coûteuses à estimer en nombre d'évaluations.

Comme dans le cas de l'optimisation mono-objective, nous pouvons essayer d'étudier les capacités des métriques du paysage de fitness multiobjectif à sélectionner l'heuristique pertinente d'un portfolio.

PORTFOLIO D'HEURISTIQUES. Dans ce scénario de portfolio simple, nous considérons les deux heuristiques précédentes G-SEMO et PLS. Il s'agit de nouveau de sélectionner l'heuristique de meilleure performance en fonction des métriques du paysage de fitness multiobjectif. Nous utilisons pour ce problème le même modèle d'apprentissage automatique, les extremely randomized trees, mais cette fois-ci dans une perspective de classification supervisée. Avec deux algorithmes, nous pensons que la classification est un problème plus facile que la régression.

Le tableau 16 donne les erreurs sur la validation croisée pour différents ensembles de métrique. Une erreur est comptée lorsque l'algorithme prédit n'est pas celui avec les meilleures performances. La taille des deux classes de prédiction (G-SEMO et PLS) n'est pas équilibrée. Sur les 1000 instances, PLS et G-SEMO ont des performances équivalentes sur 187 instances, sur 88 instances PLS est significativement meilleur que G-SEMO, et sur 725 instances G-SEMO est significativement meilleur que PLS selon un test statistique de Mann-Whitney au risque de 5%. Un algorithme optimal de tuning qui choisit le meilleur algorithme sur l'ensemble des instances a un taux d'erreur de 8.8%.

Comme pour la régression, les paramètres {M, N} et { ρ , M, N, ρ } ne sont pas suffisant pour obtenir une prédiction convenable. Leur taux d'erreur respectif, 8.2% et 13.3%, sont supérieurs ou équivalent à une méthode de tuning optimale. Pour les autres ensembles de métriques, les taux d'erreur sont équivalents et très faibles, inférieurs à 1.5%. Cet exemple simple montre les potentialités de ce type d'approche qui tient compte de l'analyse de pay-

TABLE 16 – Taux d'erreur sur la validation croisée random subsampling (50 itérations avec 90/10 de séparation entre apprentissage et test) sur les instances de grande taille de paysage ρ MNK dans un scénario de portfolio pour différents ensembles de métriques.

Ens. de métriques	Taux d'erreur		Rang
	avg	std	
all	0.012800	0.011787	1
sampling all	0.013800	0.011586	1
sampling rw	0.015000	0.012817	1
sampling phc	0.014400	0.012961	1
$\{\rho, M, N, K/N\}$	0.013400	0.011537	1
$\{\rho, M, N\}$	0.082400	0.027669	2
$\{M, N\}$	0.132800	0.030442	3

sage de fitness multiobjectif pour effectuer un choix d'heuristique. Sur des ensembles d'instances hétérogènes, il est possible d'obtenir un meilleur choix que les méthodes de tuning qui sélectionne a priori le meilleur algorithme sur l'ensemble des instances d'apprentissage.

Nous avons atteint l'un des buts de l'analyse de paysage de fitness par la prédiction de performances pour des problèmes d'optimisation multiobjectifs de grande dimension. Les métriques extraites du paysage sont basées sur une définition simple du paysage de fitness multiobjectif. Nous allons maintenant revoir cette définition dans la section suivante.

3.2.4 Paysage de fitness multiobjectif basé sur les ensembles

La définition implicite des paysages de fitness multiobjectif utilise la fonction de fitness multiobjective sans changer l'espace de recherche ni la relation de voisinage [KC02]. Les études et les mesures qui en découlent comme nous avons pu le montrer également tendent à analyser l'espace de recherche au niveau des solutions elles-mêmes. Dans cette section, nous proposons une nouvelle définition de paysage de fitness multiobjectif dont l'espace de recherche est celui des populations comme dans les approches de résolution basées sur les indicateurs de qualité.

DÉFINITION D'UN PAYSAGE DE FITNESS MULTIOBJECTIF. La réponse d'un problème multiobjectif est un ensemble de solutions qui approxime l'ensemble Pareto optimal. Ce point de vue ensembliste n'a été clairement proposé que récemment par les travaux de Zitzler [ZTB10]. En effet, dans un certain sens, les heuristiques multiobjectives peuvent être vues comme de hill-climbers sur les ensembles de solutions. Autrement dit, dans cette approche, l'espace de recherche est défini par l'ensemble de tous les ensembles de solutions potentielles $\Sigma \subset 2^X$. Pour éviter les confusions, un élément de $\sigma \in \Sigma$ de ce nouvel espace de recherche est appelé une solution-ensemble.

Pour comparer et trouver un bon ensemble d'approximation, il existe deux types de comparaison de solution-ensemble. L'une utilise la relation de dominance étendue aux ensembles, l'autre utilise un indicateur de qualité (voir section 1.1.4). Comme dans les approches de résolution basé indicateur (IBEA, SMS-MOEA, etc. voir la section 1.1.4), nous allons utiliser un indicateur de qualité pour comparer les solutions-ensembles qui associe à chaque ensemble un nombre réel : $I : \Sigma \rightarrow \mathbb{R}$. Un problème d'optimisation multiobjectif s'exprime alors comme un problème d'optimisation mono-objectif [ZTB10] où il s'agit de trouver une solution-ensemble optimale pour l'indicateur : $\arg \max_{\sigma \in \Sigma} I(\sigma)$. Les travaux de Zitzler donnent le cadre pour définir l'optimisation multiobjective comme un problème mono-objectif. Seulement aucune mention n'est faite de ce que pourrait être une recherche locale multiobjective sur les ensembles et donc encore moins sur les relations possibles de voisinage entre ces ensembles qui permettrait d'étudier la structure de l'espace de recherche multiobjectif.

Aussi, nous avons proposé dans [Ver+10; VLD11] une définition de paysage de fitness multiobjectif basée sur l'espace de recherche des ensembles et dont il est possible de mener l'analyse par des techniques classiques en utilisant un indicateur de qualité comme fonction de fitness.

Définition 9 *Paysage de fitness multiobjectif [VLD11]* Un *paysage de fitness multiobjectif* basé sur les ensembles est un triplet (Σ, N, I) tel que :

- $\Sigma \subset 2^{\mathcal{X}}$ est l'ensemble des solutions-ensembles (\mathcal{X} est l'ensemble des solutions potentielles);
- $N : \Sigma \rightarrow 2^{\Sigma}$ est une relation de voisinage entre solutions-ensembles;
- $I : \Sigma \rightarrow \mathbb{R}$ est un indicateur de qualité, c'est-à-dire une fonction de fitness qui mesure la qualité des solutions-ensembles.

Plusieurs définitions d'espace de recherche sont possibles selon les algorithmes utilisés comme :

- les espaces de recherche à base de population de taille donnée μ :
 $\Sigma = \{\sigma \in 2^{\mathcal{X}} : |\sigma| = \mu\}$,
- les espaces de recherche utilisant des archives de taille bornée μ :
 $\Sigma = \{\sigma \in 2^{\mathcal{X}} : |\sigma| \leq \mu\}$,
- les espaces de recherche qui ne manipulent que les ensembles de solutions non mutuellement dominées :
 $\Sigma = \{\sigma \in 2^{\mathcal{X}} : \forall x, x' \in \sigma, x \not\prec x'\}$,
- les espaces de recherche qui combinent des deux dernières possibilités :
 $\Sigma = \{\sigma \in 2^{\mathcal{X}} : |\sigma| \leq \mu \text{ et } \forall x, x' \in \sigma, x \not\prec x'\}$,
- l'espace de recherche avec aucune restriction :
 $\Sigma = 2^{\mathcal{X}}$.

Les indicateurs de qualité qui servent de fonctions de fitness peuvent être aussi divers. Néanmoins, nous choisirons en priorité, pour leurs propriétés en adéquation avec la relation de dominance, l'indicateur hypervolume qui mesure le volume couvert par l'ensemble dans l'espace objectif, ou encore l'indicateur ϵ qui mesure la translation pour dominer un ensemble de référence. Dans un premier temps, on peut définir des voisinages entre solutions-ensembles qui remplacent, insèrent ou suppriment une seule solution. Ce sont des types de voisinage élémentaire que l'on trouve également dans les algorithmes évolutionnaires dont le codage des solutions est de taille variable (en programmation génétique entre autres). Par exemple, lorsqu'on considère N comme la relation de voisinage initiale entre solutions, une solution-ensemble σ' est voisine d'une solution-ensemble σ lorsque :

- pour un voisinage de remplacement :
 $\sigma' = \sigma \cup \{x'\} \setminus \{x\}$ tel que $x \in \sigma$ et $x' \in N(x)$
- pour un voisinage d'insertion :
 $\sigma' = \sigma \cup \{x'\}$ tel que $x \in \sigma$ et $x' \in N(x)$
- pour un voisinage de suppression :
 $\sigma' = \sigma \setminus \{x\}$ tel que $x \in \sigma$

Nous verrons d'autres manières de définir des voisinages dans un paragraphe suivant.

L'algorithme de recherche locale multiobjectif générique 9 (Set-based Local Search, SbLS) montre le pseudo-code de la classe d'algorithme capable d'être analysé par l'analyse du paysage de fitness multiobjectif. Un grand nombre d'algorithmes multiobjectifs actuels suit cet algorithme.

EXEMPLE D'ANALYSE DE PAYSAGE DE FITNESS MULTI-OBJECTIF. À partir de la définition du paysage de fitness multiobjectif, il est possible d'utiliser les techniques classiques d'analyse de paysage de fitness afin de mettre en avant les sources de difficulté des problèmes multiobjectifs. Aussi, nous utilisons dans cette première analyse, l'autocorrélation de la fitness et les longueurs des marches adaptatives comme métrique pour caractériser le paysage de fitness. Pour valider ces outils, nous utilisons de nouveau les paysages ρ MNK de taille

Algorithm 9 Set-based Local Search (SbSL)

Commencer avec une solution-ensemble $\sigma \in \Sigma$
Évaluer σ avec l'indicateur I
repeat
 Sélectionner une solution-ensemble voisine $\sigma' \in N(\sigma)$
 Évaluer σ' avec l'indicateur I
 if accepter(σ, σ') **then**
 $\sigma \leftarrow \sigma'$
 end if
until (stopper(σ))
Retourner l'ensemble des solutions non dominées de σ

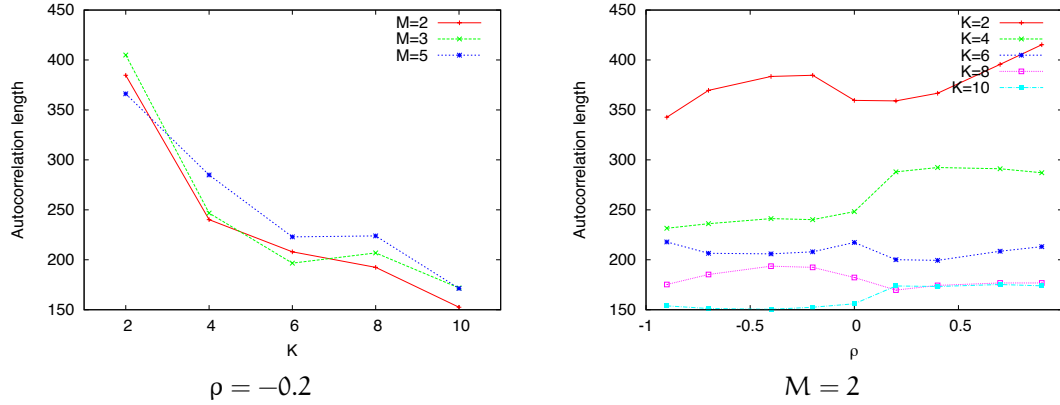


FIGURE 34 – Longueur d'autocorrélation de l'hypervolume pour des paysages ρ MNK avec $N = 64$.

intermédiaire $N = 64$. Si les outils d'analyse et la définition du paysage sont pertinents alors les métriques doivent refléter la difficulté connue de cette classe de problèmes.

La figure 34 montre la longueur d'autocorrélation qui est l'opposé de l'inverse du logarithme du premier coefficient d'autocorrélation ($-1/\log(\rho(1))$). Dans ces expériences, la fitness est donnée par l'indicateur d'hypervolume et le voisinage entre les solutions-ensembles consiste à modifier un bit d'une des solutions de l'ensemble : $\sigma' \in N(\sigma)$ si et seulement si $\exists x \in \sigma, \exists x' \in \mathcal{X}$ tel que $d_{\text{Hamming}}(x', x) = 1$ et $\sigma' = \sigma \cup \{x'\} \setminus \{x\}$. La taille maximale de ce voisinage est donc $(|\sigma| \cdot N)$. Comme dans le cas mono-objectif, et comme attendu pour ces versions multiobjectives, la longueur d'autocorrélation décroît avec le paramètre de non-linéarité K . Par contre, comme déjà observés dans l'autre type d'autocorrélation d'hypervolume local (voir tableau 30), la corrélation entre les objectifs ou le nombre d'objectifs ne sont pas corrélés avec la longueur d'autocorrélation. La rugosité du paysage multiobjectif est bien mesurée par l'autocorrélation de l'hypervolume. À noter que les longueurs sont grandes et que ce type de paysage multiobjectif semble très lisse, la variation d'hypervolume entre ensembles voisins semble faible.

Il est maintenant possible de définir simplement un optimum local pour les ensembles à l'aide de la définition existante en optimisation mono-objective.

Définition 10 *Optimum local multiobjectif basé ensemble.* Une solution-ensemble est un **optimum local multiobjectif** si et seulement si $\forall \sigma' \in N(\sigma), I(\sigma') \leq I(\sigma)$.

En suivant l'algorithme 9 des recherches locales multiobjectives, on peut définir un hill-climber first-improvement basé sur les ensembles en acceptant une solution-ensemble voisine générée aléatoirement strictement améliorante par rapport l'indicateur d'hypervolume. Partant d'un ensemble aléatoire, l'algorithme stoppe lorsqu'un optimum local multiobjectif est rencontré. La longueur de cette marche adaptative multiobjective est un indicateur de la

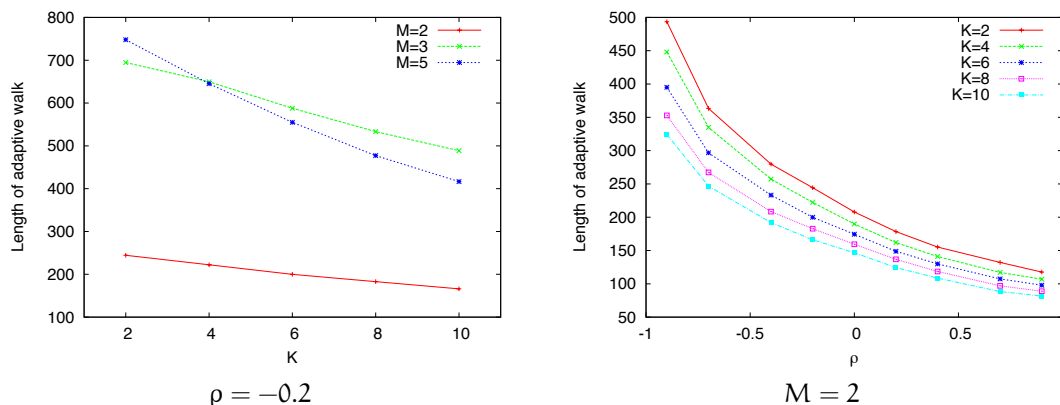


FIGURE 35 – Longueur des marches adaptatives multiobjectives pour les paysages ρMNK avec $N = 64$.

multimodalité du problème comme dans le cas mono-objectif. La figure 35 montre la longueur des marches adaptatives multiobjectives. Comme attendu, les longueurs diminuent avec le degré d'épistasie K indiquant qu'il est plus rapide un optimum local multiobjectif et que leur nombre augmente très certainement avec le degré K . Par contre de manière surprenante au contraire de l'intuition et des résultats sur les longueurs des marches adaptatives Pareto, la longueur aussi diminue aussi avec la corrélation entre objectifs. En fait, il existe de nombreuses solutions dans une solution-ensemble qui sont dominées, qui ne contribuent pas à l'hypervolume et dont les solutions voisines n'améliorent pas non plus l'hypervolume. Cette tendance est accrue lorsque la corrélation entre objectifs augmente où peu de solutions sont non dominées dans les solutions-ensembles. Il y a beaucoup de neutralité autour de ce type de solution-ensemble aléatoire qui devra être étudiée plus précisément et prise en compte dans de futures analyses [LLV14].

PROBLÈME D'EXPLORATION DE VOISINAGE. Le cadre Set-based Local Search défini par l'algorithme 9 permet de comparer différents types de voisinage [Bas+13]. Tout d'abord, formalisons une propriété attendue pour les voisinages entre solutions-ensembles.

Définition 11 *Voisinage conforme.* Une relation de voisinage \mathcal{N} est dite **conforme** à la relation de voisinage \mathcal{N} si et seulement si $\forall \sigma \in \Sigma, \forall \sigma' \in \mathcal{N}(\sigma), \forall x' \in (\sigma' \setminus \sigma), \exists x \in \sigma$ tel que $x' \in \mathcal{N}(x)$

Les voisinages conformes sont ceux qui utilisent la relation de voisinage initiale entre solutions pour générer une solution-ensemble voisine. Parmi toutes les possibilités de relation de voisinage conforme, on peut classer ces voisinages en fonction du nombre de solutions utilisées de l'ensemble pour produire de nouvelle solution et du nombre de solutions produite à partir de chacune de celles-ci. De façon générale, quatre types de voisinages peuvent être définis à partir d'une relation de voisinage initiale entre solutions \mathcal{N} . À partir d'une solution-ensemble σ , une solution-ensemble voisine peut être définie par :

1. $\sigma' \in \mathcal{N}^{(1,1)}(\sigma) \Rightarrow |\sigma' \setminus \sigma| \leq 1$ et $\forall x' \in \sigma' \setminus \sigma, \exists x \in \sigma, x' \in \mathcal{N}(x)$.
2. $\sigma' \in \mathcal{N}^{(1,*)}(\sigma) \Rightarrow \exists x \in \sigma, \forall x' \in \sigma' \setminus \sigma, x' \in \mathcal{N}(x)$.
3. $\sigma' \in \mathcal{N}^{(*,1)}(\sigma) \Rightarrow \exists \sigma_0 \subset \sigma, \exists \phi : \sigma_0 \rightarrow \sigma' \setminus \sigma$ avec $\forall x' \in \sigma' \setminus \sigma, \exists! x \in \sigma_0$ tel que $\phi(x) = x'$ et $x' \in \mathcal{N}(x)$.
4. $\sigma' \in \mathcal{N}^{(*,*)}(\sigma) \Rightarrow \forall x' \in \sigma' \setminus \sigma, \exists x \in \sigma$ tel que $x' \in \mathcal{N}(x)$.

La figure 36 illustre ces différents types de voisinage. Pour faire un lien avec les heuristiques que nous avons déjà rencontré, l'heuristique G-SEMO utilise un voisinage de type $\mathcal{N}^{(1,1)}$ alors que PLS utilise un voisinage de type $\mathcal{N}^{(1,*)}$. Le nombre de voisins pour les différents types de voisinage peut être exponentiel en fonction de la taille des solutions-ensembles. Tout l'art des heuristiques consiste à choisir un voisinage de qualité et une façon de l'explorer de faible complexité.

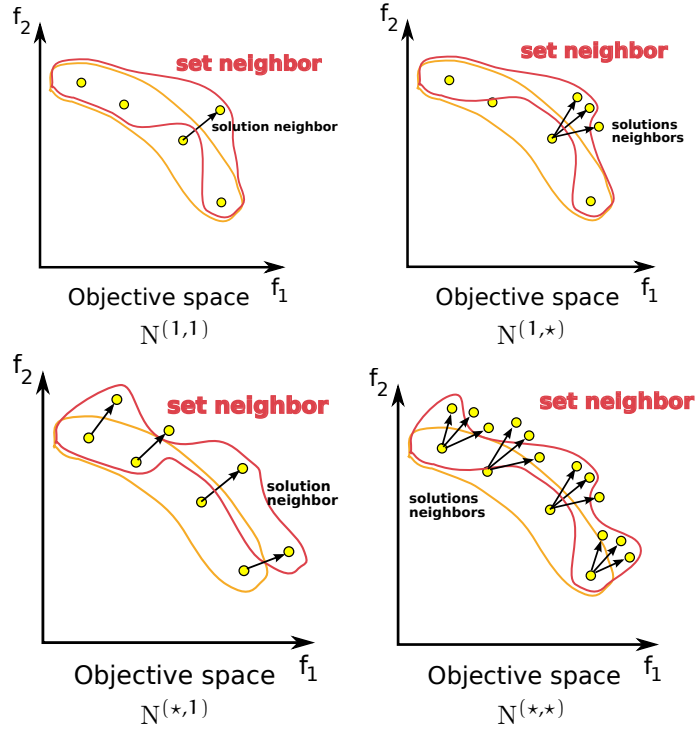


FIGURE 36 – Illustration des différents types de voisinage N basés sur les ensembles dans l'espace objectif de dimension 2. Les différents types produisent les nouvelles solutions soit à partir d'une seule ou de plusieurs solutions et le nombre de solutions produites est une ou plusieurs à partir d'une même solution.

Pour identifier les propriétés des différents types de voisinage, nous avons comparé ces voisinages dans le cadre d'une même heuristique simple de type hill-climber first-improvement. L'espace de recherche est celui des ensembles de taille μ , aussi afin de maintenir cette taille nous utilisons un procédé, noté hgh pour hypervolume-based greedy heuristic [ZTB10], de réduction qui supprime itérativement de l'ensemble la solution qui dégrade le moins l'hypervolume jusqu'à obtenir un ensemble de taille μ . Les voisinages expérimentés sont donc :

1. $\hat{N}_{\text{hgh}}^{(1,1)}(\sigma) = \{\text{hgh}(\sigma \cup \{x'\}) : x \in \sigma, x' \in N(x)\}$.
2. $\hat{N}_{\text{hgh}}^{(1,*)}(\sigma) = \{\text{hgh}(\sigma \cup N(x)) : x \in \sigma\}$.
3. $\hat{N}_{\text{hgh}}^{(*,1)}(\sigma) = \{\text{hgh}(\sigma \cup_{x \in \sigma} \{x'_x\}) : \forall x \in \sigma, x'_x \in N(x)\}$.
4. $\hat{N}_{\text{hgh}}^{(*,*)}(\sigma) = \{\text{hgh}(\sigma \cup_{x \in \sigma} N(x))\}$.

Les paysages ρMNK servent de nouveau de cadre d'étude. Du fait du coût de l'heuristique hgh, les comparaisons sont réalisées à temps identiques de 10^4 secondes. Le tableau 17 résume les hypervolumes obtenus. Le voisinage de type $(*,*)$ obtient globalement les meilleures performances. Les voisinages de type $(1,*)$ utilisés par exemple dans PLS n'obtiennent finalement pas de bonnes performances. Un voisinage de type $(*,1)$, peu utilisé dans les algorithmes actuels, peut obtenir des hypervolumes compétitifs. Néanmoins, ces résultats sont à mettre en parallèle avec la complexité de chaque voisinage et le temps de résolution. La figure 37 donne l'hypervolume atteint en fonction du temps de calcul. Le temps de convergence du voisinage de type $(*,*)$ est notamment le plus lent, alors que celui $(*,1)$ converge de manière beaucoup plus rapide.

Ce cadre d'étude permet de comparer les voisinages entre ensembles eux-mêmes et les opérateurs associés indépendamment d'une heuristique particulière comme on pourrait le

TABLE 17 – Hypervolume moyen obtenu par les SBLS pour différents types de voisinage et à titre de comparaison par NSGA-II. Les hypervolumes les plus grands sont les meilleurs.

		$N_{\text{hgh}}^{(1,1)}$	$N_{\text{hgh}}^{(1,*)}$	$N_{\text{hgh}}^{(*,1)}$	$N_{\text{hgh}}^{(*,*)}$	NSGA-II
$\rho = -0.5$						
N = 256	K = 2	0.5068	0.5019	0.5074	0.5074	0.4438
	K = 4	0.5189	0.5124	0.5203	0.5231	0.4806
N = 512	K = 2	0.4915	0.4880	0.4897	0.4905	0.4012
	K = 4	0.5186	0.5106	0.5180	0.5187	0.4440
$\rho = 0.0$						
N = 256	K = 2	0.5161	0.5129	0.5169	0.5193	0.5032
	K = 4	0.5453	0.5395	0.5444	0.5433	0.5373
N = 512	K = 2	0.5155	0.5095	0.5156	0.5159	0.4791
	K = 4	0.5417	0.5396	0.5441	0.5430	0.5201
$\rho = 0.5$						
N = 256	K = 2	0.5264	0.5231	0.5276	0.5296	0.5333
	K = 4	0.5518	0.5494	0.5539	0.5535	0.5593
N = 512	K = 2	0.5213	0.5161	0.5208	0.5228	0.5184
	K = 4	0.5526	0.5444	0.5528	0.5538	0.5589

faire classiquement lors d'une analyse de paysage de fitness d'un problème d'optimisation mono-objectif. La nouvelle définition de paysage de fitness multiobjectif permet ainsi de comprendre et d'étudier la structure d'un problème d'optimisation multiobjective.

3.2.5 Approches scalaires par décomposition

Les approches scalaires par décomposition comme l'algorithme MOEA/D (algorithme 3 section 1.1.4) coordonnent la recherche de multiples optimisations mono-objectives. Les points critiques de conception qui vont influencer les performances sont donc (i) le choix des types de fonctions scalaires qui agrègent les objectifs, (ii) le choix des directions de recherche de chaque fonction scalaire qui décompose le problème multiobjectif en de multiples problèmes mono-objectifs et (iii) la coordination des différentes recherches mono-objectives. Dans cette section, nous étudions les points critiques du choix de la fonction scalaire et des directions de recherche, puis nous proposons une coopération originale des directions de recherche dans le cadre d'un algorithme distribué. Enfin, nous proposons un algorithme multiobjectif hydrique qui combine une recherche locale Pareto avec une approche scalaire pour résoudre un problème classique mono-objectif (UBQP) dont nous avons proposé la variante multiobjective.

FONCTION SCALAIRES ET DIRECTION DE RECHERCHE DISTRIBUÉE. Les fonctions scalaires agrègent les objectifs en une nouvelle fonction mono-objective. De nombreuses fonctions scalaires ont été proposées dans la littérature (voir par exemple [Mie99] pour une revue). Parmi les plus utilisées, nous trouvons la fonction qui calcule une somme pondérée des objectifs ou encore la fonction de Tchebichev (augmentée ou non) qui correspond à la norme infinie des objectifs. Ces fonctions ont été étudiées à l'aide de comparaisons directes de leurs performances à atteindre de bonnes approximations en fonction de leurs paramètres respectifs [IAN13]. Dans cette section, nous étudions individuellement les différentes fonctions scalaires à travers les dynamiques de recherche qu'elles entraînent dans l'espace objectif en particulier en fonction des directions de recherche [Der+14a].

L'étude a pour but de montrer que les dynamiques de recherche des approches scalaires dépendent de la forme, c'est-à-dire de la direction, des lignes de niveau déduites des fonctions scalaires, plutôt que du type particulier de la fonction scalaire ou de la valeur de

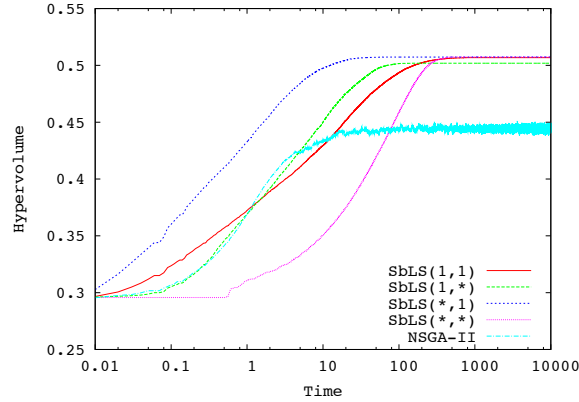


FIGURE 37 – Hypervolume (à maximiser) en fonction du temps d'exécution (en seconde) pour les différents types de voisinage. Paysages ρ MNK avec $M = 2$, $N = 256$, $K = 2$ et $\rho = -0.5$.

TABLE 18 – Résumé des fonctions scalaires considérées, des paramètres qui établissent la correspondance avec la fonction générique S_{gen} et des angles polaires des lignes d'iso-fitness avec l'axe des abscisses.

fonction scalaire	paramètres dans S_{gen}	angle avec $[O, x)$	référence
$WS(z) = w_1 \bar{z}_1 - z_1 + w_2 \bar{z}_2 - z_2 $	$\alpha = 0, \epsilon = 1$	$\theta_1 = \arctan\left(-\frac{w_1}{w_2}\right)$ $\theta_2 = \frac{\pi}{2} + \arctan\left(\frac{w_1}{w_2}\right)$	[Mie99, Eq. 3.1.1]
$T(z) = \max\{\lambda_1 \bar{z}_1 - z_1 , \lambda_2 \bar{z}_2 - z_2 \}$	$\alpha = 1, \epsilon = 0$	$\theta_1 = 0$ $\theta_2 = \pi/2$	[Mie99, Eq. 3.4.2]
$S_{\text{aug}}(z) = T(z) + \epsilon (\bar{z}_1 - z_1 + \bar{z}_2 - z_2)$	$\alpha = 1,$ $w_1 = w_2 = 1$	$\theta_1 = \arctan\left(-\frac{\epsilon}{\lambda_1 + \epsilon}\right)$ $\theta_2 = \frac{\pi}{2} + \arctan\left(\frac{\epsilon}{\lambda_2 + \epsilon}\right)$	[Mie99, Eq. 3.4.5]
$S_{\text{norm}}(z) = (1 - \epsilon)T(z) + \epsilon WS(z)$	$\alpha = 1 - \epsilon,$ $w_i = 1/\lambda_i$	$\theta_1 = \arctan\left(-\frac{\epsilon w_1}{(1 - \epsilon)\lambda_2 + \epsilon w_2}\right)$ $\theta_2 = \frac{\pi}{2} + \arctan\left(\frac{\epsilon w_2}{(1 - \epsilon)\lambda_1 + \epsilon w_1}\right)$	[Der+14a]

leurs paramètres comme dans [IAN13]. Pour cela, nous restreignons l'étude au cas biobjectif et nous définissons une fonction scalaire générique dont l'expression couvre plusieurs fonctions couramment utilisées. Puis nous montrons expérimentalement sur les paysages ρ MNK que la convergence est guidée globalement par les lignes de niveau de la fonction scalaire. Nous considérons donc la fonction scalaire S_{gen} à minimiser suivante définie pour un problème bi-objectif :

$$S_{\text{gen}}(z) = \alpha \cdot \max\{\lambda_1 \cdot |\bar{z}_1 - z_1|, \lambda_2 \cdot |\bar{z}_2 - z_2|\} + \epsilon (w_1 \cdot |\bar{z}_1 - z_1| + w_2 \cdot |\bar{z}_2 - z_2|)$$

où $z = (z_1, z_2)$ est le vecteur objectif d'une solution, $\bar{z} = (\bar{z}_1, \bar{z}_2)$ est un point de référence utopique, $\lambda_1, \lambda_2, w_1$, et $w_2 > 0$ les coefficients qui paramètrent la direction de recherche et $\alpha \geq 0$ et $\epsilon \geq 0$ des paramètres qui donnent le compromis entre la partie linéaire et non linéaire de la somme. La fonction S_{gen} est une généralisation des cas particuliers de somme pondérée, noté WS, lorsque $\epsilon = 0$, ou encore de Tchebichev, noté T, lorsque $\alpha = 0$, et Tchebichev augmenté, noté S_{aug} , pour des valeurs intermédiaires et $w_1 = w_2 = 1$ (voir le tableau 18 pour un résumé des équivalences).

Chaque fonction scalaire définit des courbes où la fonction prend les mêmes valeurs que nous appelons des lignes d'iso-fitness. Intuitivement, la dynamique de recherche selon une fonction scalaire dans l'espace objectif est un compromis entre la distribution des solutions obtenues par la variation aléatoire (opérateur de recherche local) et la direction du gradient des fonctions scalaires qui est orthogonale à ces lignes d'iso-fitness. Pour la fonction S_{gen} ,

ces courbes sont de deux demi-droites symétriques par rapport à la direction de recherche dont on peut calculer l'angle polaire avec l'axe des abscisses :

Proposition 1 Soit \bar{z} un point de référence utopique, $\lambda_1, \lambda_2, w_1$, et $w_2 > 0$ les coefficients de poids, $\alpha \geq 0$ et $\epsilon \geq 0$, avec $\alpha\epsilon > 0$. Alors les angles polaires entre les lignes d'iso-fitness de S_{gen} et l'axe f_1 des abscisses sont $\theta_1 = \arctan(-\frac{\epsilon w_1}{\alpha \lambda_2 + \epsilon w_2})$ et $\theta_2 = \frac{\pi}{2} + \arctan(\frac{\epsilon w_2}{\alpha \lambda_1 + \epsilon w_1})$.

La figure 38 donne des exemples d'optimisation sur les paysages ρ MNK de fonctions scalaires S_{gen} par un algorithme $(1 + \lambda)$ -EA. Partant d'une solution aléatoire, la recherche suit en partie la direction du gradient orienté par les lignes d'iso-fitness modulées par les solutions produites par l'opérateur de recherche local (bit-flip avec un taux de $1/N$).

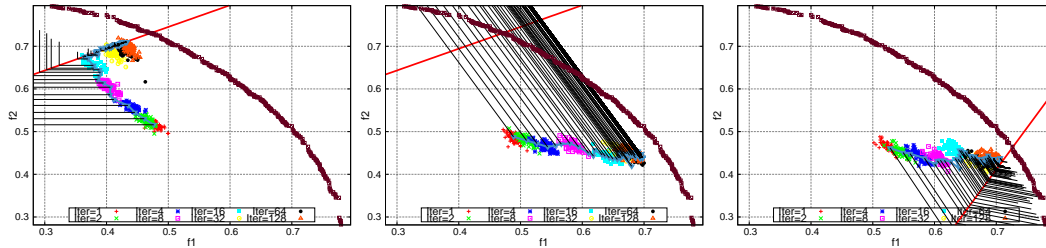


FIGURE 38 – Exemple d'optimisation par un algorithme $(1 + \lambda)$ -EA pour différentes directions de recherche δ (line rouge) et différentes valeurs de ϵ (S_{norm} , $\rho = -0.7$). La meilleure approximation connue du front de Pareto est représentée des points en noir ainsi que les lignes d'iso-fitness pour chaque fonction scalaire. Les λ solutions voisines sont représentées pour des générations successives (une couleur par génération), la première génération est en rouge aux environs du point de coordonnées $(0.5, 0.5)$. À gauche : $\epsilon = 0$, $\delta = \frac{3}{10} \cdot \frac{\pi}{2}$. Au centre : $\epsilon = 1$, $\delta = \frac{7}{10} \cdot \frac{\pi}{2}$. À droite : $\epsilon = 0.6$, $\delta = \frac{7}{10} \cdot \frac{\pi}{2}$.

Nous pouvons montrer que les directions suivies par un algorithme de type $(1 + \lambda)$ -EA sont guidées par la direction des lignes d'iso-fitness. La figure 39 montre la corrélation entre l'angle polaire formé par la solution finale obtenue et l'angle polaire des lignes d'iso-fitness avec l'axe des abscisses. Pour les deux types de fonctions, Tchebichev normalisé S_{norm} et Tchebichev augmenté S_{aug} , la corrélation est très forte, quelque soit la valeur des autres paramètres. Pour le montrer d'une autre manière, nous avons choisi pour ces deux types de fonctions, a priori considérées comme différentes, les valeurs de paramètres pour obtenir les mêmes directions de lignes iso-fitness, et nous avons comparé l'angle de la position finale de la solution dans la figure 39 (à droite). Bien que la valeur (norme) des gradients soit certainement différente, les positions des solutions finales sont identiques.

Pour la conception d'un algorithme à décomposition, le choix des fonctions scalaires et de leur orientation doit résider dans la forme des courbes d'iso-fitness où la fonction prend les mêmes valeurs numériques. Bien que ce résultat paraisse très naturel, il permet très certainement de reconsidérer certains choix dans les algorithmes pour les améliorer. De plus, en connaissant et estimant la dynamique locale de l'algorithme, les directions de ces lignes d'iso-fitness, mieux que les paramètres des fonctions scalaires directement, pourront être réglé de manière adaptative en cours de l'optimisation.

ALGORITHME DISTRIBUÉ MULTIOBJECTIF ADAPTATIF. Cette section ne concerne pas l'analyse de paysage de fitness multiobjectif, pas de nouvel outil ou d'étude d'un problème particulier. Il s'agit de proposer un algorithme avec une approche scalaire dans un environnement de calcul distribué où chaque optimisation locale mono-objective communique et se coordonne localement avec les autres optimisations mono-objectives de façon à obtenir une optimisation globale performante [Der+14c].

Dans cet algorithme, nous ne cherchons pas à communiquer entre les recherches mono-objectives les solutions qui pourraient être utiles aux autres problèmes mono-objectifs comme

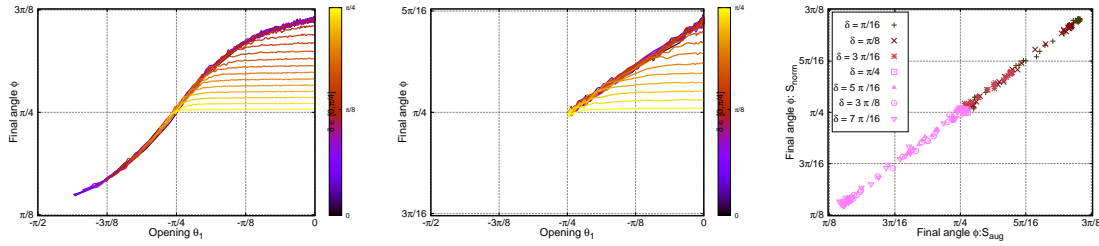


FIGURE 39 – À gauche et au centre, nuage de points entre l’angle polaire final ϕ de la solution obtenue par un $(1 + \lambda)$ -EA et l’angle polaire des lignes d’iso-fitness (opening θ_1), respectivement pour S_{norm} et S_{aug} avec une corrélation entre objectifs de $\rho = -0.7$. À droite, nuage de points entre les angles polaires finaux $(\phi(S_{norm}), \phi(S_{aug}))$ pour des angles polaires identiques des lignes d’iso-fitness.

Algorithm 10 DLBS – Pseudo-code pour tous les n nœuds de calcul v^i de la structure linéaire de communication L_n

```

 $x^i \leftarrow$  initial solution corresponding to node  $v^i$ 
repeat
  /* communicate positions */
   $z^i \leftarrow (z_1^i, z_2^i)$  the position of solution  $x^i$  in the bi-objective space,  $z^i = f(x^i)$ 
  Send  $z^i$  to neighboring nodes
   $Z^i \leftarrow$  receive neighboring positions
  /* variation */

   $S^i \leftarrow$  New_Solutions( $x^i$ )
  /* selection for replacement */
   $x^i \leftarrow$  Select( $S^i, \mathcal{LF}Z^i$ )
until STOPPING_CONDITION

```

dans le cadre d’un algorithme MOEA/D. Cette question importante est mise de côté dans ce travail. Nous cherchons à coordonner les directions de recherche, c’est-à-dire à proposer au cours de la recherche les bonnes fonctions scalaires et leurs paramètres. Une technique parmi les plus simples serait de fixer les directions de recherche initialement au début de la recherche puis de les laisser fixes et indépendantes pendant tout le temps de l’optimisation. Cette technique est souvent utilisée par les algorithmes qui cherchent alors à avoir la meilleure répartition initiale de direction de recherche [ZLo7]. Il est pourtant possible d’obtenir de meilleures performances en adaptant les directions de recherche en fonction de l’état de la recherche [Qi+14]. Seulement ces méthodes adaptatives ne permettent pas la distribution de l’algorithme pour un grand nombre de cœurs de calcul [Dur+11].

Nous adoptons l’algorithme général dédié à l’optimisation biobjective décrit dans l’algorithme 10. Une structure linéaire de communication est définie entre les solutions : chaque solution peut communiquer avec ses deux solutions voisines excepté pour les solutions aux extrémités. Chaque solution est associée à une fonction locale scalaire \mathcal{LF} et une direction de recherche. Chaque solution envoie sa position dans l’espace objectif à ses deux solutions voisines (ou une seule solution voisine dans le cas des solutions aux extrémités). En fonction de ces positions, les paramètres de la fonction locale scalaire sont adaptés, puis une nouvelle solution est sélectionnée selon la nouvelle fonction scalaire. De cette façon, les communications entre solutions sont réduites (seulement 4 nombres réels) et les directions de recherche définies par les paramètres sont localement adaptées.

Les fonctions locales scalaires testées sont de deux types. Chacune prend en compte la position des solutions voisines et tente de trouver des solutions dans une direction orthogonale intermédiaire entre les deux solutions voisines. La figure 40 résume les deux types de fonctions scalaires. Dans les deux cas, pour les solutions à l’extrémité de la communica-

tion les fonctions scalaires sont respectivement les fonctions objectifs f_1 et f_2 . Le premier de type fonction scalaire \mathcal{LF}_{OD} est une somme pondérée des objectifs. Les poids de la somme sont choisis de manière à orienter les lignes d'iso-fitness parallèlement aux deux solutions voisines et ainsi de trouver une nouvelle solution dans la direction orthogonale à celles-ci. L'autre type de fonction scalaire \mathcal{LF}_H est basé sur l'hypervolume local. À partir des deux solutions voisines, le point de référence est calculé comme le minimum des abscisses et ordonnées (dans le cas de maximisation). La fonction scalaire correspond à l'aire du rectangle depuis ce point de référence. Les lignes d'iso-fitness sont alors des hyperboles dont l'axe focal qui indique le gradient passe par le point de référence et est orienté à 45 degrés. Ces deux fonctions scalaires ont pour but de répartir localement au mieux les solutions dans une direction «entre» les deux solutions voisines.

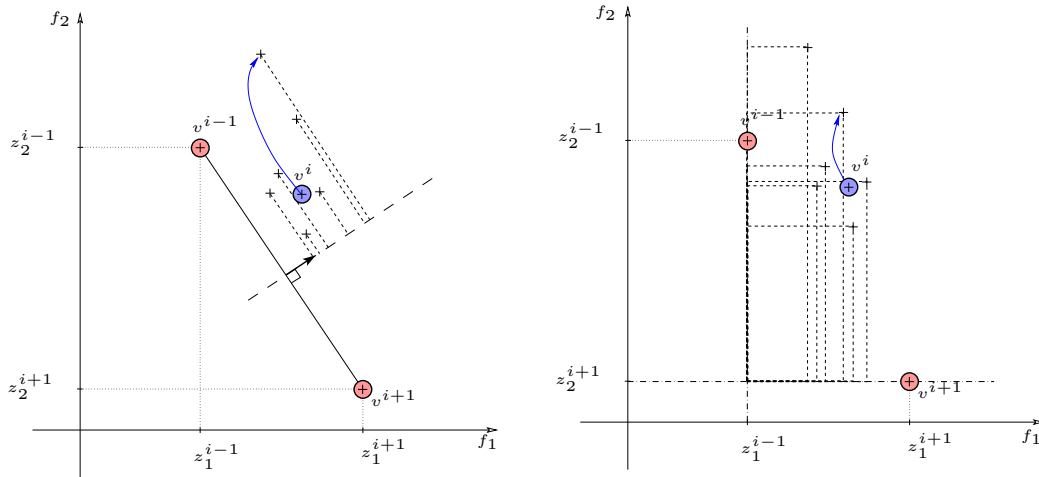


FIGURE 40 – Illustration pour une solution i de la sélection pour le remplacement en utilisant les fonctions locales scalaires de somme pondérée dans la direction orthogonale \mathcal{LF}_{OD} (gauche) et d'hypervolume \mathcal{LF}_H (droite). Les croix sans cercle sont les λ solutions obtenues par variation aléatoire et la flèche indique la solution sélectionnée.

Les deux versions de l'algorithme $DLBS_{OD}$ et $DLBS_H$ ont été testées sur les paysages ρMNK pour différentes valeurs de longueur de chaîne de bits $N \in \{128, 256\}$, de degré d'épistasie $K \in \{4, 8\}$ et de corrélation entre objectifs $\rho \in \{-0.7, 0, 0.7\}$. Le nombre de cœurs de calcul n est égal à celui du nombre de directions $\mu = 128$. Les algorithmes ont été comparés à une version $Prws$ dont les calculs sont totalement indépendants avec une fonction scalaire de somme pondérée dont les directions de recherche sont équitablement réparties dans le quart de plan supérieur. Ils ont été également comparés à un algorithme séquentiel global qui ne s'exécute que sur un seul nœud de calcul $HEMO$ qui est une variante $(\mu + \lambda)$ du $SMS-MOEA$ [BNE07]. Dans cet algorithme $SMS-MOEA$, la sélection utilise le rang de dominance comme premier critère puis la contribution d'hypervolume comme second critère. Une partie des résultats sont reportés dans la table 19. Sur toutes les instances, la version de l'algorithme $DLBS_{OD}$ avec la somme pondérée a de meilleurs résultats que la version $DLBS_H$ basée sur l'hypervolume. La version avec l'hypervolume doit sans doute être reconsidérée avec l'étude précédente (qui a eu lieu postérieurement à ces travaux) sur les directions de recherche dans les approches scalaires pour définir des lignes d'iso-fitness mieux orientées. La figure 41 montre un exemple de trajectoire dans l'espace objectif de l'algorithme $DLBS_{OD}$ basé sur la somme pondérée. Lorsqu'on compare les performances avec l'algorithme $Prws$ où les directions sont statiques, l'algorithme proposé $DLBS_{OD}$ avec les directions adaptatives obtient de bien meilleures performances ce qui suggère que le réglage adaptatif des directions est pertinent. Enfin, les performances de $DLBS_{OD}$ sont meilleures que l'algorithme

TABLE 19 – Comparaison des différents algorithmes mesurée par la différence d'hypervolume (I_H^-) pour les paysages ρ MNK. La première valeur indique le nombre d'algorithmes qui sont significativement meilleurs que l'algorithme considéré. Les nombres entre parenthèses sont les valeurs de différence d'hypervolume (les valeurs proches de zéro sont les meilleures). La taille de population est $\mu = 128$.

ρ	N	K	DLBS _{OD} ($n = 128, \mu = 128$)		DLBS _H ($n = 128, \mu = 128$)		PIWS		HEMO	
					$I_H^- (\times 10^{-2})$					
-0.7	128	4	0	(1.846)	0	(1.919)	2	(2.365)	3	(3.667)
-0.7	128	8	0	(1.914)	0	(1.984)	2	(2.275)	2	(2.375)
-0.7	256	4	0	(1.529)	1	(1.618)	2	(1.779)	3	(4.190)
-0.7	256	8	0	(1.580)	1	(1.680)	2	(1.771)	3	(2.906)
-0.7	512	4	0	(0.985)	1	(1.107)	2	(1.253)	3	(3.352)
-0.7	512	8	0	(1.248)	1	(1.318)	2	(1.461)	3	(2.836)
0.0	128	4	1	(1.778)	1	(1.876)	3	(2.491)	0	(1.406)
0.0	128	8	1	(1.677)	2	(1.821)	3	(2.178)	0	(1.043)
0.0	256	4	0	(1.272)	2	(1.390)	3	(1.613)	0	(1.284)
0.0	256	8	1	(1.219)	2	(1.349)	3	(1.582)	0	(0.667)
0.0	512	4	0	(1.038)	1	(1.115)	3	(1.339)	0	(1.068)
0.0	512	8	1	(1.107)	2	(1.214)	3	(1.379)	0	(0.822)
+0.7	128	4	1	(1.518)	2	(1.651)	3	(2.277)	0	(1.255)
+0.7	128	8	0	(0.629)	2	(0.743)	3	(0.968)	0	(0.567)
+0.7	256	4	1	(0.618)	2	(0.695)	3	(0.804)	0	(0.378)
+0.7	256	8	1	(0.526)	2	(0.609)	3	(0.721)	0	(0.329)
+0.7	512	4	1	(0.521)	2	(0.571)	3	(0.647)	0	(0.252)
+0.7	512	8	1	(0.556)	2	(0.623)	3	(0.673)	0	(0.316)

global HEMO lorsque les objectifs sont en conflit ($\rho < 0$) et le contraire lorsque les objectifs sont corrélés positivement ($\rho > 0$). DLBS est capable de trouver des solutions plus diverses lorsque les objectifs sont en conflits même sans information globale, alors que HEMO est capable de concentrer la recherche dans la région centrale de l'espace objectif ce qui est a priori une bonne stratégie lorsque les objectifs sont corrélés et que le front de Pareto se situe à cet endroit.

Nous avons aussi comparé les temps d'exécution des différents algorithmes. Naturellement par la nature distribuée de l'algorithme DLBS, son temps d'exécution est de l'ordre de plusieurs centaines de fois inférieures à l'algorithme séquentiel basé sur le calcul d'hypervolume HEMO. La taille des messages communiqués au sein de DLBS est faible et le temps de communication devient faible comparé au temps de calcul de la fonction d'évaluation. La figure 42 (gauche) montre le temps de calcul de la fonction d'évaluation par rapport au temps d'exécution global pour différente valeur de N c'est-à-dire de complexité de calcul. Pour une dimension de $N = 512$, déjà plus de 90% du temps est employé au calcul plutôt qu'à la communication. Enfin, nous pouvons observer une croissance linéaire du speed-up, c'est-à-dire du ratio entre le temps d'exécution parallèle sur n nœud et du même algorithme exécuté séquentiellement sur un seul nœud.

Nous avons donc proposé un algorithme distribué qui permet d'adapter les directions des fonctions scalaires selon l'état de la recherche. Les premiers résultats montrent l'avantage par rapport à des directions statiques. Même si ce travail ne concerne pas directement l'analyse des paysages de fitness, il encourage à l'analyse les fonctions scalaires comme nous l'avons fait dans le travail précédent, mais aussi à développer de nouveaux outils d'analyse qui permettent de mieux comprendre la coopération entre différentes optimisations mono-objectives.

PROBLÈME QUADRATIQUE BINAIRE NON CONTRAINT MULTIOBJECTIF. Le problème quadratique binaire non contraint, noté UBQP pour Unconstrained Binary Quadratic Programming, est l'un des problèmes remarquables pour l'optimisation mono-objective [Koc+14]. Étant donné une collection de n éléments telle qu'un profit est associé à chaque paire d'éléments, le profit pouvant être positif, négatif ou nul, le problème UBQP consiste à déterminer le sous-ensemble d'éléments qui maximise la somme des profits. Le problème UBQP est un problème binaire NP-difficile dans lequel de nombreux problèmes combinatoires peuvent se

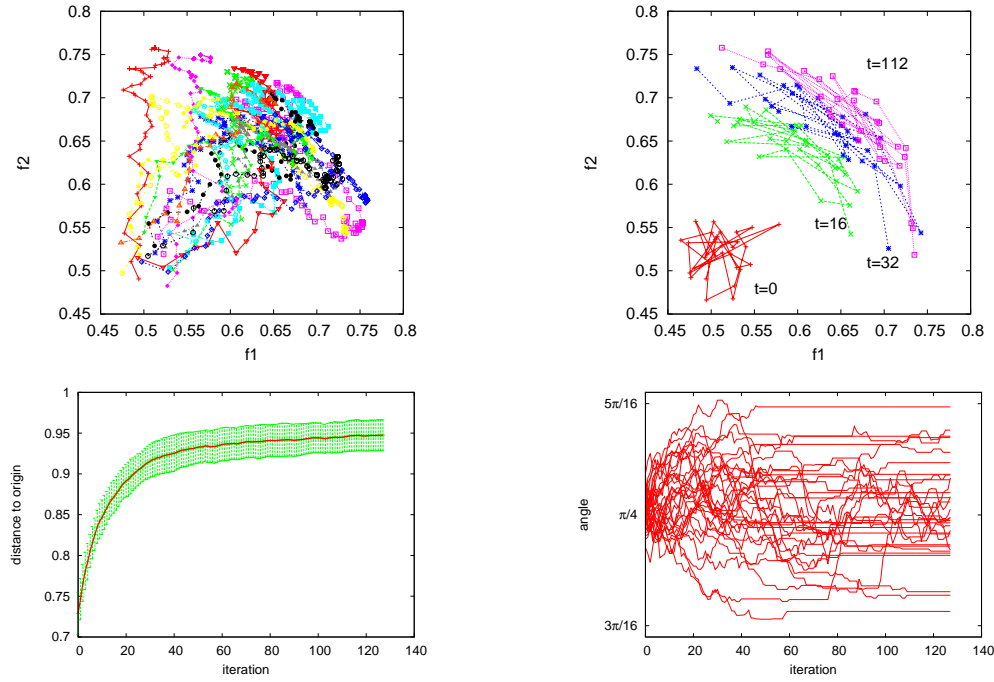


FIGURE 41 – Dynamique de $DLBS_{OD}$ pour $n = 128$ nœuds et $\mu = 128$ directions pour les paysages ρMNK avec $K = 4$, $\rho = 0.0$, $N = 128$ et $n = \mu = 128$. En haut à gauche, l'évolution des trajectoires individuelles dans l'espace objectif de chaque optimisation mono-objective. En haut à droite, l'évolution du graphe de communication qui progressivement s'élargit pour couvrir une bonne approximation du front de Pareto. En bas à gauche : évolution de la distance moyenne de toutes les solutions à l'origine qui se rapproche donc du front de Pareto. En bas à droite : évolution des différents angles polaires des solutions qui occupent toute l'étendue du front de Pareto en fin de convergence.

représenter comme la coloration de graphe, le problème de coupe maximum, le problème de partitionnement, parmi bien d'autres [Koc+14]. Dans ce travail, nous proposons d'abord une version multiobjective de ce problème UBQP [LVH14]. Puis, nous montrons comment l'analyse du paysage de fitness conduit à définir un algorithme d'optimisation hybride efficace qui combine une approche scalaire et une approche Pareto [Lie+15b].

Le problème quadratique binaire non contraint multiobjectif (mUBQP) se formule par la fonction objective multiobjective suivante dont chaque objectif f_k est à maximiser :

$$\forall x \in \{0, 1\}^n, k \in \{1, \dots, m\}, f_k(x) = \sum_{i=1}^n \sum_{j=1}^n q_{ij}^k x_i x_j$$

où $f = (f_1, f_2, \dots, f_m)$ est le vecteur de fonctions objectifs avec $m \geq 2$, n la dimension du problème et $Q^k = (q_{ij}^k)$ m matrices carrées de dimension n de nombres positifs, négatifs ou nuls. L'espace de recherche \mathcal{X} du problème est l'ensemble des chaînes binaires de longueur n : $\{0, 1\}^n$. Les instances classiques du problème UBQP considèrent généralement un paramètre d de densité de nombre égale à zéro dans la matrice Q . Également dans ces instances, les nombres non nuls sont des nombres entiers distribués uniformément dans l'intervalle $[-100, 100]$. Pour les instances mUBQP, nous avons respecté cela en ajoutant une corrélation entre les nombres non nuls. Selon le taux $d \in [0, 1]$, pour toutes les m matrices Q_k , les nombres q_{ij}^k sont nuls et la corrélation ρ entre les nombres non nuls est définie de la même

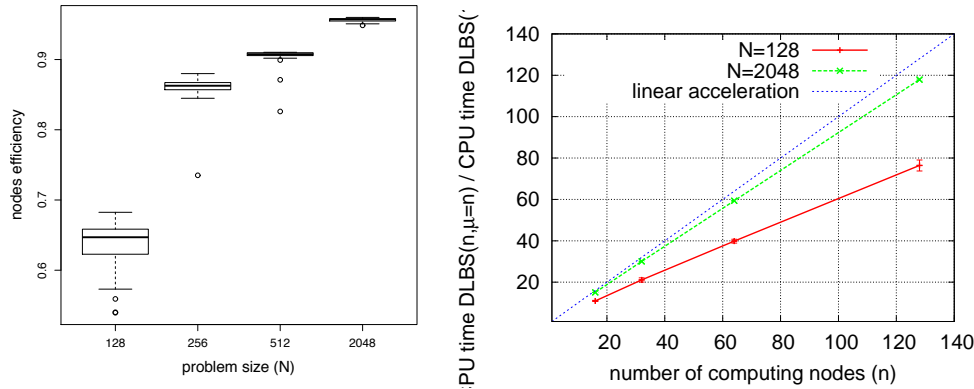


FIGURE 42 – À gauche, influence du temps d'évaluation de la fitness : ratio moyen du temps d'évaluation sur le temps global en fonction de la dimension du problème N . Le nombre de nœuds est $\mu = n = 128$, $\lambda = N$ et le nombre total d'évaluations est $\lambda \times N$ par nœud de calcul. À droite, ratio de l'accélération pour DLBS avec $n = \mu$ nœuds de calcul comparé à DLBS avec un seul nœud de calcul en fonction de la taille de la population μ . Les résultats sont pour $N \in \{128, 2048\}$, $\lambda = N$.

manière que dans les paysages ρ MNK. Les instances sont également répertoriées dans la plateforme mocoBench³ avec les instances ρ MNK.

Nous avons montré que le problème mUBQP est insoluble (intractable), c'est-à-dire que le nombre de solutions de l'ensemble Pareto optimal peut être exponentiel en fonction de la dimension du problème. Nous avons également analysé certaines propriétés du paysage de fitness multiobjectif pour de petites instances du problème avec $n = 18$. La figure 43 montre les principales métriques en fonction du paramètre de densité d et du paramètre de corrélation ρ . En résumé, le paramètre de corrélation ρ permet de régler précisément la corrélation entre les objectifs. Le nombre de solutions Pareto optimales diminue avec la corrélation ρ et ne varie pas avec la densité d . La proportion du nombre de solutions supportées augmente avec la corrélation entre objectifs. Le ratio de la taille de la plus grande composante connexe du graphe Pareto est important au-dessus de 60%, et la distance minimale pour rendre connexe le graphe Pareto est seulement environ égal à 2.5. Enfin, le nombre d'optima locaux Pareto diminue avec la corrélation entre objectifs. Pour ces métriques, la densité d des matrices influence très peu les métriques du paysage de fitness. Globalement, nous pouvons aussi remarquer que les propriétés suivent celles déjà notées dans le cas des paysages ρ MNK.

L'une des meilleures métaheuristiques pour résoudre le problème mono-objectif UBQP se base sur la recherche taboue [GLH10] qui permet, en outre du compromis entre exploration et exploitation de la métaheuristique, de bénéficier d'une évaluation incrémentale rapide pour le problème quadratique UBQP. Pour concevoir une heuristique multiobjective pour le problème mUBQP, l'idée est d'utiliser au sein d'une approche scalaire la recherche taboue existante reconnue efficace où la fonction scalaire est une fonction de somme pondérée. On peut remarquer que la somme pondérée de problème UBQP est encore un problème UBQP mono-objectif où la recherche taboue devrait être efficace. De plus au vu du nombre de solutions supportées, une approche scalaire devrait être efficace pour le problème multiobjectif mUBQP. Nous utilisons donc le principe de la recherche dichotomique [AN79]. Il consiste à rechercher dans un premier temps les meilleures solutions possibles selon chaque objectif f_1 et f_2 . Puis de rechercher une solution «au milieu» de ces deux solutions à l'aide d'une optimisation de la somme pondérée des objectifs dans la direction orthogonale aux deux

3. <http://mocabench.sourceforge.net/>

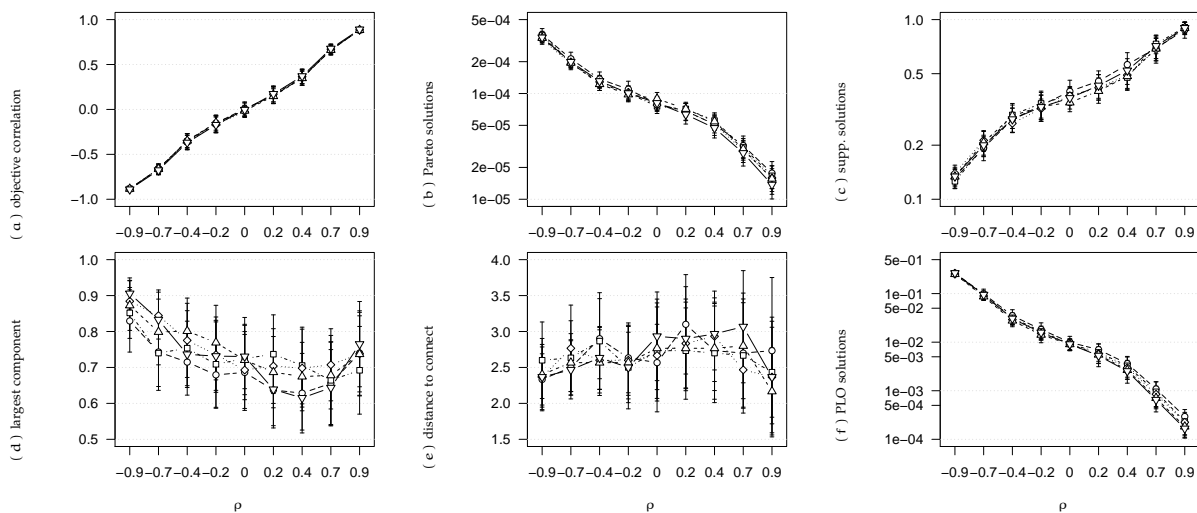


FIGURE 43 – En fonction du paramètre ρ , (a) coefficient de corrélation de Spearman entre les objectifs, (b) ratio du nombre de solutions Pareto optimales dans l'espace de recherche, (c) ratio du nombre de solutions supportées parmi les solutions Pareto optimales, (d) ratio de la taille de la plus grande composante connexe du graphe Pareto parmi les solutions Pareto optimales, (e) distance de Hamming minimale pour obtenir un graphe Pareto connexe, (f) ratio du nombre d'optima locaux Pareto dans l'espace de recherche, pour $d = 0.2$ (\circ), $d = 0.4$ (\square), $d = 0.6$ (\diamond), $d = 0.8$ (\triangle) et $d = 1.0$ (∇). Pour chaque jeu de paramètre, la valeur moyenne et l'intervalle de confiance à 10^{-2} sont reportés, mesurés sur 30 instances indépendantes. La dimension du problème est $n = 18$. À noter l'échelle logarithmique pour l'axe des ordonnées pour les figures (b), (c) et (f).

solutions trouvées précédemment. Ainsi de suite récursivement jusqu'à ne plus trouver de nouvelles solutions entre les solutions précédentes.

L'autre propriété notable du paysage de fitness du problème mUBQP est la faible distance entre les solutions Pareto optimale ce qui suggère qu'une approche Pareto devrait être efficace pour ce problème. L'algorithme TP-LS que nous proposons est un algorithme deux-phases [PS06] qui combine une première phase où une recherche scalaire dichotomique est effectuée, et une seconde phase où une recherche PLS (voir algorithme 1) est exécutée à partir des solutions trouvées dans la première phase par la recherche dichotomique.

Quelques résultats expérimentaux de comparaison entre différents algorithmes sont reportés dans la figure 44. Les instances mUBQP sont biobjectives de grande dimension avec $n \in \{1000, 2000, 3000, 4000, 5000\}$, une même densité $d = 0.8$ et différentes valeurs de corrélation entre objectifs. À titre de comparaison, nous avons choisi une approche scalaire SLS_{unif} qui exécute $\mu = 101$ optimisations scalaires indépendantes avec une recherche taboue. Nous avons naturellement comparé avec les algorithmes qui composent notre proposition, à savoir l'algorithme dichotomique SLS_{dicho} et l'algorithme de recherche Pareto PLS. L'algorithme hybride deux-phases proposées obtient clairement les meilleures performances selon l'indicateur d'hypervolume et d'épsilon quelque soit la valeur des paramètres.

Les approches scalaires par décomposition sont une des approches qui utilise un principe de diviser pour régner avec une coopération entre les sous-problèmes définis. L'analyse précise de la dynamique des algorithmes permet de définir des algorithmes pertinents soit dans un contexte de calcul distribué, soit pour une classe de problème multiobjectif qui modélisent potentiellement de nombreux autres problèmes multiobjectifs rencontrés en recherche opérationnelle.

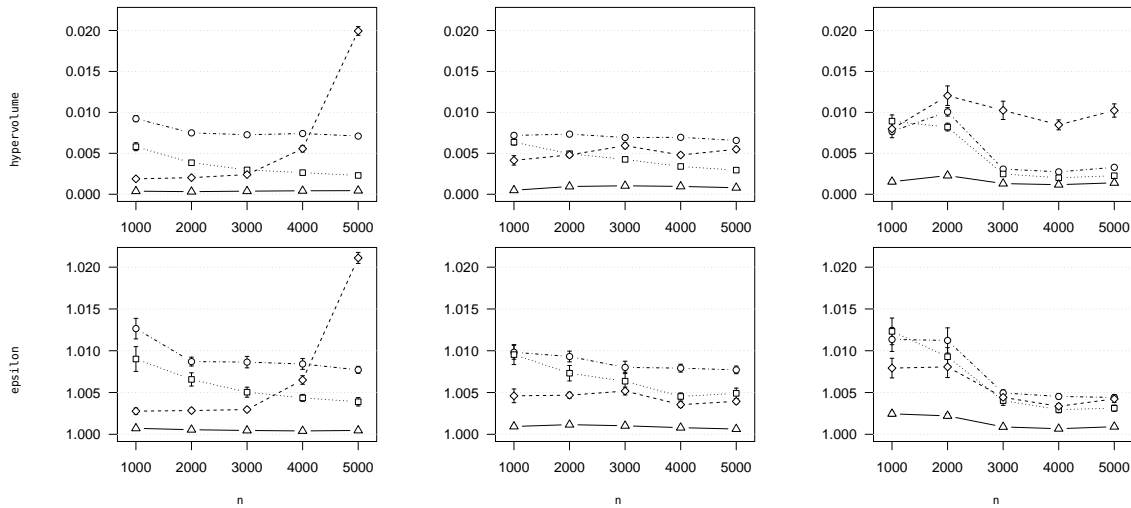


FIGURE 44 – Comparaison des performances de SLS_{unif} (\circ), SLS_{dicho} (\square), PLS (\diamond) et $TP-LS$ (\triangle) pour les indicateurs de qualité d'hypervolume (haut) et epsilon (bas) pour $\rho = -0.5$ (gauche), $\rho = 0.0$ (centre) et $\rho = +0.5$ (droit). Plus la valeur est petite, plus la performance de l'algorithme est meilleure.

3.3 CONCLUSION ET PERSPECTIVES

CONCLUSION. De nombreuses méthodes de résolution des problèmes d'optimisation multiobjectif ont été proposées lors de ces dernières années pour répondre aux nombreux nouveaux défis de l'optimisation multiobjectif. Peu de travaux analysent finement les propriétés de ces algorithmes et encore moins les propriétés des problèmes d'optimisation multiobjectif. À ce jour, il est possible de classer les heuristiques d'optimisation en 3 types d'approches (scalaire, Pareto, indicateur) dont les sources de difficulté potentielle ont été identifiées.

À partir de ce premier travail d'analyse, nous avons pu mettre en place une méthodologie d'analyse de paysage de fitness. Le but est de comprendre le rapport entre les principales caractéristiques d'un problème d'optimisation multiobjectif, à savoir sa dimension, son nombre d'objectifs, sa corrélation entre objectifs et sa non-linéarité, et la structure de l'espace de recherche. Nous avons étendu les paysages NK multiobjectifs en ajoutant la possibilité de régler la corrélation entre objectifs (paysages ρMNK). Sur les petites instances de ce problème générique, nous avons étudié des métriques de paysage de fitness pertinentes par rapport aux types d'approches de résolution. Elles concernent principalement l'ensemble Pareto optimale et les optima locaux Pareto. En parallèle, nous avons proposé de nouvelles métriques qui transposent les concepts de marche adaptative et marche aléatoire au contexte multiobjectif. Cette analyse a permis d'établir un aperçu global des paysages de fitness multiobjectifs et surtout des recommandations de conception pour les différents types de méthode de résolution.

Dans un deuxième temps, nous avons adopté une approche plus orientée donnée où l'on tente d'expliquer les performances d'algorithmes d'optimisation multiobjectif à partir des métriques précédentes. Dans ces travaux, les performances des algorithmes sont des valeurs numériques qui sont modélisées par une régression multilinéaire où les prédicteurs sont les métriques précédentes du paysage de fitness. De cette façon, on peut montrer que les performances sont corrélées aux métriques et interpréter le signe des coefficients de la régression dans la variation des performances. De plus par une technique de sélection des prédicteurs les plus pertinents pour expliquer la performance, nous avons mis en avant les métriques qui influencent le plus les performances. Enfin, afin d'enlever l'effet de groupe induit par les valeurs discrètes des paramètres du benchmark des paysages ρMNK , nous

avons également étudié un modèle mixte qui permet d'expliquer plus finement l'impact des métriques sur les performances indépendamment de la valeur particulière des paramètres du benchmark.

L'étape suivante consiste naturellement à confirmer les résultats obtenus sur des petites instances à de larges instances, ce qui présente le défi supplémentaire de l'échantillonnage de métriques pertinentes. Principalement, nous avons proposé pour cela des métriques relatives au voisinage des solutions que nous avons associées aux techniques d'échantillonnage par des marches aléatoires et des marches adaptatives qui avaient été déjà testées sur les petites instances. Le but de l'analyse est prédictif : à partir de ces nouvelles métriques estimées, nous pouvons prédire avec une bonne qualité les performances de deux types d'algorithmes. De plus, dans un scénario de portfolio d'algorithmes, nous avons aussi montré qu'il était possible de sélectionner correctement avec un faible taux d'erreur l'algorithme le plus performant pour une instance donnée.

Une autre approche d'étude consiste à changer radicalement de définition de paysage de fitness multiobjectif. En s'appuyant sur les travaux qui expriment un problème d'optimisation multiobjectif comme un problème mono-objectif où il s'agit de trouver un ensemble de solutions qui maximise un indicateur de qualité, nous avons proposé une nouvelle définition de paysage de fitness où l'on considère l'espace de recherche des ensembles et une fonction de fitness mono-objective. Bien qu'augmentant grandement la dimension de l'espace de recherche, cette définition permet néanmoins de mener une étude «classique» de paysage de fitness en utilisant les outils classiques existants pour l'optimisation mono-objective. Nous avons ainsi pu montrer les différences entre différents choix de relation de voisinage entre ensembles.

L'une des approches de résolution prometteuse de ces dernières années sont les approches scalaires par décomposition du problème d'optimisation multiobjectif en une famille de problèmes mono-objectifs qui sont résolus de manière coopérative. Nous avons pu montrer que la dynamique de ces algorithmes est guidée par les lignes iso-fitness des fonctions scalaires choisies, indépendamment du type de fonction. Ceci ouvre sur des possibilités nouvelles de contrôle adaptatif des directions de recherche de ces fonctions scalaires pendant la résolution. Ainsi dans un contexte de calcul distribué où plusieurs centaines nœuds de calculs peuvent être disponibles, nous avons pu également mettre au point un algorithme adaptatif qui modifie les directions de recherche en fonction de l'état de la recherche. Cet algorithme est capable de s'adapter à de nombreuses formes de front tout en gardant de bonnes performances lorsque le nombre de nœuds de calcul est important. Enfin, par l'analyse du paysage de fitness d'un nouveau problème d'optimisation multiobjective qui généralise le problème mono-objectif UBQP (Unconstrained Binary Quadratic Problem), nous avons proposé un algorithme hybride qui utilise à la fois les bonnes propriétés des approches scalaires et des approches Pareto.

PERSPECTIVES. Comme dans le cas des paysages de fitness mono-objectifs, l'une des perspectives naturelles est de proposer des métriques encore plus porteuses d'information sur la structure de l'espace de recherche et les dynamiques des algorithmes d'optimisation. Les degrés de liberté pour concevoir de nouvelles métriques sont grands, la diversité des algorithmes en optimisation multiobjective amène aussi à proposer des métriques diverses. Un exemple parmi d'autres, nous avons choisi plusieurs métriques basées sur l'hypervolume pour mesurer la qualité du voisinage des solutions et la rugosité. Nous pourrions imaginer tout simplement de remplacer l'indicateur hypervolume par un autre indicateur comme l'indicateur epsilon.

De même, puisqu'on nous avons mené des directions de recherche similaires entre l'optimisation mono- et multiobjective, une perspective consiste à réduire le coût de l'estimation des métriques pour les problèmes d'optimisation multiobjective. Pour les grands espaces de recherche, les échantillonnages sont basés sur des marches aléatoires ou adaptatives Pareto. Des recherches pourront être menées dans les manières de biaiser ces marches afin d'amé-

liorer la qualité des estimateurs. De même au cas mono-objectif, on pourra imaginer des techniques qui combinent l'estimation de métriques et l'exécution des algorithmes d'optimisation dans un portfolio. Là aussi, les solutions découvertes pendant la phase d'estimation peuvent de nouveau servir à initialiser ou à guider les algorithmes de recherche qui sont utilisés à la suite.

Pour analyser les approches scalaires, nous avons principalement observé les trajectoires dans l'espace objectif des solutions lors des optimisations des fonctions d'agrégation scalaires. À ce jour, peu de travaux expliquent les performances des approches scalaires. Une hypothèse réside dans le fait que ces approches profitent de la diversité des solutions induite naturellement par la décomposition pour trouver une bonne solution à partir d'une solution d'un problème voisin. D'ailleurs, des algorithmes récents proposent des améliorations en maintenant cette diversité. À partir de cette hypothèse, il reste donc à étudier les paysages de fitness conjoints de plusieurs sous-problèmes mono-objectifs, en particulier il faudrait analyser l'évolvabilité des solutions d'un problème mono-objectif à un autre.

Une question ouverte reste de savoir si un problème d'optimisation multiobjective est difficile parce qu'il est difficile d'échapper aux optima locaux, comme on le rencontre souvent en optimisation mono-objective, ou parce qu'il est difficile de converger vers un optimum local. De futurs travaux pourront encore mieux étudier la distribution des optima locaux définis dans le cadre des espaces de recherche d'ensemble. Il serait d'ailleurs possible d'étudier le réseau des optima locaux d'un problème d'optimisation multiobjective. De plus, l'une des difficultés de l'analyse de paysage de fitness multiobjectif basé sur les ensembles est la taille du voisinage qui augmente avec la taille des ensembles tout en réduisant la différence de fitness entre solutions-ensembles voisins. A contrario, une étude des métriques seulement basées sur une seule solution ne tient pas compte de l'interdépendance des solutions dans un ensemble. L'une des perspectives serait donc d'étudier les métriques de paysage de fitness multiobjectif lorsque les ensembles sont réduits à deux ou seulement quelques solutions pour observer l'influence de la covariance de ces solutions sur les métriques.

Enfin, nous menons actuellement en collaboration avec l'université de Shinshu au Japon des travaux en optimisation many-objective [Agu+13a; Agu+13b; Agu+14; Agu+15] qui n'ont pas été exposés dans ce manuscrit par souci de concision et de clarté. L'optimisation many-objectif consiste à optimiser conjointement un grand d'objectifs qui après les succès dans l'industrie de l'optimisation multiobjective avec un faible d'objectif devient un nouveau besoin. Il reste de nombreuses recherches dans le domaine des paysages de fitness à mener pour caractériser ces problèmes dont la grande dimension de l'espace objectif offre de nouveaux défis du point de vue fondamental.

4

PERSPECTIVES GÉNÉRALES

Les conclusions sont données pour chacun des chapitres et ne sont pas reprises à la fin de ce document. Toutefois, au-delà des perspectives propres à chacun des sujets exposés dans ce manuscrit sur le réseau des optima locaux et la structure des problèmes d'optimisation multiobjective, des perspectives plus générales au domaine de recherche peuvent se dégager à la suite de ces travaux. J'ai tenté de les classer en trois catégories.

PAYSAGE DE FITNESS. Une question ouverte depuis de longues années en évolution artificielle est la possibilité d'expliquer la dynamique et les performances des opérateurs de croisement à partir de la notion de paysages de fitness. En effet, l'opérateur de voisinage défini dans les paysages de fitness est un opérateur unaire qui associe un ensemble de solutions voisines à partir d'une solution. Au contraire, l'opérateur de croisement est un opérateur binaire (ou n-aire) qui *a priori* ne peut pas être expliqué directement par un opérateur unaire de mutation. Des travaux ont tenté de proposer des alternatives à la définition des paysages en définissant des P-structure [SW97] ou des croisements particuliers avec certaines solutions déduites de la solution courante [HR96] sans réellement apporter de résultats définitifs. À partir de nos travaux sur les paysages de fitness basés sur les ensembles, nous pourrions proposer une définition des paysages de fitness qui tient compte de toute la population avec une relation de voisinage entre populations obtenue par un opérateur de croisement. De nombreuses études sur la structure des espaces de recherche induite par ces opérateurs de croisement pourraient être conduites à partir d'une telle définition.

Une autre difficulté rencontrée lorsqu'on confronte les études de paysage de fitness avec les nombreux algorithmes utilisés en pratique est l'utilisation d'opérateurs stochastiques, très souvent ergodiques, dont le nombre potentiel de solutions voisines est quasiment tout l'espace de recherche. Avec de tels opérateurs, la notion d'ensemble fini et restreint de solutions voisines perd de sa pertinence et la structure induite par la relation de voisinage du paysage de fitness est perdue. En effet, quasiment toutes les solutions sont alors voisines entre elles. Pour aborder cette problématique, il s'agit de changer la définition de voisinage qui prend en compte la probabilité d'obtenir une solution voisine. Ainsi le paysage de fitness devra être analysé en utilisant les outils d'étude des graphes pondérés. De plus, dans cette perspective il s'agira sûrement de changer la notion d'optimum local en prenant en compte les probabilités induites par l'opérateur stochastique de recherche locale. Ainsi de nouveaux réseaux des optima locaux pourront être mis en évidence et étudiés.

Les travaux que j'ai pu exposer dans ce manuscrit concernent l'optimisation combinatoire. De récents travaux s'intéressent aux techniques d'analyse de paysage de fitness pour l'optimisation continue [Mer+11; ME14b]. Les algorithmes visés sont des algorithmes stochastiques dans un contexte boîte noire (expression de la fonction et de son gradient non disponible). Il paraît alors naturel d'utiliser pour l'optimisation continue les mêmes notions de paysage de fitness définies pour l'optimisation combinatoire. Plusieurs spécificités apparaissent qui peuvent faire l'objet de recherche comme la définition d'un voisinage, la définition même d'un optimum local relatif à un opérateur, la définition d'évolvabilité des solutions, de mesure de régularité de l'espace [LV98; LV06], l'utilisation d'échantillonnage en prenant en compte du «step size», ou encore celui de direction dans l'espace de recherche qui n'existe pas à proprement parler en optimisation combinatoire.

OPTIMISATION ET CONCEPTION À PARTIR DE SIMULATION. L'industrie et le secteur des services ont recours de plus en plus à des modèles et des simulateurs numériques pour

concevoir de nouveaux produits et services. Ils cherchent à tester de nouvelles conceptions afin d'améliorer leur produit ou d'en proposer de nouveaux pour de nouvelles situations. Par exemple, on peut imaginer de mieux régler les feux tricolores de circulation sur une ville entière afin d'améliorer le temps de parcours et les émissions polluantes [AAZ16]. On peut également citer un autre exemple dans le domaine nucléaire que je développe en collaboration avec le CEA dans le cadre de la thèse co-encadrée de Mathieu Muniglia [Mun+16]. L'objectif est de trouver un autre mode de contrôle du cœur d'une centrale nucléaire qui a la possibilité de suivre une demande plus fluctuante afin de pouvoir introduire une plus grande part d'énergie renouvelable dans le bouquet de production en France.

L'un des points communs des problèmes d'optimisation à partir de simulation est le temps important d'évaluation des solutions potentielles : de quelques secondes à quelques heures d'exécution pour de nombreux problèmes. Ce temps est bien plus important que dans les problèmes d'optimisation combinatoire classique comme les problèmes SAT. D'un point de vue de l'optimisation, la conception des algorithmes doit tenir compte du nombre d'évaluations disponibles pour l'algorithme a priori faible. En particulier, ces algorithmes doivent être parallèles ou distribués pour pouvoir s'exécuter sur des plateformes de calcul parallèle. Par ailleurs, tous ces problèmes s'inscrivent pleinement dans l'optimisation boîte noire. En effet, très peu de connaissances sont disponibles sur la définition de ces problèmes. À la fois, la conception distribuée et le peu de connaissances du problème posent des défis sur le plan de la conception. Par exemple, la définition d'un opérateur de voisinage ou du réglage de ces paramètres associés est souvent délicate. La conception d'un algorithme et de son paramètre s'effectue souvent alors par essai-erreur. Les outils de paysage de fitness doivent pouvoir apporter des solutions dans la conception raisonnée d'un algorithme distribué adapté à chaque type de simulation.

Afin de réduire le nombre d'évaluation et donc de simulation utilisé par l'algorithme, une voie est d'utiliser un méta modèle (surrogate model) qui se substitue à la fonction d'évaluation au cours de l'exécution de l'algorithme. Au cours de l'exécution de l'algorithme, un méta modèle est appris par l'algorithme et aussi utilisé en remplacement de la fonction d'évaluation qui peut être mono- ou multiobjective. Jusqu'alors, peu de travaux existent en optimisation combinatoire et multi-objective où les méta modèles à variables discrètes sont moins courants et pourtant ces problèmes existent comme les deux exemples cités plus haut.

CONCEPTION AUTOMATIQUE D'ALGORITHME D'OPTIMISATION. Le réglage des paramètres, au sens large un paramètre est un composant d'un algorithme ou la valeur d'un paramètre numérique (discret ou continu), est un sujet de recherche existant depuis le début des algorithmes évolutionnaires [Gre86] et toujours très actuels [Bel+16]. Classiquement, on distingue le réglage «off-line» avant l'exécution de l'algorithme, appelé tuning, du réglage «on-line» pendant l'exécution, appelé contrôle (voir la section 1.2.2 pour rappel des principales techniques). Dans le cas des algorithmes distribués en environnement de calcul parallèle de plus en plus utilisés en pratique [Lou+09], des paramètres supplémentaires sont à régler comme la distribution des algorithmes sur la plateforme, le choix des algorithmes, la communication entre eux (migration pour les modèles en îles), etc. Les méthodes de tuning off-line paraissent peu adaptées à ce contexte du fait du grand nombre de tests requis par ces méthodes. Actuellement dans le cadre de la thèse de Christopher Jankee [Jan+15b; Jan+15a; Jan+16], nous étudions les méthodes de contrôle on-line. L'approche paysage de fitness peut apporter à deux niveaux. Tout d'abord pendant l'optimisation, des mesures locales du paysage peuvent fournir des indications sur la dynamique de la recherche au-delà de la simple mesure de performances des solutions qui composent la population de recherche. De plus *a posteriori* après l'optimisation, le recueil des mesures locales pendant la recherche peut aussi fournir des indications sur le paysage de fitness de manière à être utilisé par la suite. La combinaison des approches on-line et off-line semble être une possibilité pour améliorer les algorithmes, en particulier pour les problèmes d'optimisation à évaluation coûteuse.

Comme dans d'autres domaines de la science, l'arrivée de masse de données peut apporter de nouvelles perspectives. L'un des freins à la fois à la conception d'algorithmes d'optimisation mono- et multi-objective plus efficace, mais également à la meilleure compréhension des structures des problèmes d'optimisation, est le faible nombre d'instances ou de benchmarks étudié de façon homogène. La connaissance est souvent parcellaire à un type de problème voir à un sous-type d'instance de problème. Par exemple, il existe de nombreuses variantes de problèmes de tournées de véhicules ou de problèmes de logistique [Mou+11; MNY12] qui sont peu étudiés de manière homogène du point des métaheuristiques. La constitution d'un corpus de problèmes d'optimisation mono- et multi-objective qui contient à la fois les performances de métaheuristiques et les métriques de paysage de fitness serait une vraie opportunité dans ce domaine de recherche. Se posent de nombreuses questions scientifiques lors de la constitution la plus effective de ce corpus : quelles instances de problèmes, quels algorithmes, quelles métriques ; et aussi lors de l'utilisation de ce corpus : peut-on utiliser les métriques d'un problème ou les performances d'un algorithme pour prédire les performances d'un algorithme différent sur un problème différent, comment concevoir des méthodes de régression à partir d'un grand nombre de métriques, etc.



NOTICE BIOGRAPHIQUE

A.1 INFORMATIONS PERSONNELLES

Dr. Verel Sébastien

Date de naissance	29 octobre 1976 (39 ans)
Unité de recherche	Laboratoire d'Informatique Signal et Image de la Côte d'Opale (LISIC) - EA 4491, Université du Littoral Côte d'Opale (ULCO)
Unité d'enseignement	Département informatique Université du Littoral Côte d'Opale (ULCO)
Adresse	50, rue Ferdinand Buisson BP 719 62228 Calais Cedex France
Page web	http://www-lisic.univ-littoral.fr/~verel/

A.2 SITUATION PROFESSIONNELLE

Dates début	Dates fin	Établissement	Fonctions et statuts
09/2013	actuel	Université du Littoral Côte d'Opale Laboratoire LISIC, EA 4491	Maître de Conférences
09/2006	08/2013	Université Nice Sophia Antipolis Laboratoire I3S, UMR 7271	Maître de Conférences
10/2005	08/2006	Université Nice Sophia Antipolis (UNS)	ATER 27 mi-temps
10/2002	09/2005	Ecole ing. (ENSISA) et dép. inf. UNS	Moniteur informatique
01/2002	06/2002	Laboratoire I3S UMR 7271	Stage de DEA
09/2000	10/2001	Collège et Lycée proche de Caen	Ens. de mathématiques (CAPES)

- Délégation : Sept. 2009 - Août 2011, à l'INRIA Lille Nord Europe, Equipe DOLPHIN, Sujet : "Structure des espaces de recherche de l'optimisation multiobjective"
- Mobilité : Sept. 2013, Mutation, de l'Université de Nice Sophia Antipolis vers l'Université du Littoral Côte d'Opale.

A.3 CURSUS UNIVERSITAIRE

- 2002 - 2005 : Doctorant au laboratoire I3S (UMR 7271), bourse MESR
Université de Nice-Sophia Antipolis
- 2001 - 2002 : DEA en informatique
Université de Nice-Sophia Antipolis
- 2000 - 2001 : Titulaire du Capes de Mathématiques
Institut Universitaire de Formation des Maîtres de Caen
- 1998 - 1999 : Maîtrise de mathématiques pures
Université de Caen
- 1997 - 1998 : Licence de mathématiques
Université de Caen
- 1994 - 1997 : Classes préparatoires aux grandes écoles d'ingénieurs
Lycée Malherbe de Caen (admis à l'ISMRA)

A.4 ACTIVITÉS SCIENTIFIQUES

Mes travaux s'inscrivent dans les domaines de l'intelligence artificielle et de la recherche opérationnelle, plus précisément dans celui de l'**optimisation combinatoire stochastique**. J'analyse et je développe des méthodes d'optimisation, regroupées selon le terme **métaheuristique**, telles que les algorithmes évolutionnaires, la recherche taboue, le recuit simulé, etc. Depuis ma thèse de doctorat, mes travaux sont principalement fondamentaux, ils ont pour but d'analyser la **structure des espaces de recherche** (paysage de fitness) en utilisant les outils des systèmes complexes afin d'améliorer les algorithmes d'optimisation en fonction du problème à résoudre. Depuis ma délégation inria, une grande part de mes travaux concerne les problèmes d'**optimisation multiobjective**. Par ailleurs, notamment dans le cadre de la thèse en collaboration avec le CEA, je développe des algorithmes distribués qui ont pour but d'optimiser la conception dont la qualité est basée sur des critères calculés à partir d'une **simulation numérique**.

A.5 ENCADREMENT DOCTORAL ET SCIENTIFIQUE

A.5.1 Co-encadrement doctoral en cours

- oct. 2014 - sept. 2017 : Christopher Jankee, "Optimisation par métaheuristique adaptative distribuée en environnement de calcul parallèle", co-encadrement (50%) avec Cyril Fonlupt. Financement : MESR - Université du Littoral Côte d'Opale (ULCO).
- oct. 2014 - sept. 2017 : Mathieu Muniglia, "Méthodologie d'optimisation des réacteurs de la filière REP dans le cadre de l'introduction des EnR (Energie Renouvelable) dans le mix énergétique français", co-encadrement (60%) avec Sylvain David. Financement : IDEX Paris Saclay (70%) / CEA,PS2E (30%).

A.5.2 Co-encadrement doctoral

- 2009 (4,5 années), Fabio Daolio (Université de Lausanne), co-encadrement 33% avec Marco Tomassini (dir.) et Gabriela Ochoa (Stirling univ.), Sujet : "Local Optima Networks of Hard Combinatorial Fitness Landscapes", Financement : Fond National Recherche Suisse. *post-doc Univ. Shinshu (Japan), puis à Stirling (UK)*

- Septembre 2006 - octobre 2009, David Simoncini (Université Nice Sophia Antipolis), co-direction 33% avec Philippe Collard (dir.) et Manuel Clergue, Sujet : "Sélection Topologique dans les Algorithmes Evolutionnaires Cellulaires : Etude du Compromis Exploration/Exploitation", *post-doc Riken (Japan), puis à l'INRA Toulouse*
- Septembre 2007 - décembre 2010, Maroun Bercachi (Université Nice Sophia Antipolis), co-direction 33% avec Philippe Collard (dir.) et Manuel Clergue, Sujet : "Algorithme évolutionnaire à états pour l'optimisation difficile".

A.5.3 Co-encadrement Master 2

- fév. 2014 (6 mois), Christopher Jankee, 50% avec M. Carrard (ULCO), "Etude à l'aide d'une simulation informatique multi-agents visant à saisir les conséquences de la réforme des aéroports pour les territoires". *Poursuite en thèse 2014 - 2017 à l'ULCO*
- janv. 2012 (6 mois), Kevin Lefèvre, 30% avec T. Baccino (Paris 8) "Extraction de processus et modélisation de l'activité de lecture silencieuse".
- mars 2008 (6 mois), Thé Van Luong, 50% avec M. Clergue (UNS) "Optimization by distributed metaheuristics". *Thèse soutenue en 2011, équipe DOLPHIN, sur la thématique de l'optimisation haute performance parallèle.*
- avril 2008 (5 mois), Andrea Scarcella, 75% avec L. Vanneshi (univ. Milano-bicocca), "Models of State-based Evolutionary Algorithm",
- mars 2007 (6 mois) Mirela Frandes, 50% avec M. Clergue (UNS) "Finding by metaheuristic of a rule of synchronisation for Firing Squad Problem", *Thèse soutenue en 2010 à l'INSA de Lyon.*
- mars 2007 (6 mois), Fabien Valoit, 33% avec M.-J. Rendas (i3s-CNRS), "Synchronization of signal of a video camera and a scan sonar",
- mars 2006 (6 mois), Sébastien Tesquet, 50% avec P. Collard (UNS), "Study and implementation of States based Evolutionary Algorithm".

A.5.4 Encadrement stage doctoral

- janv. 2015 (3 mois), Noel Rodriguez Maya (Universidad Michoacana, Mexique), 100%, "Performance Prediction of Evolutionary Algorithms on High Dimensional Optimization Problems",
- nov. 2013 (1 an), Martin Drozdik (Shinshu University, Japon), 50% avec A. Liefooghe (Lille 1), "Differential evolution for multiobjective optimization".

A.5.5 Co-encadrement Post-docdoral

- oct. 2014 (1 an), Fabio Daolio, 50% avec A. Liefooghe (Lille 1), "Fitness landscape analysis for multiobjective optimization".

A.6 DIFFUSION SCIENTIFIQUE

Tutoriels

- Janv. 2015, conférence LION'9, Lille, France, "Fitness landscape : the metaphor and beyond"
- Juil. 2013, conférence GECCO'13, Amsterdam, Pays-Bas, "Fitness Landscapes and graphs : Multimodularity, ruggedness and Neutrality".
- Juil. 2012, conférence GECCO'12, Philadelphia, USA,
- Juil. 2010, conférence WCCI 2010, Barcelona, Espagne,
- Juil. 2009, conférence GECCO'09, Montreal, Canada,
- Juil. 2007, conférence GECCO'07, Londres, UK, "Fitness landscapes and problem hardness in evolutionary computation".

Conférencier invité

- juin 2015, Séminaire du TVES, Dunkerque, France, "Modélisation et simulation informatique en sciences humaines et sociales".
- avr. 2015, Séminaire MC3, Sophia-Antipolis, France, "Firing Squad Synchronization Problem : Optimization methods to synchronize as much as possible".
- fév. 2014, Service de Etudes de Réacteurs et de Math. App. (SERMA), CEA, Saclay, "Multiobjective optimization with evolutionary algorithms and metaheuristics in distributed computing environment".
- sept. 2013, Lorentz Center, International Center for Workshops in the Sciences, Set-Oriented and Indicator-Based Multi-Criteria Optimization (SIMCO).
- fév. 2013, Laboratoire d'Informatique Signal et Image de la Côte d'Opale, Calais, France, "Design of optimization algorithms : tuning and control of parameters".
- janv. 2012, LAMIA, Université des Antilles et de la Guyane (UAG), France, "Problèmes combinatoires multi-objectif et paysages de fitness".
- sept. 2010, Dagstuhl seminar on Theory of Evolutionary Algorithms, Warden, Deutschland, "Population-based Fitness Landscapes".
- mars 2010, 4th workshop on Theory of Randomized Search Heuristics, Paris, "Set-based Multiobjective Fitness Landscapes : definition and properties".
- mars 2010, Université d'Angers, Angers, Fr, "Fitness Landscapes and Local Optima Network".
- fev 9, 2010, ASAP group University of Nottingham, Nottingham, UK, "Fitness Landscapes and graphs in combinatorial optimization".
- juil 2007, FRAC, Nice, "Towards a resolution of the firing squad problem with 5 states by metaheuristics".

Développement logiciel

- ParadisEO (<http://paradiseo.gforge.inria.fr>) : Plate-forme logiciel open-source écrite en c++ permettant de résoudre des problèmes d'optimisation mono- ou multi-objectifs à l'aide d'algorithmes d'optimisation stochastiques séquentiels ou parallèles tels que les algorithmes évolutionnaires ou les metaheuristiques.
 - Co-administrateur et co-développeur du projet Paradiseo
 - Coordinateur et principal développeur de la partie recherche locale (paradiseo-MO) qui comprend conception, implémentation, tests, documentation, tutoriel. Actuellement 148 classes et 15453 lignes de code

- MoCObench (mocobench.sourceforge.com) :

Répertoire de problèmes, de générateurs et d'instances de problèmes d'optimisation combinatoire multiobjective.

- Fonctions : co-administrateur et co-développeur du projet depuis la création en 2012. 2 types de problème d'optimisation écrits en 4 langages comprenant 615 instances.

A.7 RAPPORTEUR DE THÈSE ET INSTANCE D'ÉVALUATION

- Rapporteur de thèse de doctorat :

- juil. 2016, Noel Enrique Rodriguez Maya, Universidad Michoacana de San Nicolas de Hidalgo, Mexique,

Titre de la thèse : "Characterizing the performances of evolutionary algorithms to solve continuous optimization problems",

dir. J. J. Flores Romero et M. Graff Guerrero,

- fév. 2014, Katherine Mary Malan, university of Pretoria, Afrique du Sud,

Titre de la thèse : "Characterising Continuous Optimisation Problems for Particle Swarm Optimisation Performance Prediction",

dir. A. P. Engelbrecht

- Membre de jury de thèse de doctorat :

- Noel Enrique Rodriguez Maya, "Characterizing the performances of evolutionary algorithms to solve continuous optimization problems", Universidad Michoacana de San Nicolas de Hidalgo (Mexico), dir. Juan José Flores Romero and Mario Graff Guerrero, July 15th, 2016.

- Fazia Aiboud, "Méthodes approchées pour la résolution de problèmes inverses : identification paramétrique et génération de formes", Université Blaise Pascal - Clermont-Ferrand II, dir. S. Norre and N. Grangeon, 29 novembre 2013.

- Marie-Eléonore Marmion, "Local search and combinatorial optimization : from structural analysis of a problem to design efficient algorithms", Université Lille 1 (France), dir. C. Dhaenens and L. Jourdan, décembre 2012.

- Maroun Bercacchi, "Algorithme évolutionnaire à états pour l'optimisation difficile" université de Nice-Sophia Antipolis (France), dir. P. Collard, décembre 2010.

- David Simoncini, "Sélection Topologique dans les Algorithmes Evolutionnaires Cellulaires : Etude du Compromis Exploration/Exploitation", université de Nice-Sophia Antipolis (France), dir. P. Collard, octobre 2009.

- Evaluations de projet :

- juil. 2013, expertise pour le Research Foundation - Flanders, Belgique.

- juil. 2012, expertise pour le Portuguese Foundation for Science and Technology (FCT).

- avr. 2012, expertise pour l'ANR programme MN.

- juin 2011, expertise pour le programme eurotalent-CEA.

- Editeur associé pour une issue spéciale de journal international :

- 2014, European Journal of Operational Research (EJOR),

"evolutionary multiobjective optimization", volume 243, issue 2

co-éditée avec Dimo Brockhoff, Bilel Derbel, Arnaud Liefooghe (INRIA/Lille 1).

- Activité de review de journaux internationaux :

Reviewer régulier de :

- IEEE transaction on Evolutionary Computation (IEEE TEC)

- Journal of Evolutionary Computation (ECJ)

Reviewer ponctuel de :

- Journal of Heuristic (JoH)

- Journal of Theoretical Biology (JTB)

- Natural Computing (NC)

- Journal of Computer Science and Technology (JCST)
- Journal of Mathematical Modeling and Algorithms (JMMA)

A.8 ADMINISTRATION ET ANIMATION DE LA SCIENCE

Projets scientifiques

- **Projet international :**
 - 2013 - 2016 (2.5 années) : projet bilatéral Japon-France, Financé par le JSPS, programme MEXT Japan Society for the Promotion of Science, "Global Research on the Framework of Evolutionary Solution Search to Accelerate Innovation" Participants : Univ. Shinshu (Japon), Univ. Tokyo (Japon), Tao (inria Saclay), Osmose (ULCO), DOLPHIN (Lille 1, inria Lille). Coordinateur : Kiyoshi Tanaka (Univ. Shinshu) Responsable de la tâche : "Landscape Analysis and Search Performance", Participant de la tâche : "evolutionary multi- and many-objective optimization", Financement en France de 3 années de postdoc et de 1 an de visite doctorale.
- **Projet GDR :**
 - 2011 : Projet bilatéral GDR RO (recherche opérationnelle) de l'INS2I (CNRS), "Set-based search for multiobjective combinatorial optimization", A. Goëffon, M. Basseur (Angers), A. Liefooghe, C. Dhaenens (Lille 1), S. Verel (UNS)
- **Participation projet ANR :**
 - 2009 - 2012 : Programme blanc. Coord. A. Guérin-Dugué (Univ. Grenoble), "GAZE and EEG : Traitements conjoints synchronisés de signaux EEG et de suivi de mouvements oculaires pour l'analyse spatiotemporelle et modélisation d'activités neuronales." Partenaires : Paris 8, TIMC-IMAG et LPNC (univ. de Grenoble).
- **Projet à financement local :**
 - 2016 : BQI de l'Université du Littoral Côte d'Opale, "ULCO-SHINSHU : développement des relations en recherche et formation avec l'université de Shinshu au Japon".
 - 2014 : BQR de l'Université du Littoral Côte d'Opale, "Combinatorial Optimization in distributed environnement".
 - 2012 - 2014 : Action de développement technologique (ADT) financé par inria, "VRP-solve : Metaheuristic-based solver for vehicle routing" Responsabilité : Réglage automatique des paramètres des algorithmes d'optimisation

Administration de la science

- **Responsabilités au sein de conférences internationales :**
 - oct. 2015 : trésorier 12th International Conference on Artificial Evolution, Lyon.
 - oct. 2013 : trésorier 11th International Conference on Artificial Evolution, Bordeaux.
 - oct. 2011 : trésorier 10th International Conference on Artificial Evolution, Angers.
 - oct. 2009 : trésorier 9th International Conference on Artificial Evolution, Strasbourg.
- **Administration et animation de structures associatives scientifiques :**
 - 2015 - actuel : Vice-président de l'association Evolution Artificielle.
 - 2008 - 2015 : Trésorier et membre du bureau de l'association Evolution Artificielle. Association scientifique qui regroupe les chercheurs en évolution artificielle qui organise séminaires, écoles d'été et conférences internationales.
- **Responsabilité au sein du laboratoire :**
 - 2014 - 2019 : Membre élu conseil du laboratoire du LISIC
 - 2014 - : Responsable des séminaires du LISIC

— 2011 - 2013 : Membre nommé conseil du laboratoire I3S

Animation de la science

- Organisation écoles d'été :

- juin 2013, 8th Artificial Evolution Summer School, Quiberon, France, co-organisée avec A. Liefoghe, B. Derbel et D. Brockhoff (Lille 1/INRIA).
- juin 2012, 7^{ème} école d'été "artificial evolution", Quiberon, France, co-organisée avec A. Liefoghe et L. Jourdan.
- juin 2009, 4^{ème} école d'été "artificial evolution", Porquerolles, France, co-organisée avec M. Clergue.
- juin 2008, 3^{ème} école d'été "artificial evolution", Porquerolles, France, co-organisée avec M. Clergue.

- Sessions spéciales de conférences internationales :

- jui. 2015, international conference CEC 2015 "Fitness landscape Analysis and Search Space Structure" Co-organizers : Hernan Aguirre (Shinshu University, Japan), Kiyoshi Tanaka (Shinshu University, Japan), Arnaud Liefoghe (univ. Lille 1 / INRIA, France).
- Special session at the international conference LION 9 : "Multiobjective Combinatorial Optimization (LION MoCO)" Co-organizers : Hernan Aguirre (Shinshu University, Japan), Kiyoshi Tanaka (Shinshu University, Japan), Arnaud Liefoghe (univ. Lille 1 / INRIA, France).
- 2013, International Conference on Multiple Criteria (MCDM'2013) "Evolutionary Multiobjective Optimization" co-organisée avec D. Brockhoff (INRIA), B. Derbel, A. Liefoghe (Lille 1).
- jan. 2013, international conference LION 7, "Problem Structure vs. Algorithm Performance in Multiobjective Comb. Optimization" co-organisée avec H. Aguirre, K. Tanaka (Shinshu University, Japan), A. Liefoghe.
- jan. 2012, international conference LION 6, Autonomous Control for Search Algorithms (LION-S*EA) co-organisée avec F. Lardeux et F. Saubion (Univ. Angers).

- Vulgarisation

Livre sur les systèmes complexes s'adressant à un large public curieux bien au delà des passionnés d'informatique. Ce livre vise à montrer les principes fondamentaux des systèmes complexes par la pratique à travers des exemples simples :

Philippe Collard, Sébastien Verel, and Manuel Clergue. *Systèmes Complexes : Une Introduction Par La Pratique*. Presses Polytechniques et Universitaires Romandes, Publisher of EPFL Press, 306 pages, 2013.

A.9 ACTIVITÉS D'ENSEIGNEMENT

Résumé des enseignements depuis la thèse

h TD	Fonction	Formation	Intitulé de l'enseignement
2015 - 2016 : Département informatique, Université du Littoral Côte d'Opale, mcf			
56 h	CM+TP	Master 1 I2L	Intelligence Artificielle et Systèmes Formels
39 h	CM+TP	Master 1 I2L	Recherche opérationnelle, optimisation
45 h	CM+TP	Master 1 ISiDIS	Recherche opérationnelle, optimisation
32,5 h	CM+TP	Master 2 ISiDIS	Data science
36 h	CM+TP	Master 2 I2L	Programmation fonctionnelle
5 h		Master 2 ISiDIS	Projet de master 2
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization
2014 - 2015 : Département informatique, Université du Littoral Côte d'Opale, mcf			
56 h	CM+TP	Master 1 I2L	Intelligence Artificielle et Systèmes Formels
39 h	CM+TP	Master 1 I2L	Recherche opérationnelle, optimisation
45 h	CM+TP	Master 1 ISiDIS	Recherche opérationnelle, optimisation
18 h	CM+TP	Master 1 ISiDIS	Systèmes Formels
36 h	CM+TP	Master 2 I2L	Programmation fonctionnelle
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization
2013 - 2014 : Département informatique, Université du Littoral Côte d'Opale, mcf			
56 h	CM+TP	Master 1 I2L	Intelligence Artificielle et Systèmes Formels
39 h	CM+TP	Master 1 I2L	Recherche opérationnelle, optimisation
45 h	CM+TP	Master 1 ISiDIS	Recherche opérationnelle, optimisation
36 h	CM+TP	Master 2 I2L	Programmation fonctionnelle
7,5 h	CM+TD	Master 1 informatique (UNS)	Systèmes Artificiels Complexes
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization
2012 - 2013 : U.F.R Sciences, Université de Nice Sophia-Antipolis, mcf			
70 h	CM+TP	Licence 1 MASS	Introduction à la programmation java
37 h	CM+TD	Licence 2 informatique	Projet Scientifique Informatique
17 h	CM+TD	Licence 3 math	Projet Scientifique Informatique
35 h	CM+TD	Licence 3 informatique	Introduction aux systèmes complexes
18 h	TP	Licence 3 informatique	Réseaux
12,5 h	CM+TD	Master 1 informatique	Systèmes Artificiels Complexes
6h	TD	Master 1 informatique	Travail d'étude et de recherche
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization
2011 - 2012 : U.F.R Sciences, Université de Nice Sophia-Antipolis, mcf			
15 h	CM+TD	Master 2 informatique, parc. complex systems	Natural Computing
72 h	TP	Licence 1 informatique	Systèmes informatique
70 h	CM+TP	Licence 1 MASS	Introduction à la programmation java
30 h	CM+TD	Licence 3 informatique	Introduction aux systèmes complexes
20 h	TD	Licence 3 MIAGE	Harmonisation web
15 h	CM+TD	Master 1 informatique	Systèmes Artificiels Complexes
4,5h	TD	Master 1 informatique	Travail d'étude et de recherche
9h	TD	Licence 3 MIAGE	Stage en entreprise
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization
7,5h	CM	Ecole Navale de Brest	Théorie des Jeux et Réseaux Sociaux
2009 - 2011 : en délégation			
8 h	CM+TD	Master 2 informatique, parc. complex systems	Natural Computing
7,5 h	CM	Master 2 psychology of cognitive systems (Paris 8 et EPHE)	Measure and Modelization

h TD	Fonction	Formation	Intitulé de l'enseignement
2008 - 2009 : U.F.R Sciences, Université de Nice Sophia-Antipolis, mcf			
15 h	CM+TD	Licence 3 informatique	Introduction aux systèmes complexes
4,5 h	CM	Master 2 informatique	Informatique bio-inspirée
13 h	CM+TP	Licence 2 informatique	Programmation logique
51 h	CM+TP	Licence 3 informatique	Programmation web
40 h	CM+TD	Licence 2 MASS	Création site web
9 h	CM	Master 1 informatique	Systèmes Artificiels Complexes
44 h	CM+TP	Licence 1 MASS	Introduction à la programmation java
26 h	TD	Licence 1 MASS	Algorithmique
2007 - 2008 : U.F.R Sciences, Université de Nice Sophia-Antipolis, mcf			
45 h	CM+TD	Licence 1 MASS	Algorithmique et Calcul Formel
16 h	TP	Licence 2 maths-info	Programmation Fonctionnelle
68 h	CM+TP	Licence 2 MASS	Algorithmique et Informatique Théorique
32 h	TP	Licence 3 informatique	Programmation Orientée Objet
12 h	TD	Licence 3 MIAGE	Suivi de stage
10 h	CM	Master 1 informatique	Systèmes Artificiels Complexes
8 h	CM	Master 2 informatique	Programmation évolutionnaire
2006 - 2007 : U.F.R Sciences, Université de Nice Sophia-Antipolis, mcf			
45 h	CM+TD	Licence 1 MASS	Algorithmique et Calcul Formel
19 h	TP	Licence 2 maths-informatique	Programmation Fonctionnelle
68 h	CM+TP	Licence 2 MASS	Algorithmique et Informatique Théorique
16 h	TP	Licence 3 informatique	Programmation Orientée Objet
18 h	TD	Licence 3 informatique	Systèmes et réseaux
10 h	CM	Master 1 informatique	Systèmes Artificiels Complexes
8 h	CM	Master 2 informatique	Programmation évolutionnaire
2005 - 2006 : U.F.R Sciences, Université de Nice Sophia-Antipolis, ater			
32 h	TP	Licence 1 MASS	Algorithmique et Calcul Formel
32 h	TD	Licence 2 MASS	Algorithmique et Informatique Théorique
4 h	CM	Master 1 STIC	Systèmes Artificiels Complexes
8 h	CM	Master 2 STIC	Programmation Évolutionnaire
24 h	TP	Licence 1 Math-info	Introduction à la Programmation (Java)

Responsabilités pédagogiques

- sept. 2013 - actuel : Directeur des études du Master 2 informatique parcours ISiDIS, Département informatique, université du Littoral Côte d'Opale :
 - Suivi pédagogique des étudiants : Etude des dossiers d'inscription, suivi des étudiants, conseils d'orientation, jury, informations sur la licence, etc.
 - Tâches de coordination : emploi du temps, absences, etc.
- sept. 2012 - août 2013 : Responsable Licence 3 informatique, Département informatique, université de Nice Sophia Antipolis :
 - Suivi pédagogique des étudiants : Etude des dossiers d'inscription, suivi des étudiants, conseils d'orientation, jury, informations sur la licence, etc.
 - Tâches de coordination : emploi du temps, absences, réservation de salles, etc.
 - Mise en place de conventions de recrutement avec d'autres départements d'enseignement
- sept. 2011 - août 2013 : Responsable Validation Acquis d'Expérience (VAE), Département informatique, université de Nice Sophia Antipolis,
- 2012 - 2013 : Participation jury de recrutement, Licence 3 informatique parcours MIAGE

B

BIBLIOGRAPHIE PERSONNELLE CLASSÉE DEPUIS LA THÈSE

B.1 LIVRE

- [CVC13] Philippe COLLARD, Sébastien VEREL et Manuel CLERGUE. *Systèmes complexes : Une introduction par la pratique*. Presses polytechniques et universitaires romandes, Publisher of the EPFL Press, 2013.

B.2 CHAPITRE DE LIVRE

- [Och+14] Gabriela OCHOA, Sébastien VEREL, Fabio DAOLIO et Marco TOMASSINI. “Local Optima Networks: A New Model of Combinatorial Fitness Landscapes”. In : *Recent Advances in the Theory and Application of Fitness Landscapes*. Sous la dir. d’Andries Engelbrecht HENDRIK RICHTER. Emergence, Complexity and Computation. Springer Berlin Heidelberg, 2014, p. 233–262.
- [Ber+08] Maroun BERCACHI, Philippe COLLARD, Manuel CLERGUE et Sébastien VEREL. “Studying the Effects of Dual Coding on the Adaptation of Representation for Linkage in Evolutionary Algorithms”. In : *Linkage in Evolutionary Computation*. Sous la dir. d’Ying-ping CHEN et Meng-Hiot LIM. Studies in Computational Intelligence. Springer Berlin / Heidelberg, 2008, p. 249–284.

B.3 ARTICLES DE JOURNAUX INTERNATIONAUX AVEC COMITÉ DE LECTURE

10 articles publiés depuis 2006 dans des journaux internationaux avec comité de lecture.

- [Dao+16] Fabio DAOLIO, Arnaud LIEFOOGHE, Sébastien VEREL, Hernán AGUIRRE et Kiyoshi TANAKA. “Problem Features vs. Algorithm Performance on Rugged Multi-objective Combinatorial Fitness Landscapes”. In : *Evolutionary Computation* (2016), to appear.
- [Der+14c] Bilel DERBEL, Jérémie HUMEAU, Arnaud LIEFOOGHE et Sébastien VEREL. “Distributed Localized Bi-objective Search”. In : *European Journal of Operational Research* 239 (2014), p. 731–743.
- [LVH14] Arnaud LIEFOOGHE, Sébastien VEREL et Jin-Kao HAO. “A hybrid metaheuristic for multiobjective unconstrained binary quadratic programming”. In : *Applied Soft Computing* 16 (2014), p. 10–19.
- [Hum+13a] Jérémie HUMEAU, Arnaud LIEFOOGHE, El-Ghazali TALBI et Sébastien VEREL. “ParadisEO-MO: From Fitness Landscape Analysis to Efficient Local Search Algorithms”. In : *Journal of Heuristics* 19.6 (2013), p. 881–915.
- [Ver+13] Sébastien VEREL, Arnaud LIEFOOGHE, Laetitia JOURDAN et Clarisse DHAENENS. “On the structure of multiobjective combinatorial search space: MNK-landscapes with correlated objectives”. In : *European Journal of Operational Research* 227.2 (2013), p. 331–342.

- [Van+12] Leonardo VANNESCHI, Yuri PIROLA, Giancarlo MAURI, Philippe COLLARD et Sébastien VEREL. “A Study of Neutrality of Boolean Function Landscapes in Genetic Programming”. In : *Journal of Theoretical Computer Science* 425 (2012), p. 34–57.
- [Dao+11] Fabio DAOLIO, Marco TOMASSINI, Sébastien VEREL et Gabriela OCHOA. “Communities of Minima in Local Optima Networks of Combinatorial Spaces”. In : *Physica A: Statistical Mechanics and its Applications* 390.9 (2011), p. 1684–1694.
- [VOT10] Sébastien VEREL, Gabriela OCHOA et Marco TOMASSINI. “Local Optima Networks of NK Landscapes with Neutrality”. In : *IEEE Transactions on Evolutionary Computation* volume 14.6 (2010), p. 783–797.
- [TVO08] Marco TOMASSINI, Sébastien VEREL et Gabriela OCHOA. “Complex-network analysis of combinatorial spaces: The NK landscape case”. In : *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* 78.6 (2008), p. 066114.
- [Ver+07] Sébastien VEREL, Philippe COLLARD, Marco TOMASSINI et Leonardo VANNESCHI. “Fitness landscape of the cellular automata majority problem: View from the Olympus”. In : *Theoretical Computer Science* 378.1 (2007), p. 54–77.

B.4 ARTICLES DE CONFÉRENCES INTERNATIONALES AVEC COMITÉ DE LECTURE

57 articles publiés depuis 2006 dans des conférences internationales avec comité.

- [Jan+16] Christopher JANKEE, Sébastien VEREL, Bilel DERBEL et Cyril FONLUPT. “A Fitness Cloud Model for Adaptive Metaheuristic Selection Methods”. In : *Parallel Problem Solving from Nature – PPSN XIV: 14th International Conference, Edinburgh, UK, September 17–21, 2016, Proceedings*. Sous la dir. de Julia HANDL, Emma HART, R. Peter LEWIS, Manuel LÓPEZ-IBÁÑEZ, Gabriela OCHOA et Ben PAECHTER. Cham : Springer International Publishing, 2016, p. 80–90.
- [Mun+16] Mathieu MUNIGLIA, Jean-Michel DO, Le Pallec JEAN-CHARLES, Hubert GRARD, Sébastien VEREL et S. DAVID. “A Multi-Physics PWR Model for the Load Following”. In : *International Congress on Advances in Nuclear Power Plants (ICAPP)*. San Francisco, United States, 2016.
- [Agu+15] Hernan AGUIRRE, Saúl ZAPOTECAS-MARTÍNEZ, Arnaud LIEFOOGHE, Sébastien VEREL et Kiyoshi TANAKA. “Approaches for many-objective optimization: analysis and comparison on MNK-landscapes”. In : *13th International Conference on Artificial Evolution (EA 2015)*. Lecture Notes in Computer Science (LNCS). Lyon, France : Springer, 2015, p. 14–28.
- [Dao+15] Fabio DAOLIO, Arnaud LIEFOOGHE, Sébastien VEREL, Hernán AGUIRRE et Kiyoshi TANAKA. “Global vs Local Search on Multi-objective NK-Landscapes: Contrasting the Impact of Problem Features”. In : *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*. GECCO '15. Madrid, Spain : ACM, 2015, p. 369–376.
- [Jan+15a] Christopher JANKEE, Sébastien VEREL, Bilel DERBEL et Cyril FONLUPT. “Distributed Adaptive Metaheuristic Selection: Comparisons of Selection Strategies”. In : *13th International Conference on Artificial Evolution (EA 2015)*. Lecture Notes in Computer Science (LNCS). Lyon, France : Springer, 2015, p. 83–96.
- [Lie+15a] Arnaud LIEFOOGHE, Sébastien VEREL, Fabio DAOLIO, Hernan AGUIRRE et Kiyoshi TANAKA. “A feature-based performance analysis in evolutionary multiobjective optimization”. In : *8th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2015)*. Lecture Notes in Computer Science. Guimarães, Portugal, 2015, p. 95–109.

- [Lie+15b] Arnaud LIEFOOGHE, Sébastien VEREL, Luis PAQUETE et Jin-Kao HAO. “Experiments on local search for bi-objective unconstrained binary quadratic programming”. In : *8th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2015)*. Lecture Notes in Computer Science. Guimarães, Portugal, 2015, p. 171–186.
- [Ver15] Sébastien VEREL. “Fitness landscape: the metaphor and beyond.” In : *9th International Conference on Learning and Optimization (LION9)*. 2015.
- [Agu+14] Hernan AGUIRRE, Arnaud LIEFOOGHE, Sébastien VEREL et Kiyoshi TANAKA. “An Analysis on Selection for High-Resolution Approximations in Many-Objective Optimization”. In : *Parallel Problem Solving from Nature - PPSN XIII*. Sous la dir. de Thomas BARTZ-BEIELSTEIN, Jürgen BRANKE, Bogdan FILIPIC et Jim SMITH. T. 8672. Lecture Notes in Computer Science. Ljubljana, Slovenia : Springer International Publishing, 2014, p. 487–497.
- [Der+14a] Bilel DERBEL, Dimo BROCKHOFF, Arnaud LIEFOOGHE et Sébastien VEREL. “On the Impact of Multiobjective Scalarizing Functions”. In : *Parallel Problem Solving from Nature - PPSN XIII*. Sous la dir. de Thomas BARTZ-BEIELSTEIN, Jürgen BRANKE, Bogdan FILIPIČ et Jim SMITH. T. 8672. Lecture Notes in Computer Science. Ljubljana, Slovenia : Springer International Publishing, 2014, p. 548–557.
- [Dro+14] Martin DROZDIK, Kiyoshi TANAKA, Hernan AGUIRRE, Sébastien VEREL, Arnaud LIEFOOGHE et Bilel DERBEL. “An Analysis of Differential Evolution Parameters on Rotated Bi-objective Optimization Functions”. In : *The Tenth International Conference on Simulated Evolution And Learning (SEAL 2014)*. Dunedin, New Zealand : Springer, 2014, p. 1–12.
- [Jan+14] Christopher JANKEE, Michel CARRARD, Sébastien VEREL et Eric RAMAT. “Etude à l’aide d’une simulation informatique multi-agents visant à saisir les conséquences de la réforme des aéroports pour les territoires”. In : *11ème Conférence Annuelle “Territoires, Espaces, et Politiques Publiques”*. Nantes, France, 2014.
- [LLV14] Manuel LÓPEZ-IBÁÑEZ, Arnaud LIEFOOGHE et Sébastien VEREL. “Local Optimal Sets and Bounded Archiving on Multi-objective NK-Landscapes with Correlated Objectives”. In : *Parallel Problem Solving from Nature - PPSN XIII*. Sous la dir. de Thomas BARTZ-BEIELSTEIN, Jürgen BRANKE, Bogdan FILIPIČ et Jim SMITH. T. 8672. Lecture Notes in Computer Science. Ljubljana, Slovenia : Springer International Publishing, 2014, p. 621–630.
- [Agu+13a] Hernan AGUIRRE, Arnaud LIEFOOGHE, Sébastien VEREL et Kiyoshi TANAKA. “A study on population size and selection lapse in many-objective optimization”. In : *IEEE Congress on Evolutionary Computation (CEC 2013)*. Cancún, Mexico, 2013, p. 1507–1514.
- [Agu+13b] Hernan AGUIRRE, Arnaud LIEFOOGHE, Sébastien VEREL et Kiyoshi TANAKA. “Effects of population size on selection and scalability in evolutionary many-objective optimization”. In : *Learning and Intelligent OptimizatioN Conference (LION 7)*. T. 7997. Lecture Notes in Computer Science. Catania, Italy, 2013, p. 450–454.
- [Bas+13] Matthieu BASSEUR, Adrien GOËFFON, Arnaud LIEFOOGHE et Sébastien VEREL. “On set-based local search for multiobjective combinatorial optimization”. In : *Genetic and Evolutionary Computation Conference (GECCO 2013)*. Amsterdam, Netherlands, 2013, p. 471–478.
- [Dao+13] Fabio DAOLIO, Sébastien VEREL, Gabriela OCHOA et Marco TOMASSINI. “Local Optima Networks of the Permutation Flow-Shop Problem”. In : *International Conference on Artificial Evolution (EA 2013)*. Sous la dir. de SPRINGER. Lecture Notes in Computer Science. Bordeaux, France, 2013, p. 41–52.

- [Lie+13] Arnaud LIEFOOGHE, Sébastien VEREL, Hernan AGUIRRE et Kiyoshi TANAKA. "What makes an instance difficult for black-box 0-1 evolutionary multiobjective optimizers?" In : *International Conference on Artificial Evolution (EA 2013)*. T. 8752. Lecture Notes in Computer Science. Bordeaux, France : Springer, 2013, p. 3–15.
- [Ver13] Sébastien VEREL. "Fitness landscapes and graphs: multimodularity, ruggedness and neutrality". In : *Proceedings of the fifth international conference on Genetic and evolutionary computation conference companion*. Amsterdam, Netherlands : ACM, 2013, p. 1013–1034.
- [Chi+12] FRANCISCO CHICANO, Fabio DAOLIO, Gabriela OCHOA, Sébastien VEREL, Marco TOMASSINI et Enrique ALBA. "Local Optima Networks, Landscape Autocorrelation and Heuristic Search Performance". In : *Parallel Problem Solving from Nature - PPSN XII*. Sous la dir. de Carlos A. COELLO, Vincenzo CUTELLO, Kalyanmoy DEB, Stephanie FORREST, Giuseppe NICOSIA et Mario PAVONE. T. 7492. Lecture Notes in Computer Science. Taormina, Italy : Springer Berlin Heidelberg, 2012, p. 337–347.
- [Dao+12] Fabio DAOLIO, Sébastien VEREL, Gabriela OCHOA et Marco TOMASSINI. "Local optima networks and the performance of iterated local search". In : *Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference*. Philadelphia, United States : ACM, 2012, p. 369–376.
- [Mar+12] Marie-Eleonore MARMION, Clarisse DHAENENS, Laetitia JOURDAN, Arnaud LIEFOOGHE et Sébastien VEREL. "Conception de recherche locale en présence de neutralité". In : *ROADEF 2012 : 13e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*. Angers, France, 2012.
- [Ver12] Sébastien VEREL. "Fitness landscapes and graphs: multimodularity, ruggedness and neutrality". In : *Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference companion*. Philadelphia, United States : ACM, 2012, p. 1013–1034.
- [DV11] Bilel DERBEL et Sébastien VEREL. "DAMS: Distributed Adaptive Metaheuristic Selection". In : *Genetic And Evolutionary Computation Conference*. Dublin, Ireland : ACM, 2011, p. 1955–1962.
- [Mar+11a] Marie-Eleonore MARMION, Clarisse DHAENENS, Laetitia JOURDAN, Arnaud LIEFOOGHE et Sébastien VEREL. "NILS: a Neutrality-based Iterated Local Search and its application to Flowshop Scheduling". In : *11th European Conference on Evolutionary Computation in Combinatorial Optimisation*. Sous la dir. de Peter MERZ et Jin-Kao HAO. T. 6622. Lecture Notes in Computer Science. Turino, Italy : Springer, 2011, p. 191–202.
- [Mar+11b] Marie-Eleonore MARMION, Clarisse DHAENENS, Laetitia JOURDAN, Arnaud LIEFOOGHE et Sébastien VEREL. "On the Neutrality of Flowshop Scheduling Fitness Landscapes". In : *Learning and Intelligent Optimization Conference (LION 5)*. T. 6683/2011. Rome, Italy : Springer, 2011, p. 238–252.
- [Mar+11c] Marie-Eleonore MARMION, Clarisse DHAENENS, Laetitia JOURDAN, Arnaud LIEFOOGHE et Sébastien VEREL. "The Road to VEGAS: Guiding the Search over Neutral Networks". In : *Genetic And Evolutionary Computation Conference*. Dublin, Ireland : ACM, 2011, p. 1979–1986.
- [Och+11] Gabriela OCHOA, Sébastien VEREL, Fabio DAOLIO et Marco TOMASSINI. "Clustering of Local Optima in Combinatorial Fitness Landscapes". In : *Learning and Intelligent Optimization Conference (LION 5)*. T. 6683/2011. Rome, Italy : Springer, 2011, p. 454–457.

- [Ver+11a] Sébastien VEREL, Fabio DAOLIO, Gabriela OCHOA et Marco TOMASSINI. “Local Optima Networks with Escape Edges”. In : *International Conference on Artificial Evolution (EA-2011)*. Angers, France, 2011, p. 10–23.
- [VLD11] Sébastien VEREL, Arnaud LIEFOOGHE et Clarisse DHAENENS. “Set-based Multiobjective Fitness Landscapes: A Preliminary Study”. In : *Genetic And Evolutionary Computation Conference*. Dublin, Ireland : ACM, 2011, p. 769–776.
- [Ver+11b] Sébastien VEREL, Arnaud LIEFOOGHE, Jérémie HUMEAU, Laetitia JOURDAN et Clarisse DHAENENS. “On the Effect of Connectedness for Biobjective Multiple and Long Path Problems”. In : *Learning and Intelligent OptimizatioN Conference (LION 5)*. Sous la dir. de C.A. Coello COELLO. T. 6683. Rome, Italy : Springer, 2011, p. 31–45.
- [Ver+11c] Sébastien VEREL, Arnaud LIEFOOGHE, Laetitia JOURDAN et Clarisse DHAENENS. “Analyzing the Effect of Objective Correlation on the Efficient Set of MNK-Landscapes”. In : *Learning and Intelligent OptimizatioN Conference (LION 5)*. Sous la dir. de C.A. Coello COELLO. T. 6683/2011. Lecture Notes in Computer Science (LNCS). Rome, Italy : Springer, 2011, p. 116–130.
- [Ver+11d] Sébastien VEREL, Arnaud LIEFOOGHE, Laetitia JOURDAN et Clarisse DHAENENS. “Pareto Local Optima of Multiobjective NK-Landscapes with Correlated Objectives”. In : *11th European Conference on Evolutionary Computation in Combinatorial Optimisation*. Sous la dir. de Peter MERZ et Jin-Kao HAO. T. 6622. Lecture Notes in Computer Science. Turino, Italy : Springer, 2011, p. 226–237.
- [Dao+10] Fabio DAOLIO, Sébastien VEREL, Gabriela OCHOA et Marco TOMASSINI. “Local Optima Networks of the Quadratic Assignment Problem”. In : *IEEE world conference on computational intelligence (WCCI - CEC)*. Barcelona, Spain, 2010, p. 3145–3152.
- [OVT10] Gabriela OCHOA, Sébastien VEREL et Marco TOMASSINI. “First-improvement vs. Best-improvement Local Optima Networks of NK Landscapes”. In : *11th International Conference on Parallel Problem Solving From Nature*. Krakow, Poland, 2010, p. 104–113.
- [Och+10] Gabriela OCHOA, Sébastien VEREL, Marco TOMASSINI et Fabio DAOLIO. “Local Optima Networks: Current Results and Perspectives”. In : *4th Workshop on Theory of Randomized Search Heuristics*. Paris, France, 2010.
- [VCC10] Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. “States based evolutionary algorithm”. In : *Workshop selfstar at conference PPSN*. Krakow, Poland, 2010.
- [Ver+10] Sébastien VEREL, Laetitia JOURDAN, Clarisse DHAENENS et Arnaud LIEFOOGHE. “Set-based Multiobjective Fitness Landscapes: definition, properties.” In : *4th Workshop on Theory of Randomized Search Heuristics*. Paris, France, 2010.
- [VO10] Sébastien VEREL et Gabriela OCHOA. “Fitness landscapes and graphs: multimodularity, ruggedness and neutrality”. In : *WCCI 2010*. Barcelona, Spain : IEEE, 2010, p. 3593–3656.
- [Ber+09] Maroun BERCAHI, Manuel CLERGUE, Philippe COLLARD et Sébastien VEREL. “Do not Choose Representation just Change: An Experimental Study in States based EA”. In : *GECCO '09 the 11th Annual conference on Genetic and evolutionary computation*. Montréal, Canada : ACM, 2009, p. 1799–1800.
- [OVT09] Gabriela OCHOA, Sébastien VEREL et Marco TOMASSINI. “Local Optima Networks of NK Landscapes with and without Neutrality”. In : *3rd Workshop on Theory of Randomized Search Heuristics*. Birmingham, United Kingdom, 2009.

- [Sim+09] David SIMONCINI, Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. “Centric selection: a way to tune the exploration/exploitation trade-off”. In : *GECCO’09*. Montreal, Canada : ACM, 2009, p. 891–898.
- [Van+09] Leonardo VANNESCHI, Sébastien VEREL, Philippe COLLARD et Marco TOMASSINI. “NK landscapes difficulty and Negative Slope Coefficient: How Sampling Influences the Results”. In : *evoNum workshop of evostar conference*. LNCS. Tubingen, Germany : Springer, 2009, p. 100–110.
- [Ver09] Sébastien VEREL. “Fitness landscapes and graphs: multimodularity, ruggedness and neutrality”. In : *GECCO ’09*. Montreal, Canada : ACM, 2009, p. 3593–3656.
- [Bac+08] Thierry BACCINO, Maud KICKA, Laurent DUMERCY et Sébastien VEREL. “Investigating Cognitive Load by Normalized Task-Evoked Pupillary Response”. In : *Workshop on Cognition and the Web 2008: Information Processing, Comprehension and Learning*. France, 2008, p. 45–52.
- [Och+08] Gabriela OCHOA, Marco TOMASSINI, Sébastien VEREL et Christian DARABOS. “A Study of NK Landscapes’ Basins and Local Optima Networks”. In : *Genetic And Evolutionary Computation Conference*. best paper nomination. Atlanta, United States : ACM New York, NY, USA, 2008, p. 555–562.
- [TOV08] Marco TOMASSINI, Gabriela OCHOA et Sébastien VEREL. “The network structure of hard combinatorial landscapes.” In : *BCNetWORKSHOP trends and perspectives in complex networks*. Barcelona, Spain, 2008.
- [VOT08] Sébastien VEREL, Gabriela OCHOA et Marco TOMASSINI. “The Connectivity of NK Landscapes’ Basins: A Network Analysis”. In : *Artificial Life XI*. Sous la dir. de S. BULLOCK, J. NOBLE, R. WATSON et M. A. BEDAU. Winchester, France : MIT Press, Cambridge, MA, 2008, p. 648–655.
- [Ber+07] Maroun BERCAHI, Philippe COLLARD, Manuel CLERGUE et Sébastien VEREL. “Evolving Dynamic Change and Exchange of Genotype Encoding in Genetic Algorithms for Difficult Optimization Problems”. In : *IEEE Congress on Evolutionary Computation CEC2007*. singapore, Singapore : IEEE Press, 2007, p. 4516–4523.
- [Def+07] Michael DEFOIN PLATEL, Sébastien VEREL, Manuel CLERGUE et Malik CHAMI. “Density estimation with Genetic Programming for Inverse Problem solving”. In : *EuroGP’07, the 10th European Conference on Genetic Programming*. Sous la dir. de Marc EBNER, Michael O’NEILL, Aniko EKART, Leonardo VANNESCHI et Anna Isabel ESPARCIA-ALCAZAR. T. 4445. Lecture Notes in Computer Science. Valencia, Spain : Springer, 2007, p. 45–54.
- [Sim+07] David SIMONCINI, Philippe COLLARD, Sébastien VEREL et Manuel CLERGUE. “On the Influence of Selection Operators on Performances in Cellular Genetic Algorithms”. In : *IEEE Congress on Evolutionary Computation CEC2007*. singapore, Singapore : IEEE Press, 2007, p. 4706–4713.
- [Van+07] Leonardo VANNESCHI, Philippe COLLARD, Sébastien VEREL, Marco TOMASSINI, Yuri PIROLA et Giancarlo MAURI. “A Comprehensive View of Fitness Landscapes with Neutrality and Fitness Clouds”. In : *EuroGP 2007*. Sous la dir. de Marc EBNER, Michael O’NEILL, Aniko EKART, Leonardo VANNESCHI et Anna Isabel ESPARCIA-ALCAZAR. T. 4445. Valencia, Spain : Springer, 2007, p. 241–250.
- [VV07] Leonardo VANNESCHI et Sébastien VEREL. “Fitness landscapes and problem hardness in evolutionary computation”. In : *Genetic And Evolutionary Computation Conference*. Tutorial. London, United Kingdom : ACM Press, 2007, p. 3690–3733.

- [Bea+06] William BEAUDOIN, Sébastien VEREL, Philippe COLLARD et Cathy ESCAZUT. "Deceptiveness and Neutrality - the ND family of fitness landscapes". In : *Genetic And Evolutionary Computation Conference*. Sous la dir. de M. Keijzer et AL. Seattle, United States : ACM Press, 2006, p. 507–514.
- [Col+06] Philippe COLLARD, Giancarlo MAURI, Yuri PIROLA, Marco TOMASSINI, Leonardo VANNESCHI et Sébastien VEREL. "A Quantitative Study of Neutrality in GP Boolean Landscapes". In : *Genetic And Evolutionary Computation Conference*. Sous la dir. de M. Keijzer et AL. Seattle, United States : ACM Press, 2006, p. 895–902.
- [Sim+06a] David SIMONCINI, Philippe COLLARD, Sébastien VEREL et Manuel CLERGUE. "From Cells to Islands: An unified Model of Cellular Parallel Genetic Algorithms". In : *7th International Conference on Cellular Automata For Research and Industry - ACRI 2006*. Sous la dir. de Samira El YACOUBI, Bastien CHOPARD et Stefania BANDINI. T. 4173. Lecture Notes in Computer Science. Perpignan, France : Springer, 2006, p. 248–257.
- [Sim+06b] David SIMONCINI, Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. "Anisotropic selection in cellular genetic algorithms". In : *Genetic And Evolutionary Computation Conference*. Sous la dir. de M. Keijzer et AL. Seattle, United States : ACM Press, 2006, p. 559–566.
- [Sim+05] David SIMONCINI, Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. "A preliminary investigation of anisotropic selection in cellular genetic algorithms". In : *Poster at Evolution Artificielle, 7th International Conference*. Lille, France, 2005.

C

BIBLIOGRAPHIE

- [AAZ16] Rolando ARMAS, Hernán AGUIRRE et Kiyoshi ZAPOTECAS-MARTÍNEZ Saúl and Tanaka. “Traffic Signal Optimization: Minimizing Travel Time and Fuel Consumption”. In : *Artificial Evolution: 12th International Conference, Evolution Artificielle, EA 2015, Lyon, France, October 26-28, 2015. Revised Selected Papers*. Sous la dir. de Stéphane BONNEVAY, Pierrick LEGRAND, Nicolas MONMARCHÉ, Evelyne LUTTON et Marc SCHOENAUER. Cham : Springer International Publishing, 2016, p. 29–43.
- [Bel+16] Nacim BELKHIR, Johann DRÉO, Pierre SAVÉANT et Marc SCHOENAUER. “Feature Based Algorithm Configuration: A Case Study with Differential Evolution”. In : *Parallel Problem Solving from Nature – PPSN XIV: 14th International Conference, Edinburgh, UK, September 17-21, 2016, Proceedings*. Sous la dir. de Julia HANDL, Emma HART, R. Peter LEWIS, Manuel LÓPEZ-IBÁÑEZ, Gabriela OCHOA et Ben PAECHTER. Cham : Springer International Publishing, 2016, p. 156–166.
- [GLS16] Adrien GOËFFON, Frédéric LARDEUX et Frédéric SAUBION. “Simulating non-stationary operators in search algorithms”. In : *Applied Soft Computing* 38 (2016), p. 257–268.
- [Bar15] A-L. BARABÁSI. *Network Science*. <http://barabasi.com/NetworkScienceBook>, 2015.
- [DD15] Benjamin DOERR et Carola DOERR. “Optimal Parameter Choices Through Self-Adjustment: Applying the 1/5-th Rule in Discrete Settings”. In : (2015).
- [Muñ+15] Mario A. MUÑOZ, Yuan SUN, Michael KIRLEY et Saman K. HALGAMUGE. “Algorithm selection for black-box continuous optimization problems: A survey on methods and challenges”. In : *Information Sciences* 317 (2015), p. 224–245.
- [TP15] M-H TAYARANI-N et Adam PRÜGEL-BENNETT. “Anatomy of the fitness landscape for dense graph-colouring problem”. In : *Swarm and Evolutionary Computation* 22 (2015), p. 47–65.
- [Whi15] Darrell WHITLEY. “Blind No More: Constant Time Non-Random Improving Moves and Exponentially Powerful Recombination”. In : *Genetic and Evolutionary Computation Conference, GECCO 2015, Madrid, Spain, July 11-15, 2015, Companion Material Proceedings*. 2015, p. 391–407.
- [AR14] Khulood ALYAHYA et Jonathan E ROWE. “Phase transition and landscape properties of the number partitioning problem”. In : *Evolutionary Computation in Combinatorial Optimisation*. Springer, 2014, p. 206–217.
- [BP14] Petr BAUDIŠ et Pet POŠIK. “Online Black-Box Algorithm Portfolios for Continuous Optimization”. In : *Parallel Problem Solving from Nature–PPSN XIII*. Springer, 2014, p. 40–49.
- [CWA14] Francisco CHICANO, Darrell WHITLEY et Enrique ALBA. “Exact computation of the expectation surfaces for uniform crossover along with bit-flip mutation”. In : *Theor. Comput. Sci.* 545 (2014), p. 76–93.
- [Hut+14] Frank HUTTER, Lin XU, Holger H HOOS et Kevin LEYTON-BROWN. “Algorithm runtime prediction: Methods and evaluation”. In : *Artificial Intelligence* 206 (2014), p. 79–111.

- [Koc+14] Gary KOCHENBERGER, Jin-Kao HAO, Fred GLOVER, Mark LEWIS, Zhipeng LÜ, Haibo WANG et Yang WANG. "The unconstrained binary quadratic programming problem: a survey". In : *Journal of Combinatorial Optimization* 28.1 (2014), p. 58–81.
- [Kot14] Lars KOTTHOFF. "Algorithm Selection for Combinatorial Search Problems: A Survey". In : *AI Magazine* (2014).
- [ME14a] Katherine M. MALAN et Andries P. ENGELBRECHT. "Fitness Landscape Analysis for Metaheuristic Performance Prediction". In : *Recent Advances in the Theory and Application of Fitness Landscapes*. Sous la dir. d'Hendrik RICHTER et Andries ENGELBRECHT. T. 6. Emergence, Complexity and Computation. Springer Berlin Heidelberg, 2014, p. 103–132.
- [ME14b] K.M. MALAN et A.P. ENGELBRECHT. "Characterising the searchability of continuous optimisation problems for PSO". English. In : *Swarm Intelligence* 8.4 (2014), p. 275–302.
- [Qi+14] Yutao QI, Xiaoliang MA, Fang LIU, Licheng JIAO, Jianyong SUN et Jianshe WU. "MOEA/D with adaptive weight adjustment". In : *Evolutionary computation* 22.2 (2014), p. 231–264.
- [TP14] Mohammad-H TAYARANI-N et Adam PRUGEL-BENNETT. "On the landscape of combinatorial optimization problems". In : *Evolutionary Computation, IEEE Transactions on* 18.3 (2014), p. 420–434.
- [Whi+14] Darrell WHITLEY, Andrew M. SUTTON, Gabriela OCHOA et Francisco CHICANO. "The component model for elementary landscapes and partial neighborhoods". In : *Theor. Comput. Sci.* 545 (2014), p. 59–75.
- [Can+13] Caner CANDAN, Adrien GOËFFON, Frédéric LARDEUX et Frédéric SAUBION. "Parameter Setting with Dynamic Island Models". In : *Learning and Intelligent Optimization*. Springer, 2013, p. 253–258.
- [IAN13] Hisao ISHIBUCHI, Naoya AKEDO et Yusuke NOJIMA. "A Study on the Specification of a Scalarizing Function in MOEA/D for Many-Objective Knapsack Problems". In : *LION7*. 2013, p. 231–246.
- [LW13] Per Kristian LEHRE et Carsten WITT. "General drift analysis with tail bounds". In : *arXiv preprint arXiv:1307.2559* (2013).
- [ME13] Katherine M. MALAN et Andries P. ENGELBRECHT. "A survey of techniques for characterising fitness landscapes and some possible ways forward". In : *Information Sciences* 241 (2013), p. 148–163.
- [Sch+13] Julien SCHLEICH, Grégoire DANOY, Bernabé DORRONSORO et Pascal BOUVRY. "An Overlay Approach for Optimising Small-World Properties in VANETs". In : *Applications of Evolutionary Computation: 16th European Conference, EvoApplications 2013, Vienna, Austria, April 3-5, 2013. Proceedings*. Sous la dir. d'Anna I. ESPARCIA-ALCÁZAR. Berlin, Heidelberg : Springer Berlin Heidelberg, 2013, p. 32–41.
- [SJ13] Christine SOLNON et Narendra JUSSIEN. "Constructive Heuristic Approaches". In : *Ant Colony Optimization and Constraint Programming*. John Wiley & Sons, 2013, p. 85–92.
- [Vaz13] Vijay V VAZIRANI. *Approximation algorithms*. Springer Science & Business Media, 2013.
- [YCZ13] Shiu Yin YUEN, Chi Kin CHOW et Xin ZHANG. "Which Algorithm Should I Choose at Any Point of the Search: An Evolutionary Portfolio Approach". In : *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation. GECCO '13*. Amsterdam, The Netherlands : ACM, 2013, p. 567–574.

- [Can+12] Caner CANDAN, Adrien GOEFFON, Frédéric LARDEUX et Frédéric SAUBION. “A dynamic island model for adaptive operator selection”. In : *Proceedings of the 14th annual conference on Genetic and evolutionary computation*. ACM. 2012, p. 1253–1260.
- [MNY12] Riadh MOUSSI, Ndèye Fatma NDIAYE et Adnan YASSINE. “Hybrid genetic simulated annealing algorithm (HGSAA) to solve storage container problem in port”. In : *Asian Conference on Intelligent Information and Database Systems*. Springer. 2012, p. 301–310.
- [PT12] Adam PRÜGEL-BENNETT et Mohammad-Hassan TAYARANI-NAJARAN. “Maximum satisfiability: Anatomy of the fitness landscape for a hard combinatorial optimization problem”. In : *Evolutionary Computation, IEEE Transactions on* 16.3 (2012), p. 319–338.
- [Dur+11] Juan J DURILLO, Qingfu ZHANG, Antonio J NEBRO et Enrique ALBA. “Distribution of computational effort in parallel MOEA/D”. In : *International Conference on Learning and Intelligent Optimization*. Springer. 2011, p. 488–502.
- [HHL11] Frank HUTTER, Holger H. HOOS et Kevin LEYTON-BROWN. “Sequential Model-based Optimization for General Algorithm Configuration”. In : *Proceedings of the 5th International Conference on Learning and Intelligent Optimization*. LION’05. Rome, Italy : Springer-Verlag, 2011, p. 507–523.
- [LLY11a] Jinlong LI, Guanzhou LU et Xin YAO. “Fitness Landscape-Based Parameter Tuning Method for Evolutionary Algorithms for Computing Unique Input Output Sequences”. In : *Neural Information Processing*. Sous la dir. de Bao-Liang LU, Liqing ZHANG et James KWOK. T. 7063. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, p. 453–460.
- [LLY11b] Guanzhou LU, Jinlong LI et Xin YAO. “Fitness-probability cloud and a measure of problem hardness for evolutionary algorithms”. In : *Evolutionary Computation in Combinatorial Optimization*. Springer, 2011, p. 108–117.
- [Mer+11] Olaf MERSMANN, Bernd BISCHL, Heike TRAUTMANN, Mike PREUSS, Claus WEIHS et Günter RUDOLPH. “Exploratory Landscape Analysis”. In : *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*. GECCO ’11. Dublin, Ireland : ACM, 2011, p. 829–836.
- [Mou+11] Riadh MOUSSI, Adnan YASSINE, Ali KANSOU et Thierry GALINHO. “Scheduling of lifting vehicles with time windows in an automated port container terminal”. In : *2011 4th International Conference on Logistics*. IEEE. 2011, p. 55–61.
- [WMC11] Yanghui WU, John MCCALL et David CORNE. “Fitness landscape analysis of Bayesian network structure learning”. In : *Evolutionary Computation (CEC), 2011 IEEE Congress on*. IEEE. 2011, p. 981–988.
- [Bir+10] Mauro BIRATTARI, Zhi YUAN, Prasanna BALAPRAKASH et Thomas STÜTZLE. “F-Race and iterated F-Race: An overview”. In : *Experimental methods for the analysis of optimization algorithms*. Springer, 2010, p. 311–336.
- [CG10] Marco CHIARANDINI et Yuri GOEGBEUR. “Mixed Models for the Analysis of Optimization Algorithms”. In : *Experimental Methods for the Analysis of Optimization Algorithms*. Springer, 2010, p. 225–264.
- [Fia+10] Álvaro FIALHO, Luis DA COSTA, Marc SCHOENAUER et Michele SEBAG. “Analyzing bandit-based adaptive operator selection mechanisms”. In : *Annals of Mathematics and Artificial Intelligence* 60.1-2 (2010), p. 25–64.
- [GLH10] Fred GLOVER, Z. LÜ et Jin-Kao HAO. “Diversification-driven tabu search for unconstrained binary quadratic problems”. In : *4OR: A Quarterly Journal of Operations Research* 8.3 (2010), p. 239–253.

- [KMT10] Maciej KURANT, Athina MARKOPOULOU et Patrick THIRAN. “On the bias of bfs (breadth first search)”. In : *Teletraffic Congress (ITC), 2010 22nd International*. IEEE. 2010, p. 1–8.
- [LMS10] Helena R LOURENÇO, Olivier C MARTIN et Thomas STÜTZLE. “Iterated local search: Framework and applications”. In : *Handbook of Metaheuristics*. Springer, 2010, p. 363–397.
- [YTS10] Mouadh YAGOUBI, Ludovic THOBOIS et Marc SCHOENAUER. “An asynchronous steady-state nsga-ii algorithm for multi-objective optimization of diesel combustion”. In : *Proceedings of the 2nd International Conference on Engineering Optimization*. 2010, p. 77.
- [ZTB10] E. ZITZLER, L. THIELE et J. BADER. “On Set-Based Multiobjective Optimization”. In : *IEEE Transactions on Evolutionary Computation* 14.1 (2010), p. 58–79.
- [AST09] Carlos ANSÓTEGUI, Meinolf SELLMANN et Kevin TIERNEY. “A gender-based genetic algorithm for the automatic configuration of algorithms”. In : *Principles and Practice of Constraint Programming-CP 2009*. Springer, 2009, p. 142–157.
- [Gar09a] Diego GARLASCHELLI. “The weighted random graph model”. In : *New Journal of Physics* 11.7 (2009), p. 073005.
- [Gar09b] Deon GARRETT. “Plateau connection structure and multiobjective metaheuristic performance”. In : *2009 IEEE Congress on Evolutionary Computation*. IEEE. 2009, p. 1281–1288.
- [Hut+09] Frank HUTTER, Holger H HOOS, Kevin LEYTON-BROWN et Thomas STÜTZLE. “ParamILS: an automatic algorithm configuration framework”. In : *Journal of Artificial Intelligence Research* 36.1 (2009), p. 267–306.
- [Jün+09] Michael JÜNGER et al. *50 Years of integer programming 1958-2008: From the early years to the state-of-the-art*. Springer Science & Business Media, 2009.
- [Lie+09] Arnaud LIEFOOGHE, Salma MESMOUDI, Jérémie HUMEAU, Laetitia JOURDAN et El-Ghazali TALBI. “A Study on Dominance-Based Local Search Approaches for Multiobjective Combinatorial Optimization”. In : *Engineering Stochastic Local Search Algorithms. Designing, Implementing and Analyzing Effective Heuristics: Second International Workshop, SLS 2009, Brussels, Belgium, September 3-4, 2009. Proceedings*. Sous la dir. de Thomas STÜTZLE, Mauro BIRATTARI et Holger H. Hoos. Berlin, Heidelberg : Springer Berlin Heidelberg, 2009, p. 120–124.
- [Lou+09] L. LOUKIL, M. MEHDI, N. MELAB, E. G. TALBI et P. BOUVRY. “A parallel hybrid genetic algorithm-simulated annealing for solving Q3AP on computational grid”. In : *Parallel Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on*. 2009, p. 1–8.
- [Mat+09] Jorge MATURANA, Álvaro FIALHO, Frédéric SAUBION, Marc SCHOENAUER et Michele SEBAG. “Extreme compass and dynamic multi-armed bandits for adaptive operator selection”. In : *Evolutionary Computation, 2009. CEC’09. IEEE Congress on*. IEEE. 2009, p. 365–372.
- [PS09] Luís PAQUETE et Thomas STÜTZLE. “Clusters of Non-dominated Solutions in Multiobjective Combinatorial Optimization: An Experimental Analysis”. In : *Multiobjective Programming and Goal Programming: Theoretical Results and Practical Applications*. T. 618. Lecture Notes in Economics and Mathematical Systems. Springer, 2009, p. 69–77.
- [SE09] Selmar K SMIT et Agoston E EIBEN. “Comparing parameter tuning methods for evolutionary algorithms”. In : *Evolutionary Computation, 2009. CEC’09. IEEE Congress on*. IEEE. 2009, p. 399–406.
- [Smio9] Kate A. SMITH-MILES. “Cross-disciplinary Perspectives on Meta-learning for Algorithm Selection”. In : *ACM Comput. Surv.* 41.1 (2009), 6:1–6:25.

- [Tal09] El-Ghazali TALBI. *Metaheuristics: from design to implementation*. T. 74. John Wiley & Sons, 2009.
- [Fia+08] Álvaro FIALHO, Luis DA COSTA, Marc SCHOENAUER et Michèle SEBAG. “Extreme value based adaptive operator selection”. In : *Parallel Problem Solving from Nature–PPSN X*. Springer, 2008, p. 175–184.
- [ITNo8] Hisao ISHIBUCHI, Noritaka TSUKAMOTO et Yusuke NOJIMA. “Evolutionary many-objective optimization: A short review”. In : *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*. 2008, p. 2419–2426.
- [Llo8] Raluca LEFTICARU et Florentin IPATE. “A comparative landscape analysis of fitness functions for search-based testing”. In : *Symbolic and Numeric Algorithms for Scientific Computing, 2008. SYNASC’08. 10th International Symposium on*. IEEE. 2008, p. 201–208.
- [TPCo8] J. TAVARES, F.B. PEREIRA et E. COSTA. “Multidimensional Knapsack Problem: A Fitness Landscape Analysis”. In : *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 38.3 (2008), p. 604–616.
- [Xu+08] Lin XU, Frank HUTTER, Holger H. HOOS et Kevin LEYTON-BROWN. “SATzilla: Portfolio-based Algorithm Selection for SAT”. In : *J. Artif. Intell. Res. (JAIR)* 32 (2008), p. 565–606.
- [ATo7] H. E. AGUIRRE et Kiyoshi TANAKA. “Working principles, behavior, and performance of MOEAs on MNK-landscapes”. In : *European Journal of Operational Research* 181.3 (2007), p. 1670–1690.
- [BNEo7] Nicola BEUME, Boris NAUJOKS et Michael EMMERICH. “SMS-EMOA: Multiobjective selection based on dominated hypervolume”. In : *European Journal of Operational Research* 181.3 (2007), p. 1653–1669.
- [Eib+07] A. E. EIBEN, Zbigniew MICHALEWICZ, Marc SCHOENAUER et J. E. SMITH. “Parameter Control in Evolutionary Algorithms”. In : *Parameter Setting in Evolutionary Algorithms*. Springer, 2007, p. 19–46.
- [GD07] Deon GARRETT et Dipankar DASGUPTA. “Multiobjective landscape analysis and the generalized assignment problem”. In : *International Conference on Learning and Intelligent Optimization*. Springer. 2007, p. 110–124.
- [ZLo7] Qingfu ZHANG et Hui LI. “MOEA/D: A multiobjective evolutionary algorithm based on decomposition”. In : *Evolutionary Computation, IEEE Transactions on* 11.6 (2007), p. 712–731.
- [Dre+06] Johann DREO, Alain PETROWSKI, Patrick SIARRY et Eric TAILLARD. *Metaheuristics for hard optimization: methods and case studies*. Springer Science & Business Media, 2006.
- [Ehro6] Matthias EHRGOTT. *Multicriteria optimization*. Springer Science & Business Media, 2006.
- [GD06] Deon GARRETT et Dipankar DASGUPTA. “Analyzing the Performance of Hybrid Evolutionary Algorithms for the Multiobjective Quadratic Assignment Problem”. In : *Congress on Evolutionary Computation (CEC 2006)*. IEEE Service Center, 2006, p. 1710–1717.
- [GEWo6] Pierre GEURTS, Damien ERNST et Louis WEHENKEL. “Extremely randomized trees”. In : *Machine Learning* 63.1 (2006), p. 3–42.
- [GKo6] Fred W GLOVER et Gary A KOCHENBERGER. *Handbook of metaheuristics*. T. 57. Springer Science & Business Media, 2006.

- [LV06] E. LUTTON et J. L. VEHEL. "Local regularity of fitness landscapes and the performance of a simple ES". In : *2006 IEEE International Conference on Evolutionary Computation*. 2006, p. 451–455.
- [PS06] Luís PAQUETE et Thomas STÜTZLE. "A study of stochastic local search algorithms for the biobjective QAP with correlated flow matrices". In : *European Journal of Operational Research* 169.3 (2006), p. 943–959.
- [Roto6] Franz ROTHLAUF. *Representations for Genetic and Evolutionary Algorithms*. Springer Berlin Heidelberg, 2006.
- [VCC06] Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. "Measuring the Evolvability Landscape to study Neutrality". In : *Genetic and Evolutionary Computation – GECCO-2006*. Sous la dir. de M. KEIJZER et et AL. Seattle, WA, United States : ACM Press, 2006, p. 613–614.
- [AH05] A. AUGER et N. HANSEN. "Performance evaluation of an advanced local search evolutionary algorithm". In : *Evolutionary Computation, 2005. The 2005 IEEE Congress on*. T. 2. IEEE. 2005, p. 1777–1784.
- [Fon+05] Carlos M FONSECA, Joshua D KNOWLES, Lothar THIELE et Eckart ZITZLER. "A tutorial on the performance assessment of stochastic multiobjective optimizers". In : *Third International Conference on Evolutionary Multi-Criterion Optimization (EMO 2005)*. T. 216. 2005, p. 240.
- [Orro5] H Allen ORR. "The genetic theory of adaptation: a brief history". In : *Nature Reviews Genetics* 6.2 (2005), p. 119–127.
- [Thio5] Dirk THIERENS. "An adaptive pursuit strategy for allocating operator probabilities". In : *Proceedings of the 7th annual conference on Genetic and evolutionary computation*. ACM. 2005, p. 1539–1546.
- [WWC05] Daniel M WEINREICH, Richard A WATSON et Lin CHAO. "Perspective: sign epistasis and genetic constraint on evolutionary trajectories". In : *Evolution* 59.6 (2005), p. 1165–1174.
- [CMT04] S. CAHON, N. MELAB et E.-G. TALBI. "ParadisEO: A Framework for the Reusable Design of Parallel and Distributed Metaheuristics". In : *Journal of Heuristics* 10.3 (2004), p. 357–380.
- [CNM04] Aaron CLAUSET, Mark EJ NEWMAN et Cristopher MOORE. "Finding community structure in very large networks". In : *Physical review E* 70.6 (2004), p. 066111.
- [LTZ04] Marco LAUMANN, Lothar THIELE et Eckart ZITZLER. "Running time analysis of evolutionary algorithms on a simplified multiobjective knapsack problem". In : *Nat Comput* 3.1 (2004), p. 37–51.
- [PCS04] L. PAQUETE, M. CHIARANDINI et T. STÜTZLE. "Pareto Local Optimum Sets in the Biobjective Traveling Salesman Problem: An Experimental Study". In : *Metaheuristics for Multiobjective Optimisation*. T. 535. Lecture Notes in Economics and Mathematical Systems. Springer, 2004. Chap. 7, p. 177–199.
- [Wato4] D. J. WATTS. "The "new" science of networks". In : *Annual Review of Sociology* 30 (2004), p. 243–270.
- [ZK04] Eckart ZITZLER et Simon KÜNZLI. "Indicator-based selection in multiobjective search". In : *Parallel Problem Solving from Nature-PPSN VIII*. Springer. 2004, p. 832–842.
- [Bas+03] U. BASTOLLA, M. PORTO, H. E. ROMAN et M. VENDRUSCOLO. "Statistical properties of neutral evolution". In : *Journal Molecular Evolution* 57.S (2003), p. 103–119.

- [KC03] Joshua KNOWLES et David CORNE. "Instance generators and test suites for the multiobjective quadratic assignment problem". In : *International Conference on Evolutionary Multi-Criterion Optimization*. Springer. 2003, p. 295–310.
- [RG03] Franz ROTHLAUF et David E GOLDBERG. "Redundant representations in evolutionary computation". In : *Evolutionary Computation* 11.4 (2003), p. 381–415.
- [VCC03] Sébastien VEREL, Philippe COLLARD et Manuel CLERGUE. "Where are Bottlenecks in NK Fitness Landscapes?" In : *Evolutionary Computation, 2003. CEC'03*. Canberra, Australia : IEEE Press, 2003, p. 273–280.
- [AZSo2] Jarmo T ALANDER, Lyudmila A ZINCHENKO et Sergey N SOROKIN. "Analysis of fitness landscape properties for evolutionary antenna design". In : *Proceedings 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002)*. IEEE. 2002, p. 363.
- [Coe02] Carlos A Coello COELLO. "Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art". In : *Computer Methods in Applied Mechanics and Engineering* 191.11–12 (2002), p. 1245–1287.
- [Deb+02] K. DEB, S. AGRAWAL, A. PRATAP et T. MEYARIVAN. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II". In : *IEEE Transactions on Evolutionary Computation* 6.2 (2002), p. 182–197.
- [Doy02] J. P. K. DOYE. "The network topology of a potential energy landscape: a static scale-free network". In : *Phys. Rev. Lett.* 88 (2002), p. 238701.
- [KC02] Joshua KNOWLES et David CORNE. "Towards landscape analyses to inform the design of a hybrid local search for the multiobjective quadratic assignment problem". In : *Soft Computing Systems: Design, Management and Applications*. T. 2002. Amsterdam, The Netherlands : IOS Press, 2002, p. 271–279.
- [LL02] Pedro LARRANAGA et Jose A LOZANO. *Estimation of distribution algorithms: A new tool for evolutionary computation*. T. 2. Springer Science & Business Media, 2002.
- [RS02a] Christian M REIDYS et Peter F STADLER. "Combinatorial landscapes". In : *SIAM review* 44.1 (2002), p. 3–54.
- [RS02b] Jan W RIVKIN et Nicolaj SIGGELKOW. "Organizational sticking points on NK landscapes". In : *Complexity* 7.5 (2002), p. 31–43.
- [RS02c] Jan W RIVKIN et Nicolaj SIGGELKOW. "Organizational sticking points on NK landscapes". In : *Complexity* 7.5 (2002), p. 31–43.
- [Stao2] Peter F STADLER. "Fitness landscapes". In : *Biological evolution and statistical physics*. Springer, 2002, p. 183–204.
- [Bre01] Leo BREIMAN. "Random Forests". In : *Machine Learning* 45.1 (2001), p. 5–32.
- [Debo1] Kalyanmoy DEB. *Multi-objective optimization using evolutionary algorithms*. T. 16. John Wiley & Sons, 2001.
- [Goto1] Jens GOTTLIEB. "On the Feasibility Problem of Penalty-Based Evolutionary Algorithms for Knapsack Problems". In : *Applications of Evolutionary Computing*. Sous la dir. d'Egbert J. W. BOERS. T. 2037. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2001, p. 50–59.
- [AP00] Pierre AUGER et Rafael Bravo de la PARRA. "Methods of aggregation of variables in population dynamics". In : *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie* 323.8 (2000), p. 665–674.
- [MF00a] Peter MERZ et Bernd FREISLEBEN. "Fitness landscape analysis and memetic algorithms for the quadratic assignment problem". In : *Evolutionary Computation, IEEE Transactions on* 4.4 (2000), p. 337–352.

- [MFoob] Peter MERZ et Bernd FREISLEBEN. "Fitness landscapes, memetic algorithms, and greedy operators for graph bipartitioning". In : *Evolutionary Computation* 8.1 (2000), p. 61–91.
- [BA99] A-L. BARABÁSI et R. ALBERT. "Emergence of scaling in random networks". In : *Science* 286 (1999), p. 509–512.
- [Fon+99] Cyril FONLUPT, Denis ROBILLIARD, Philippe PREUX et El-Ghazali TALBI. "Fitness Landscapes and Performance of Meta-Heuristics". In : *Meta-Heuristics*. Sous la dir. de Stefan VOSS, Silvano MARTELLO, IbrahimH. OSMAN et Catherine ROUCAIROL. Springer US, 1999, p. 257–268.
- [Mie99] K. MIETTINEN. *Nonlinear Multiobjective Optimization*. Boston, MA, USA : Kluwer, 1999.
- [BH98] P.C. BORGES et M.P. HANSEN. *A basis for future successes in multiobjective combinatorial optimization*. Rapp. tech. IMM-REP-1998-8. Lyngby, Denmark : Institute of Mathematical Modelling, Technical University of Denmark, 1998.
- [LV98] E. LUTTON et J. Levy VEHEL. "Holder functions and deception of genetic algorithms". In : *IEEE Transactions on Evolutionary Computation* 2.2 (1998), p. 56–71.
- [WS98] D. J. WATTS et S. H. STROGATZ. "Collective dynamics of 'small-world' networks". In : *Nature* 393 (1998), p. 440–442.
- [HW97] Robert B HECKENDORN et L Darrell WHITLEY. "A Walsh Analysis of NK-Landscapes." In : *ICGA*. Citeseer. 1997, p. 41–48.
- [HME97] Robert HINTERDING, Zbigniew MICHALEWICZ et Agoston E EIBEN. "Adaptation in evolutionary computation: A survey". In : *Evolutionary Computation, 1997., IEEE International Conference on*. IEEE. 1997, p. 65–69.
- [SW97] Peter F STADLER et Günter P WAGNER. "Algebraic theory of recombination spaces". In : *Evolutionary computation* 5.3 (1997), p. 241–275.
- [WM97] David H WOLPERT et William G MACREADY. "No free lunch theorems for optimization". In : *Evolutionary Computation, IEEE Transactions on* 1.1 (1997), p. 67–82.
- [Gil96] Steven G GILMOUR. "The Interpretation of Mallows's C_p -Statistic". In : *The Statistician* (1996), p. 49–56.
- [HR96] Christian HÖHN et Colin REEVES. "The crossover landscape for the onemax problem". In : *Proceedings of the 2nd Nordic Workshop on Genetic Algorithms*. Sous la dir. de Finland UNIVERSITY OF VAASA PRESS. J. Alander, 1996, p. 27–43.
- [MS96] Zbigniew MICHALEWICZ et Marc SCHOENAUER. "Evolutionary Algorithms for Constrained Parameter Optimization Problems". In : *Evol. Comput.* 4.1 (1996), p. 1–32.
- [REA96] Helge ROSÉ, Werner EBELING et Torsten ASSELMAYER. "The Density of States - A Measure of the Difficulty of Optimisation Problems". In : *Parallel Problem Solving from Nature*. 1996, p. 208–217.
- [Sta96] Peter F STADLER. "Landscapes and their correlation functions". In : *Journal of Mathematical chemistry* 20.1 (1996), p. 1–45.
- [Wei+96] Edward D WEINBERGER et al. *N_p completeness of kauffman's nk model, a tuneable rugged fitness landscape*. Rapp. tech. 96-02-003. Sante Fee Institute, 1996.
- [Gre95] John J. GREFENSTETTE. "Predictive Models Using Fitness Distributions of Genetic Operators". In : *Foundations of Genetic Algorithms* 3. Morgan Kaufmann, 1995, p. 139–161.

- [Jon95a] T. JONES. "Evolutionary Algorithms, Fitness Landscapes and Search". Thèse de doct. University of New Mexico, Albuquerque, 1995.
- [Jon95b] Terry JONES. *One operator, one landscape*. Rapp. tech. 95-02. Santa Fe Institute, 1995.
- [JF95] Terry JONES et Stephanie FORREST. "Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms." In : *ICGA*. T. 95. 1995, p. 184–192.
- [KE95] J. KENNEDY et R. EBERHART. "Particle swarm optimization". In : *Neural Networks, 1995. Proceedings., IEEE International Conference on*. T. 4. 1995, p. 1942–1948.
- [Sta95] Peter F. STADLER. "Towards a Theory of Landscapes". In : *Complex Systems and Binary Networks*. Sous la dir. de R. LOPEZ-PENA, R. CAPOVILLA, R. GARCIA-PELAYO, H. WAELBROECK et F. ZERTUCHE. T. 461. Berlin, New York : Springer Verlag, 1995, p. 77–163.
- [Tai95] Eric D TAILLARD. "Comparison of iterative searches for the quadratic assignment problem". In : *Location science* 3.2 (1995), p. 87–105.
- [WM95] David H WOLPERT et William G MACREADY. *No free lunch theorems for search*. Rapp. tech. Technical Report SFI-TR-95-02-010, Santa Fe Institute, 1995.
- [HGD94] Jeffrey HORN, David E. GOLDBERG et Kalyanmoy DEB. "Long Path Problems". In : *PPSN*. T. 866. LNCS. Springer, 1994, p. 149–158.
- [Sch+94] Peter SCHUSTER, Walter FONTANA, Peter F STADLER et Ivo L HOFACKER. "From sequences to shapes and back: a case study in RNA secondary structures". In : *Proceedings of the Royal Society of London B: Biological Sciences* 255.1344 (1994), p. 279–284.
- [SP94] Mandavilli SRINIVAS et Lalit M PATNAIK. "Adaptive probabilities of crossover and mutation in genetic algorithms". In : *Systems, Man and Cybernetics, IEEE Transactions on* 24.4 (1994), p. 656–667.
- [Kau93] S. A. KAUFFMAN. *The Origins of Order*. New York : Oxford University Press, 1993.
- [Dor92] Marco DORIGO. "Optimization, Learning and Natural Algorithms". Thèse de doct. Italy : Politecnico di Milano, 1992.
- [MW92] Keith E. MATHIAS et L. Darrell WHITLEY. "Genetic Operators, the Fitness Landscape and the Traveling Salesman Problem". In : *Parallel Problem Solving from Nature 2, PPSN-II, Brussels, Belgium, September 28-30, 1992*. 1992, p. 221–230.
- [MFH92] Melanie MITCHELL, Stephanie FORREST et John H HOLLAND. "The royal road for genetic algorithms: Fitness landscapes and GA performance". In : *Proceedings of the first european conference on artificial life* (1992), p. 245–254.
- [CD91] Hue Sun CHAN et Ken A. DILL. "Sequence space soup of proteins and copolymers". In : *The Journal of Chemical Physics* 95.5 (1991), p. 3775–3787.
- [MWS91] Bernard MANDERICK, Mark de WEGER et Piet SPIESSENS. "The genetic algorithm and the structure of the fitness landscape". In : *Proceedings of the fourth international conference on genetic algorithms*. 1991, p. 143–150.
- [Koz90] John R KOZA. *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*. Stanford University, Department of Computer Science, 1990.
- [Weig90] Edward WEINBERGER. "Correlated and uncorrelated fitness landscapes and how to tell the difference". In : *Biological cybernetics* 63.5 (1990), p. 325–336.

- [Gol89] David E. GOLDBERG. *Genetic algorithms in search, optimization, and machine learning*. 2. Addison-Wesley, Reading, MA, 1989.
- [NW88] George L NEMHAUSER et Laurence A WOLSEY. "Integer programming and combinatorial optimization". In : *Wiley, Chichester. GL Nemhauser, MWP Savelsbergh, GS Sigismondi (1992). Constraint Classification for Mixed Integer Programming Formulations. COAL Bulletin* 20 (1988), p. 8–12.
- [KL87] Stuart KAUFFMAN et Simon LEVIN. "Towards a general theory of adaptive walks on rugged landscapes". In : *Journal of theoretical Biology* 128.1 (1987), p. 11–45.
- [Ser87] P. SERAFINI. "Some considerations about computational complexity for multiobjective combinatorial problems". In : *Recent advances and historical development of vector optimization*. T. 294. Lecture Notes in Economics and Mathematical Systems. Springer, 1987, p. 222–232.
- [Glo86] Fred GLOVER. "Future paths for integer programming and links to artificial intelligence". In : *Computers & operations research* 13.5 (1986), p. 533–549.
- [Gre86] John J GREFENSTETTE. "Optimization of control parameters for genetic algorithms". In : *Systems, Man and Cybernetics, IEEE Transactions on* 16.1 (1986), p. 122–128.
- [Par86] Giorgio PARISI. "Spin glasses theory". In : *Physica A: Statistical Mechanics and its Applications* 140.1–2 (1986), p. 312–318.
- [SW84] Frank H STILLINGER et Thomas A WEBER. "Packing structures and transitions in liquids and solids". In : *Science(Washington, DC)* 225.4666 (1984), p. 983–9.
- [KGV83] S. KIRKPATRICK, C. D. GELATT et M. P. VECCHI. "Optimization by simulated annealing". In : *Science* 220.4598 (1983), p. 671–680.
- [Mez82] Paul G. MEZEY. "Topology of energy hypersurfaces". In : *Theoretica chimica acta* 62.2 (1982), p. 133–161.
- [AN79] Y. P. ANEJA et K. P. K. NAIR. "Bicriteria Transportation Problem". In : *Management Science* 25.1 (1979).
- [GJS76] Michael R GAREY, David S JOHNSON et Ravi SETHI. "The complexity of flowshop and jobshop scheduling". In : *Mathematics of operations research* 1.2 (1976), p. 117–129.
- [Ric76] John R. RICE. "The Algorithm Selection Problem." In : *Advances in Computers* 15 (1976), p. 65–118.
- [Hol75] John H HOLLAND. "Adaptation in natural and artificial system". In : *Ann Arbor, University of Michigan Press* (1975).
- [Rec73] Ingo RECHENBERG. "Evolutionsstrategie – Optimierung technischer Systeme nach Prinzipien der biologischen Evolution". Thèse de doct. Technical University of Berlin, 1973.
- [KB57] T. C. KOOPMANS et M. BECKMANN. "Assignment Problems and the Location of Economic Activities". In : *Econometrica* 25.1 (1957), p. 53–76.
- [Wri32] S. WRIGHT. "The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution". In : *Proceedings of the Sixth International Congress of Genetics* 1. 1932, p. 356–366.